Summary of Dose-Response Modeling for Developmental Toxicity Studies

Daniel L. Hunt
Shesh N. Rai
Chin-Shang Li

Abstract

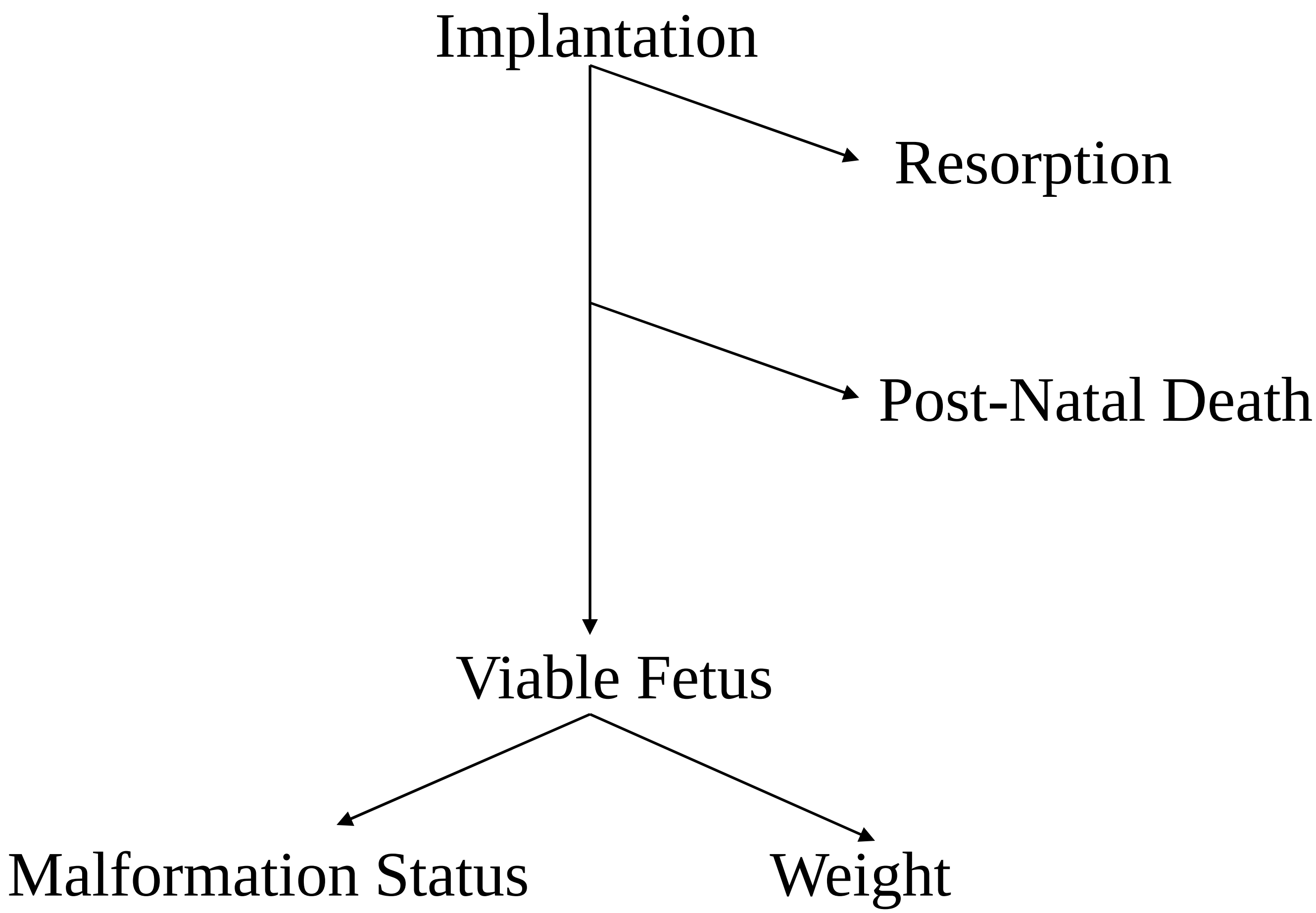
- Developmental toxicity studies (DTS) are an important area in the field of toxicology. In a DTS, fetal litters are indirectly exposed to various levels of non-carcinogenic toxic substances through direct exposure to the host animals. Endpoints that are recorded in these studies include fetal weight and length, as well as indicators of abnormality and death. Endpoints are then measured to determine litter responses, which include average weight, malformation and death rate. The dose-response pattern in these studies typically appears to exhibit at least the existence of a threshold effect. The threshold dose-response (DR) model is the default model for non-carcinogenic risk assessment, according to the USEPA, and is encouraged by the agency for the use in the risk assessment process. Several statistical models are proposed to estimate the threshold dose and to account for other important aspects of the developmental toxicity study. Use of these models to different applications will be summarized. The advantages and disadvantages of these models, and the comparison to other alternative models are discussed. We, also, summarize potentials for future research in this field.

