Combining Multiple Comparisons and Modeling Techniques in Dose Response Studies

Frank Bretz and **José Pinheiro**

Biostatistics, Novartis Pharmaceuticals

frank.bretz@novartis.com, jose.pinheiro@novartis.com

BASS XII, November 07-11, 2005, Savannah GA

F. Bretz, J. C. Pinheiro, and M. Branson (2005) Combining Multiple Comparisons and Modeling Techniques in Dose-Response Studies. *Biometrics*, **61**: 738–748.

Outline

- Motivation
- Classical approaches: multiple comparisons and modeling
- Example: a dose finding study
- Common dose-response models and their evaluation
- Dose finding example re-visited
- Combining multiple comparisons and modeling
- Conclusions

Motivation

- Conducting confirmatory phase III trails is expensive
- Dose-finding trials are a critical component of decision process
- Identifying right dose is a key goal of clinical development:
  - too high a dose can result in unacceptable toxicity
  - too low a dose decreases chance of showing efficacy

Motivation

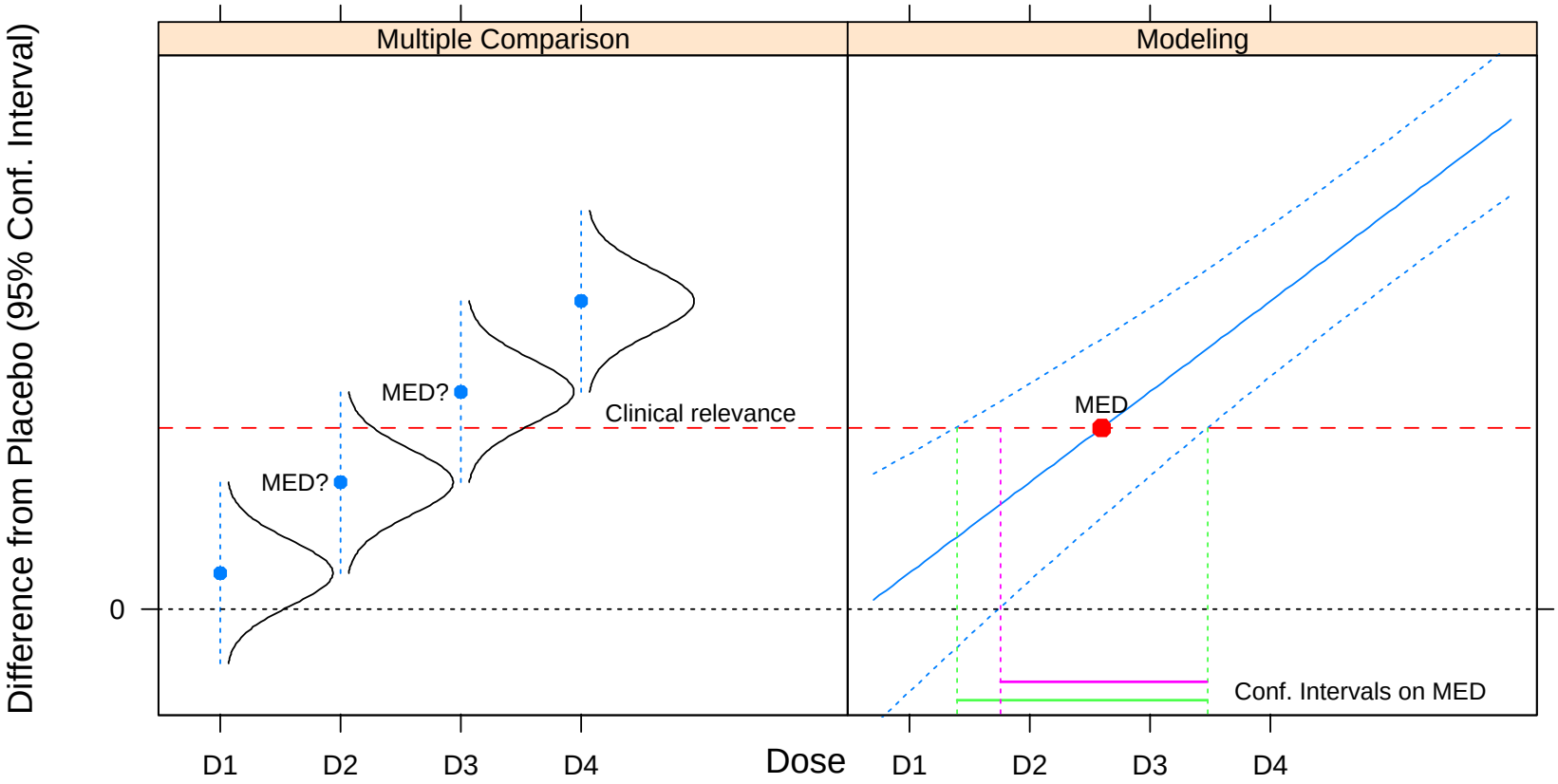
- FDA reported that 20% of drugs approved between 1980 and 1999 had initial dose changed by $> 33\%$ (80% decreased)
- FDA/DIA meeting on "Good Dose Response", 10/2004: Failure rate of current Phase III trials is 50%; ten years ago, 20% $\Rightarrow$ improper dose selection in Phase II blamed for many such failed trials
- At same meeting, it was agreed that better understanding and evaluation of dose-response should be critical part of Phase II trials