- *Keywords: Developmental toxicity, Dose-group variability, Estimation, Splines, Threshold*

Developmental Toxicity Study (DTS)

- Randomly assign pregnant animals to exposure levels of toxin
- Outcomes measured on litters (fetuses)
- Fetal endpoints of interest:
  - Live endpoints
    - body weight and length
    - structural malformations
  - Death enpoints
    - Resorption
    - post-natal death

Outcomes in a DTS



Modeling of Data

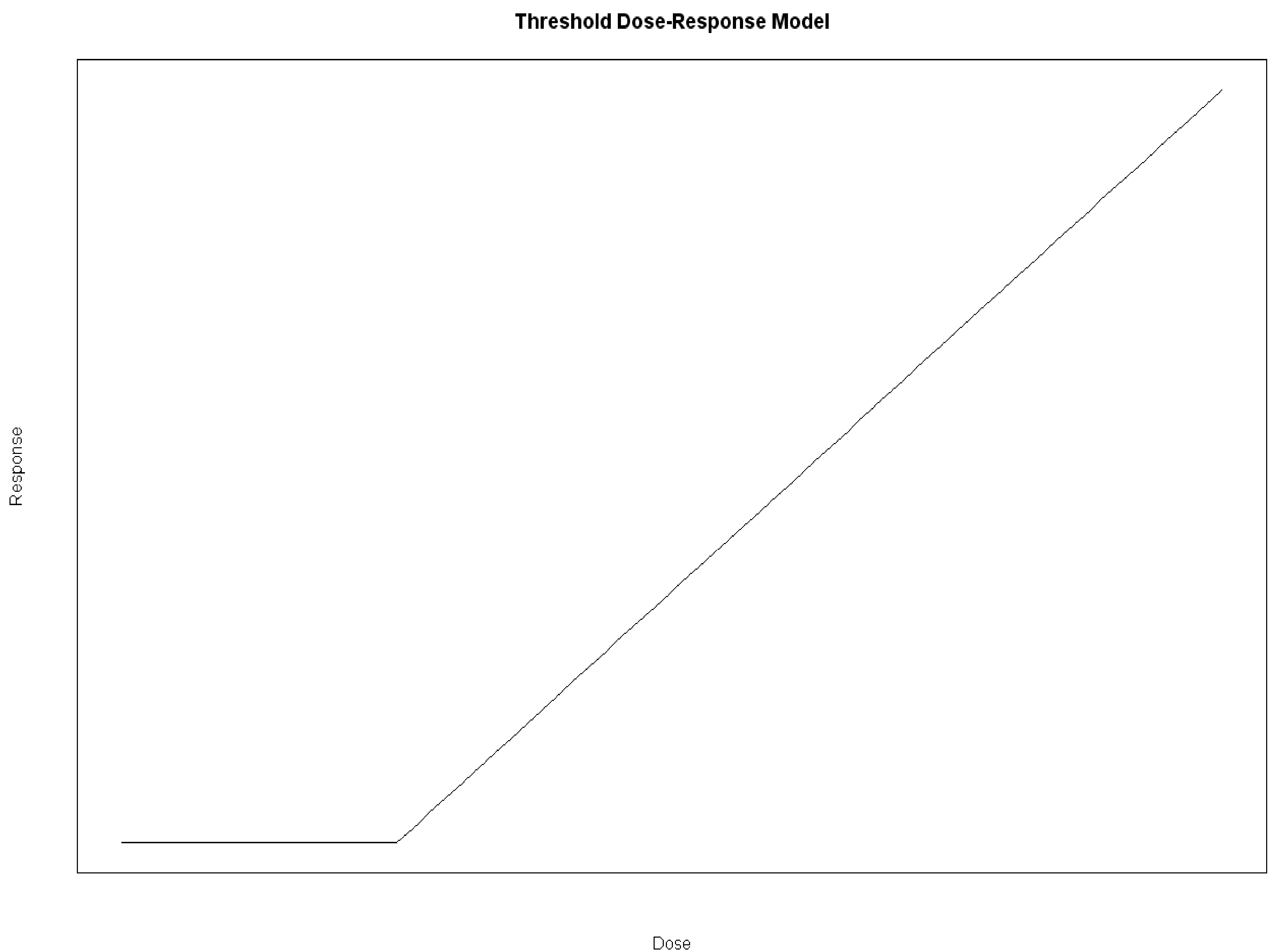
- Equate implantation site number to litter size
- Categorize malformations,resorptions,and post-natal deaths together as *adverse events*
- Measure outcome per litter as *proportion of adversely affected fetuses*
- Predict dose-response relationship by modeling P(d)
  - Dose:level of toxic substance (*d*)
  - Response rate: probability toxic response (*P*)

Threshold

- Definition: *The largest (non-zero) dose level which yields a toxic response that is equivalent to response at the the control level*

- 1991 USEPA Guidelines for Developmental Toxicity Risk Assessment:
  - *"…In general, a threshold is assumed for the dose-response curve for agents that produce developmental toxicity…"*

- USEPA models of non-carcinogen studies
  - NOAEL (no-observed-adverse-effects-level):*highest experimental dose at which response not (statistically) different from control*
  - Benchmark dose: *lower statistical confidence limit for dose corresponding to specified increased level of adverse effect over background level (excess risk)*


Threshold Dose-Response Model
Response
Dose

Functional Form of Threshold DR Model

$$P(d) = F\left[\theta_0 + \theta_1(d-\tau)\times I(d>\tau)\right]$$

OR

$$P(d) = \begin{cases} F\left(\theta_0\right), & d \le \tau \\ F\left[\theta_0 + \theta_1(d-\tau)\right], & d > \tau \end{cases}$$

Model Properties

- Continuous, piecewise function
- Figure slide
  - F would simply be an identity function, appropriate for continuous endpoint, such as weight; probability P could be replaced with W, for weight
- Discrete, binary endpoints
  - F could be logistic or probit function
  - P(d) is the probability of response at dose level d

Linear Spline Model

- Although above-threshold model may not be, below-threshold model is linear
- However, it is limited to only one pattern
- To make more robust, one could fit a linear spline model in lieu of the threshold model
- Spline model gives more flexibility in being able to accomodate below-threshold patterns

Polynomial Regression B-spline model

- Theory: DeBoor (2001)
- A regression spline of order m (degree m-1) and k interior knots yields a function of the form

$$s(d;\boldsymbol{\theta},\boldsymbol{\varepsilon}) = \sum_{i=1}^{m+k} \theta_i B_i(d;\boldsymbol{\varepsilon})$$

$\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_k)'$ is the set of interior knots

$\{B_i(d;\boldsymbol{\varepsilon}) : i = 1, \ldots, m+k\}$ is the set of (order m) B-splines

$\boldsymbol{\theta} = (\theta_1, \ldots, \theta_{m+k})'$ are the regression coefficients of the B-splines