Motivation (cont.'d)

- Two main goals in phase II studies:
  - proof-of-concept (PoC) – any evidence of treatment effect
  - dose-selection – which dose(s) to take into phase III?
    minimum effective dose (MED), maximum safe dose (MSD)
- ICH-E4: Purpose of dose-response information is to find the

  Smallest dose with a discernible useful effect
- Emphasis is placed on identifying or estimating the MED
  - Assurance that a desired effect size is plausible
- Analysis strategies categorized into two broad classes:
  multiple comparisons (MCP) of contrasts between doses and
  modeling of dose response relationship

Finding the MED – an illustration



- Either D2 or D3 could be chosen as the MED in the MCP case
- Modeling is more flexible, but requires additional assumptions

Objectives of this presentation

- Discuss the modeling approach to dose finding: interpretation of model parameters, initial estimates, and estimation of parameters
- Methods for estimating target doses of interest (e.g., MED) from dose response models
- Introduce a unified approach for more efficient and robust dose finding statistical analyses, based on a combination of multiple comparison and modeling ideas.

Multiple comparisons and Modeling approaches

Multiple comparisons (MCP)

- Uses contrasts between responses at different dose levels: dose treated as categorical variable
- Main goals: test PoC and obtain minimum efficacious dose MED or maximum safe dose MSD, while controlling FWER

Modeling

- When enough doses are present and some prior knowledge of the dose response profile is available, a parametric dose-response model can be used: $y = f(d, \theta) + \epsilon$
- Dose is treated as a continuous variable
- Estimation of MED and other target doses is done by inverse regression (identify dose achieving a specific response)

Advantages of each approach

MCP

- Allows strong control of FWER
- Easy to implement and interpret
- Does not require much prior knowledge of dose response relationship: less sensitive to assumptions

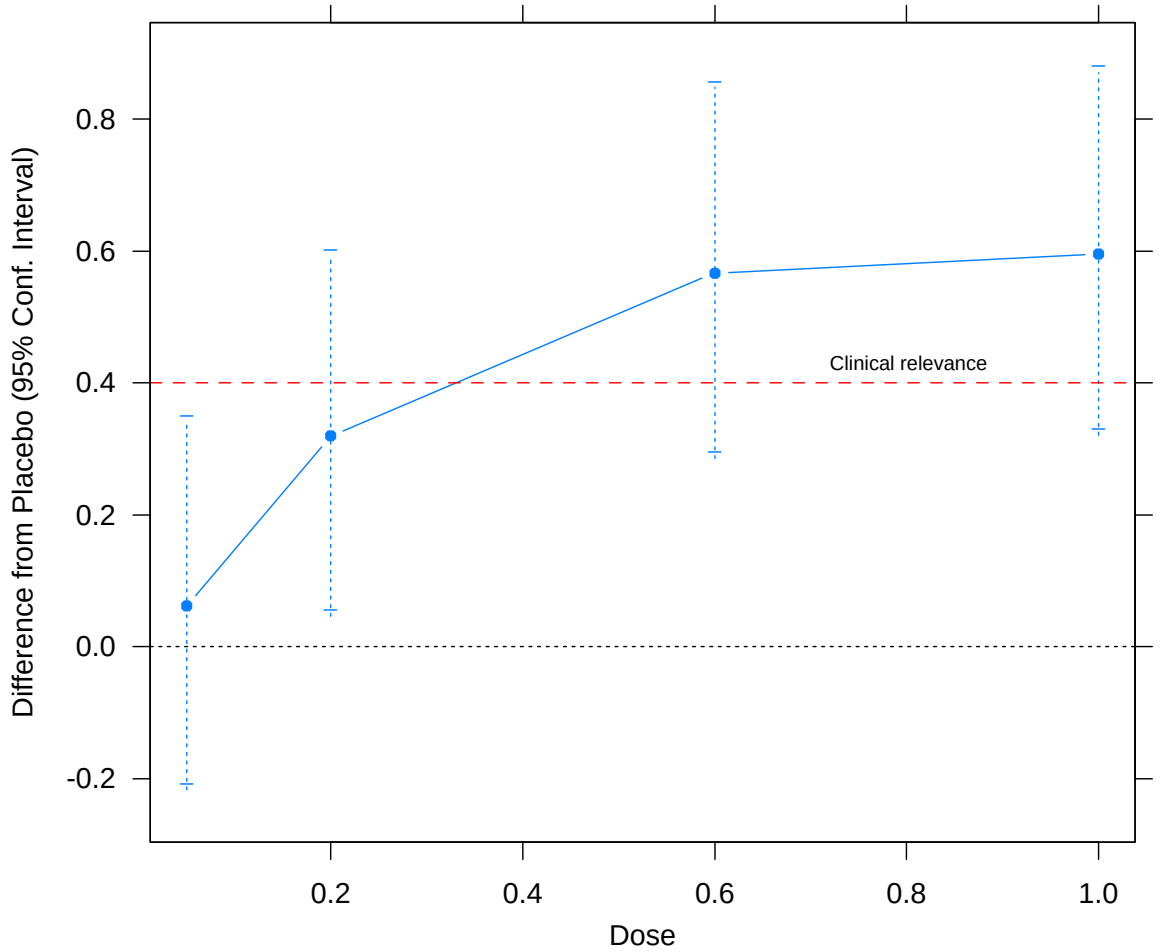
Modeling

- MED and other target doses: any dose in observed range
- Provides confidence intervals on estimated doses
- Easy to include requirements on clinical relevance
- Better understanding of dose-response relationship: useful for planning future studies and simulations

Example: a phase II dose-finding study

- Randomized double-blind parallel group trial with about 250 patients equally allocated to placebo or one of four active doses: 0.05, 0.2, 0.6, or 1
- Normally distributed, homoscedastic primary endpoint
- Step-down procedure (hierarchical) used to preserve FWER at 5% two-sided level
- All doses were well-tolerated: MSD ≥ 1

Example (cont.'d)



Which dose should be considered MED?