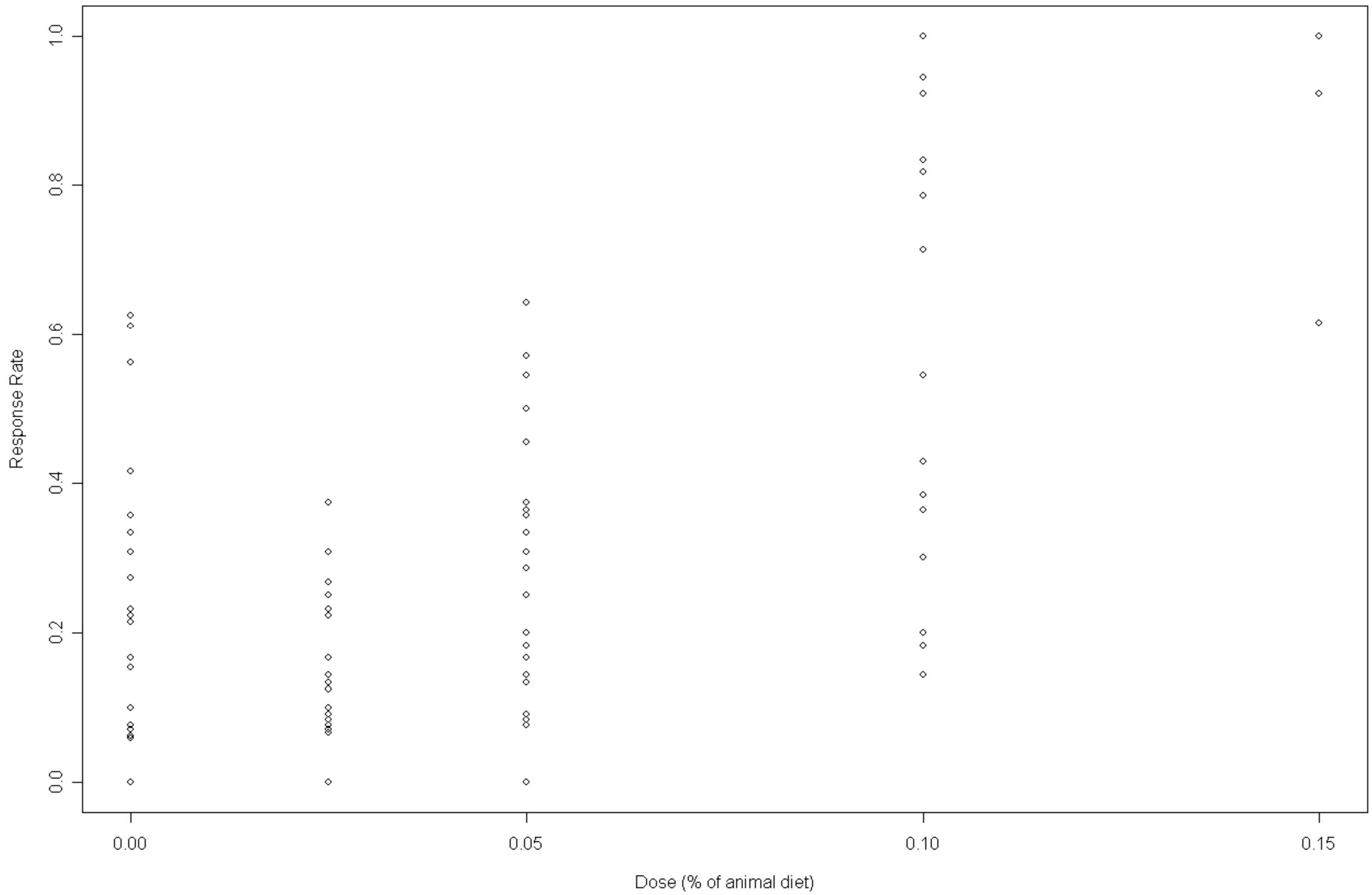
Spline model for developmental study

- Linear (m=2) B-spline with k=1 interior knot ε

$$P_1(d) = G\left[\sum_{i=1}^{3} \theta_i B_i(d;\varepsilon)\right]$$

- G is appropriate function to accommodate the type of data (identity, logistic, probit, etc.)

Example DTS: DEHP administered to pregnant rats (Tyl et al., 1983)



Response Variation

- As noted in the prior slide, there is noticeable degree of variability in the litter responses
- This variability should be accounted for in the modeling process
- Simple modification to the fixed effects models P(d) and $P_1(d)$:
- Add random effect to model this response variability

Models modified

$$P(d) = F\left[\theta_0 + \theta_1(d-\tau)\times I(d>\tau) + \sigma Z\right]$$

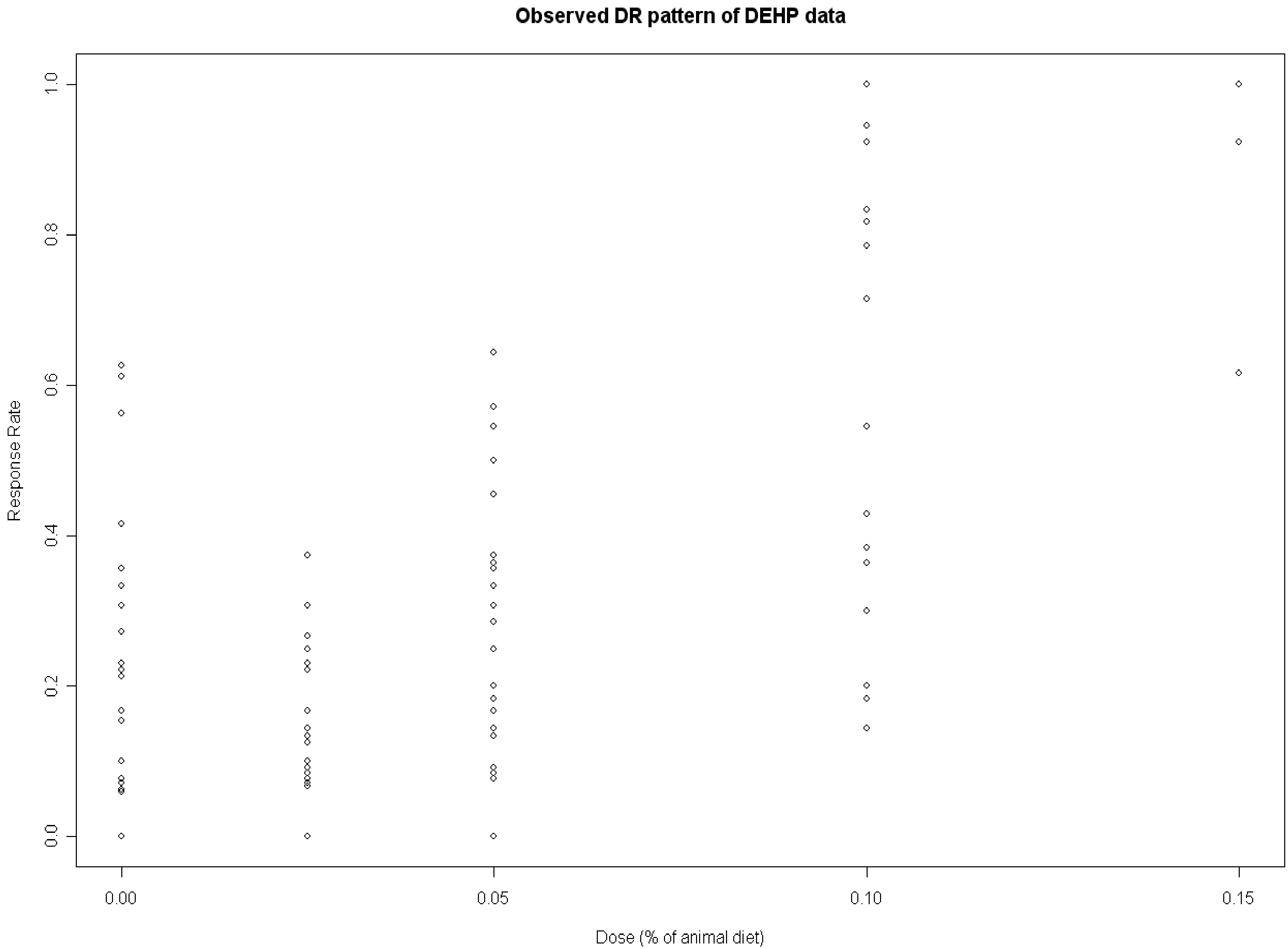
$$P_1(d) = G\left[\sum_{i=1}^{3}\theta_i B_i(d;\varepsilon) + \sigma Z\right]$$