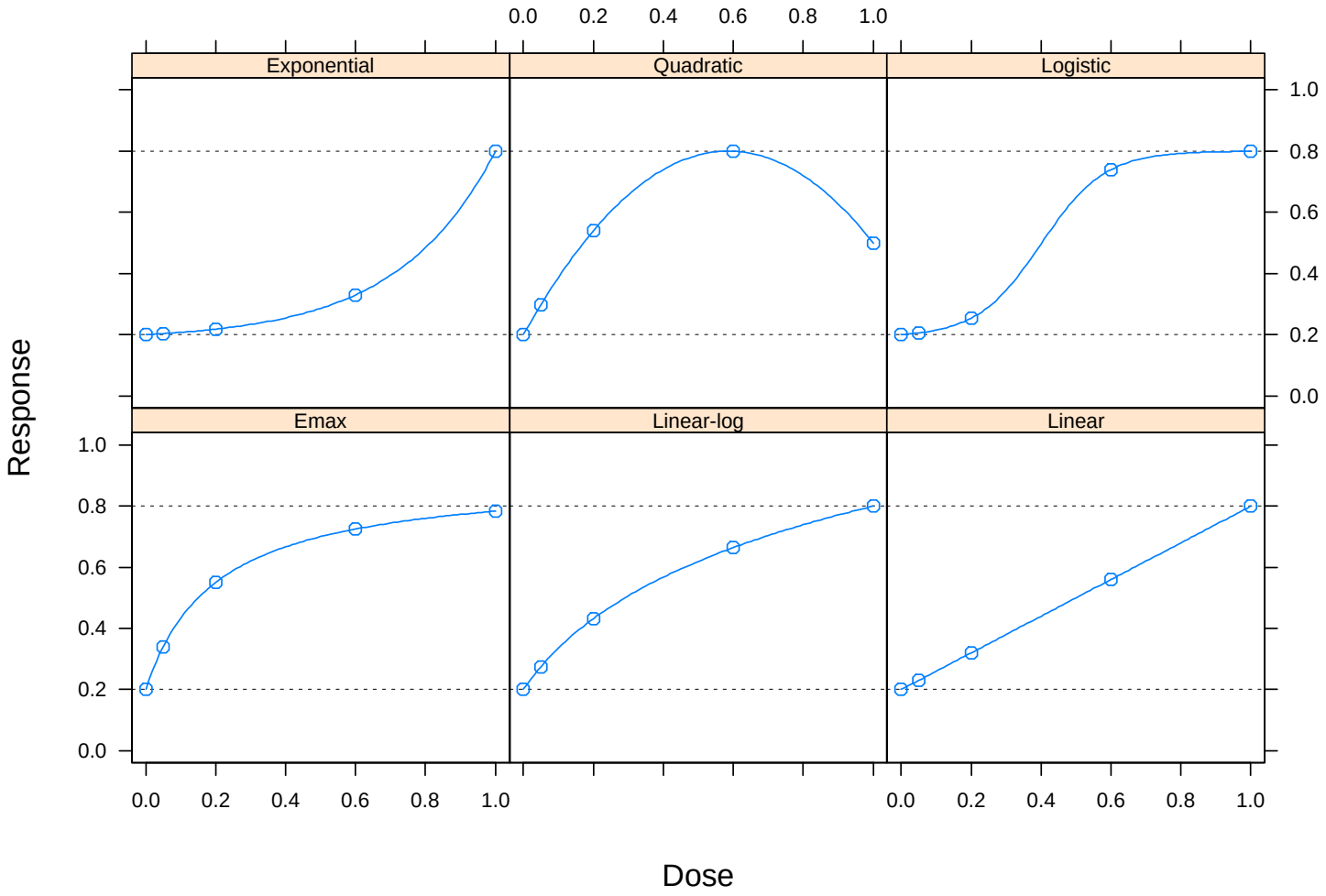
Dose-response models

- Framework considered: response Y (efficacy or safety) observed for parallel groups corresponding to ordered doses $d_1 < d_2 < \cdots < d_k$ (d_1 typically placebo)
- Methods can be extended to repeated measures data, such as cross-over designs, factorial drug combinations, and other more complex trial designs
- General dose-response model for parallel group (one-way) case

$$Y_{ij} = f(d_i, \boldsymbol{\theta}) + \epsilon_{ij}, \quad \epsilon_{ij} \overset{\text{ind}}{\sim} \mathcal{N}(0, \sigma^2)$$

- Can be often be expressed as $f(d, \boldsymbol{\theta}) = \theta_0 + \theta_1 f^0(d, \boldsymbol{\theta}^0)$, with f^0 representing standardized model; initial values only required for $\boldsymbol{\theta}^0$