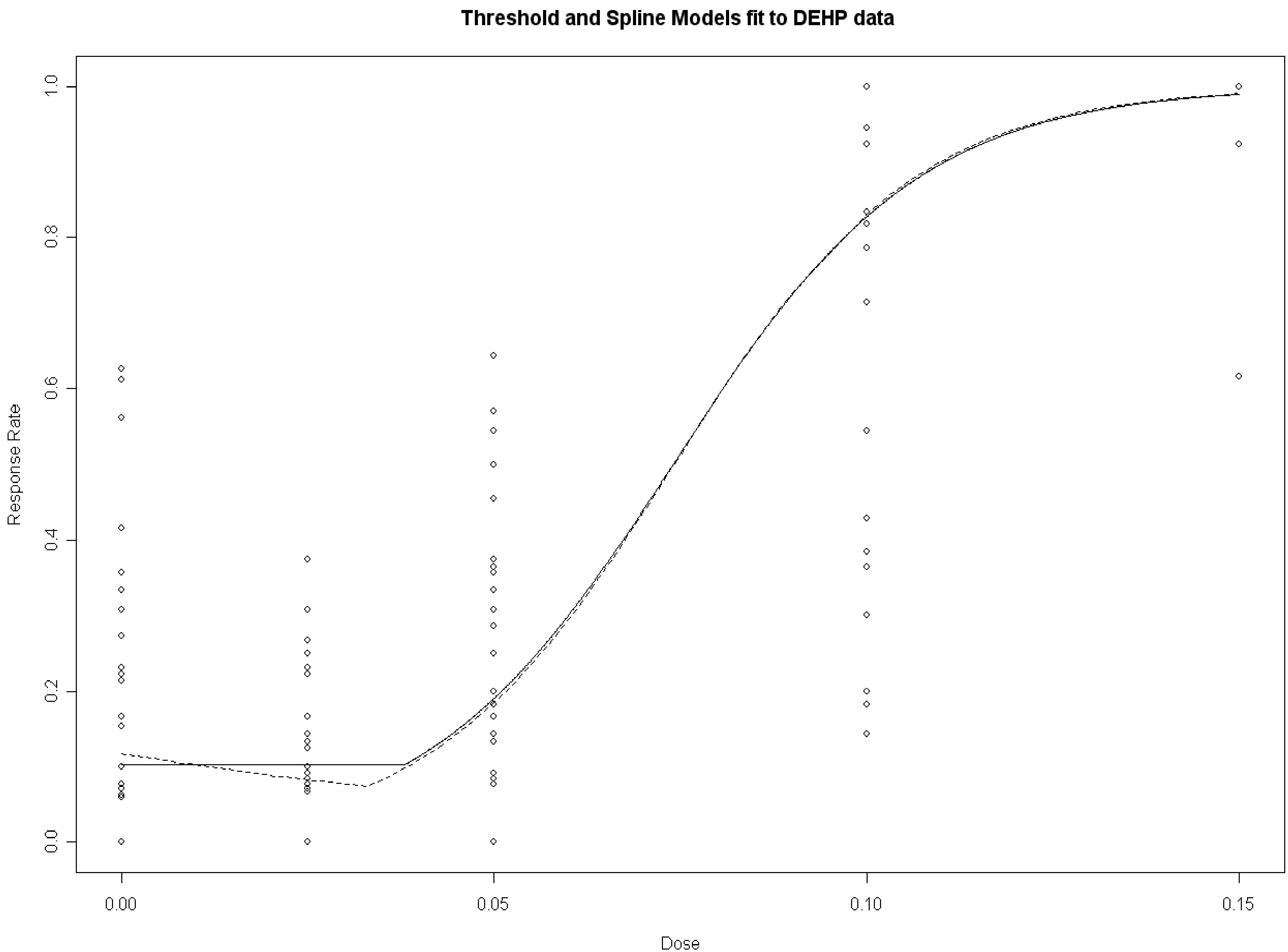
Add random component

The random component σZ is such that Z~N(0,1) and Var(Z)=σ²

DEHP data (Tyl et al., 1983)

Dose of DEHP	Number Litters	Number Fetuses	Average Litter Size	Number Affected Fetuses	Proportion Affected Fetuses
0.00	30	396	13.2	75	.189
0.025	26	320	12.3	37	.116
0.05	26	319	12.3	80	.251
0.10	24	276	11.5	192	.696
0.15	25	308	12.3	302	.981


Observed DR pattern of DEHP data
Response Rate
0.0
0.2
0.4
0.6
0.8
1.0
0.00
0.05
0.10
0.15
Dose (% of animal diet)



Threshold and Spline Models fit to DEHP data
Response Rate
0.0
0.2
0.4
0.6
0.8
1.0
0.00
0.05
0.10
0.15
Dose

Comparing Models

- From deBoor (2001), alternate representation of the spline model
  - $s(d,\theta,\varepsilon)=\theta_1+\theta_2 d+\theta_3(d-\varepsilon)_+$

- If $\theta_2=0$, the model becomes
  - $s_1(d,\theta,\varepsilon)=\theta_1+\theta_3(d-\varepsilon)_+$

- which is equivalent to the threshold model

Likelihood Ratio (LR) testing

- Dealing with parametric, nested models, hence LR test for significance

- Result
  - p-value =0.42
  - Indicative of non-significance (of the spline model)

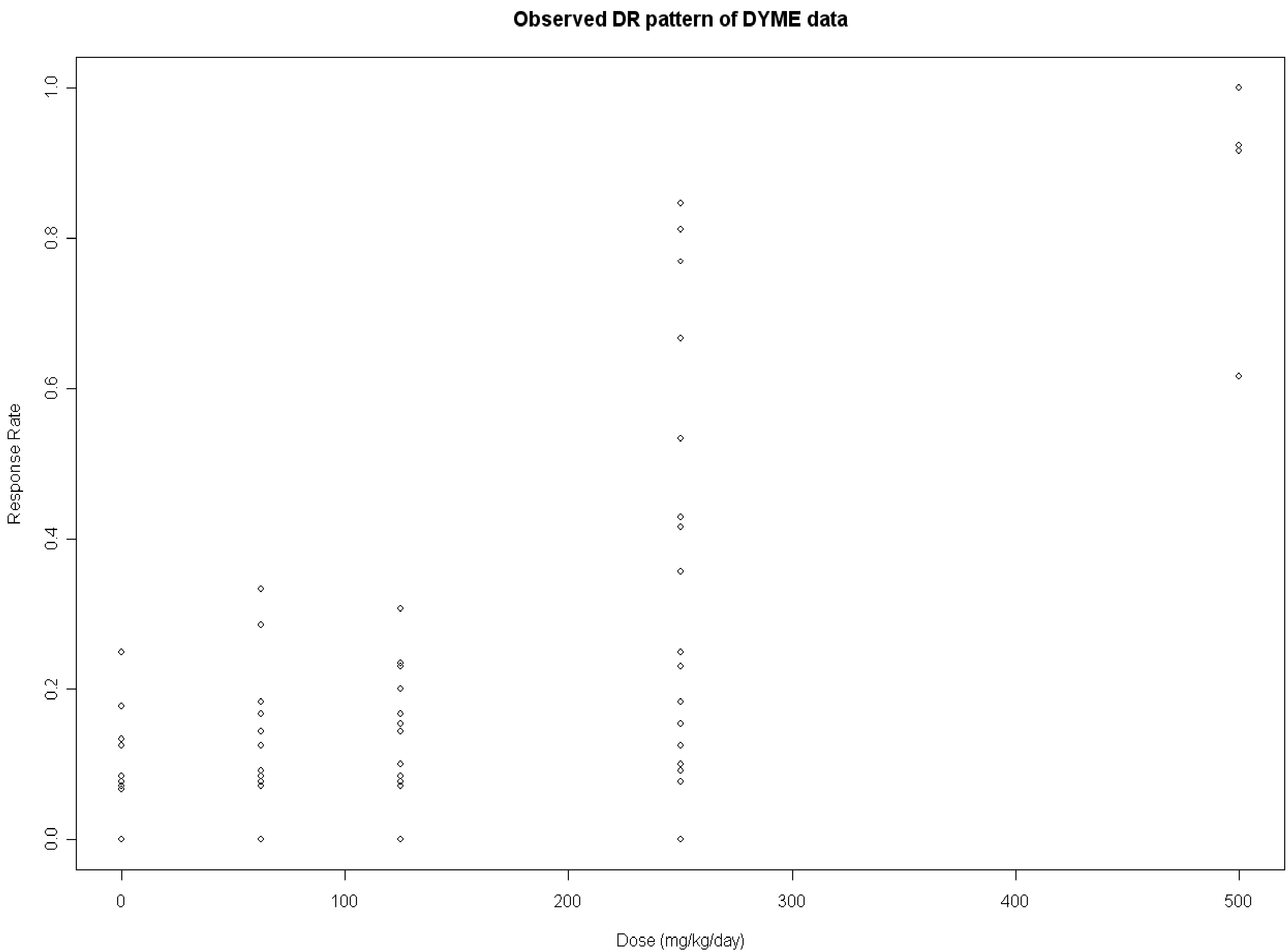
- Note the limited # dose groups below threshold

Factors to Improve Study Power

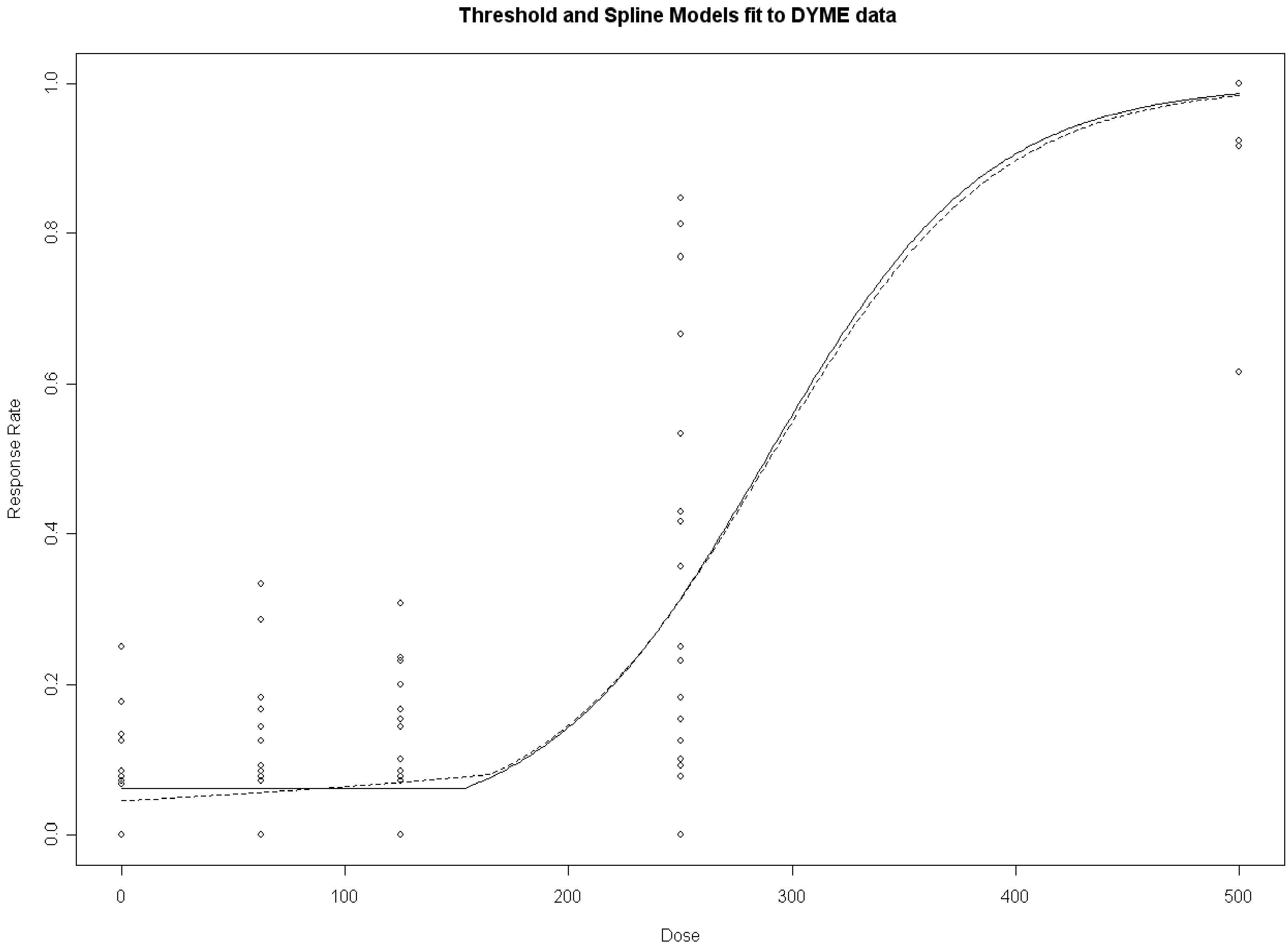
- To better estimate threshold & other effects:
- More dose groups (below threshold)
- Larger sample size (at dose groups below threshold)
- Adequate dose spacing
- Authors
  - Sielken and Stevenson, 1998
  - Teeguarden *et al.*, 2000
  - Hunt (2002)

DYME data (Price et al., 1987)

Dose of DYME (mg/kg/day)	Number of Litters	Number of Fetuses	Average Litter Size	Number of Affected Fetuses	Proportion of Affected Fetuses
0	21	297	14.1	17	.057
62.5	20	242	12.1	20	.083
125	24	312	13.0	35	.112
250	23	299	13.0	102	.341
500	23	285	12.4	277	.972



Observed DR pattern of DYME data
Response Rate
0.0
0.2
0.4
0.6
0.8
1.0
0
100
200
300
400
500
Dose (mg/kg/day)


Threshold and Spline Models fit to DYME data
Response Rate
Dose
0.0
0.2
0.4
0.6
0.8
1.0
0
100
200
300
400
500

LR testing: DYME data

- Significance of spline over threshold model
- p-value=0.091
- Much lower than for DEHP data
- Reasons? Now, 3 dose groups below threshold instead of 2
- Conclusion? More dose groups needed to adequately assess significance

Considerations when using Spline Approach

- In DTS, there are few dose groups
- The number of noticeable changepoints in the pattern of the DR data will be few (if not non-existent)
- Spline order and #knots should be based on these factors
- Current design of DTS seems to accommodate this

Advantages of these models

- Direct estimation of threshold or changepoint
- Inclusion of random effects into DR function to counter models such as the BB; facilitates estimation by allowing common methods to find parameter estimates
- Direct modeling of DR pattern

Disadvantages

- Threshold can be difficult to estimate
- Current design of DTS accomodates use of threshold and low-order spline model; however, modified design DTS may require more complex models; can lead to estimation difficulties

Future Work

- Modify models to include multiple σ parms.
- Easier method for estimating SEs, e.g., bootstrapping
- Higher order spline (perhaps quadratic) model
- More general spline model (higher order, more knots)

References

- Chen JJ and Kodell RL. 1989. Quantitative risk assessment for teratological effects.
- J Am Stat Assoc 84: 966-971
- Collett D. 1991. Modelling Binary Data. CRC Press, Boca Raton, FL
- Cox C. 1987. Threshold dose-response models in toxicology. Biometrics 43**:** 511-523
- Crump KS. 1984. A new method for determining allowable daily intakes. Fundam Appl Toxicol 4: 854-871
- De Boor C. 2001. A Practical Guide to Splines (revised edition). Sprnger-Verlag, New York
- Doull J, Cattley R, Elcombe C, Lake BG, Swenberg J, Wilkinson C, Williams G, and van Gemert M. 1999. A cancer risk assessment of di(2-ethylhexyl)phthalate: application of the new USEPA risk assessment guidelines. Regul Toxicol Pharmacol 29**:** 27-357
- Haseman JK and Kupper LL. 1979. Analysis of dichotomous response data from certain toxicological experiments. Biometrics 35: 281-293
- Hunt D. 2002. Dose and litter allocations in the design of teratological studies for detecting hormesis. Teratology 66: 309-314

References, cont.

- Hunt D and Rai SN. 2003. A threshold dose-response model with random effects in teratological experiments. Commun Stat Theory Meth 32: 1439-1457
- Hunt DL and Bowman D. 2004. A parametric model for detecting hormetic effects in developmental toxicity studies. Risk Anal 24: 65-72
- Hunt DL and Rai SN. 2008. Interlitter response variability in a threshold dose-response model. In Press Commun Stat Theory Meth
- Kupper LL, Portier C, Hogan MD, and Yamamoto E. 1986. The impact of litter effects on dose-response modeling in teratology. Biometrics 42**:** 85-98
- Li CS and Hunt D. 2004. Regression splines for threshold selection with application to a random-effects logistic dose-response model. Comput Stat Data Anal 46: 1-9
- McCullagh P and Nelder JA. Generalized Linear Models (second edition). Chapman and Hall: London
- Paul SR. 1982. Analysis of proportions of affected fetuses in teratological experiments. Biometrics 38: 361-370
- Price CJ, Kimmel CA, George JD, and Marr MC. 1987. The developmental toxicity of diethylene glycol dimethyl ether in mice. Fundam Appl Toxicol 81**:** 113-127

References, cont.

- Ryan LM. 2000. Statistical issues in toxicology. J Am Stat Assoc 95: 304-308
- Schwartz PF, Gennings C, and Chinchilli VM. 1995. Threshold models for combination data from reproductive and developmental experiments. J Am Stat Assoc 90**:** 862-870
- Sielken RL and Stevenson DE. 1998. Some implications for quantitative risk assessment if hormesis exists. Hum ExpToxicol 17: 259-262
- Teeguarden JG, Dragan Y, and Pitot HC. 2000. Hazard assessment of chemical carcinogens: the impact of hormesis. J Appl Toxicol 20: 113-120
- Tyl RW, Jones-Price C, Marr MC, and Kimmel CA. 1983. Teratological evaluation of diethylhexyl phthalate (CAS No. 117-81-7) in CD-1 mice. Final Study Report for NCTR/NTP Contract NO. 222-80-2031 9(c). NTIS NO PB85105674.
- National Technical Information Service, Springfield, VA
- USEPA. 1986. Guidelines for mutagenicity risk assessment. Federal Register 51**:** 34006-34012
- USEPA. 1991. Guidelines for developmental toxicity risk assessment. Federal Register 56**:** 63798-63826

References, cont.

- USEPA. 1992. Guidelines for exposure assessment. Federal Register 57**:** 22888-22938
- USEPA. 1996. Guidelines reproductive toxicity risk assessment. Federal Register 61**:** 56274-56322
- USEPA. 1998. Guidelines neurotoxicity risk assessment. Federal Register 63**:** 26926-26954
- USEPA. 2005. Guidelines carcinogen risk assessment. Federal Register 70**:** 17765-17817
- Williams DA. 1975. The analysis of binary responses from toxicological experiments
- involving reproduction and teratogenicity. Biometrics 31: 949-952

Biomedical References for Applications of the Spline Model

- Molinari N, Daures J, Durand J. 2001. Regression splines for threshold selection in survival data analysis. Stat.Med. 20: 237-247.

- Bessaoud F, Daures JP, Molinari N. 2005. Free knot splines for logistic models and threshold selection. Computer Methods and Programs in Biomedicine 77: 1-9.

QUESTIONS???