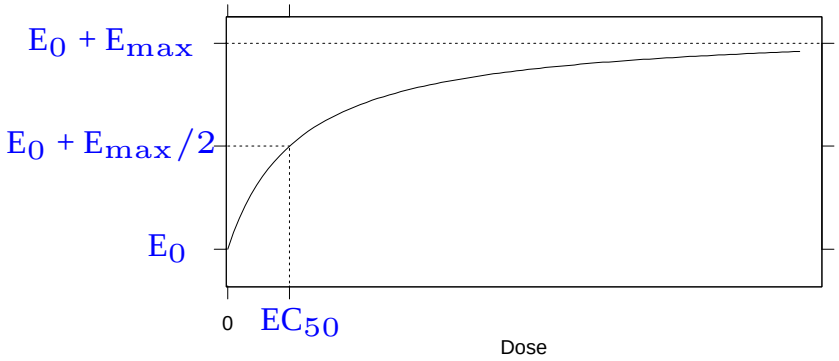
Some typical dose-response models



Cover a wide range of possible shapes, most monotonic

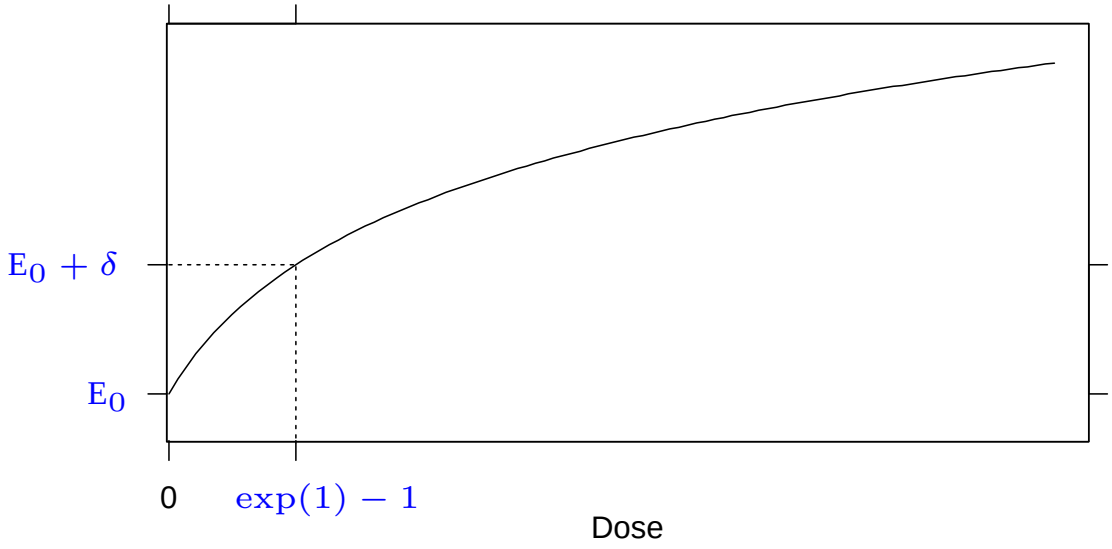
$\mathbf{E_{max}}$ model

- $f(d, \boldsymbol{\theta}) = E_0 + E_{\max} d/(EC_{50} + d)$
- Standardized form: $f^0(d, \boldsymbol{\theta}^0) = d/(EC_{50} + d)$
  $\Longrightarrow$ % of max change
- E_0: basal effect (at $d = 0$)
- $E_{\max}$: max change in effect
- EC_{50}: dose at half of max change
- Initial estimate for EC_{50}: from percentage of maximum effect p^* associated with dose d^*: $\widehat{EC}_{50} = d^*(1 - p^*)/p^*$



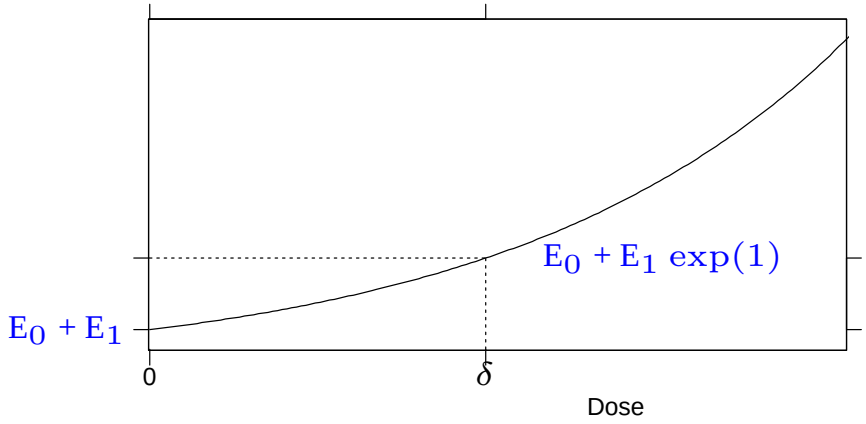
Linear in log-dose model

- $f(d, \boldsymbol{\theta}) = E_0 + \delta \log(d + c)$
- Standardized form: $f^0(d, \boldsymbol{\theta}^0) = \log(d + 1)$
- E_0: basal effect (at $d = 0$)
- δ: log-dose slope
- No need for initial estimates



Exponential (power) model

- $f(d, \boldsymbol{\theta}) = E_0 + E_1 \exp(d/\delta)$
- Standardized form: $f^0(d, \boldsymbol{\theta}^0) = \exp(d/\delta)$
- Basal effect (at $d = 0$): $\mathrm{E}_0 + \mathrm{E}_1$
- δ determines rate of increase
- Initial estimate for δ: from percentage of effect p^* associated with dose d^*: $\widehat{\delta} = d^* / \log(1 + p^*)$



Quadratic model

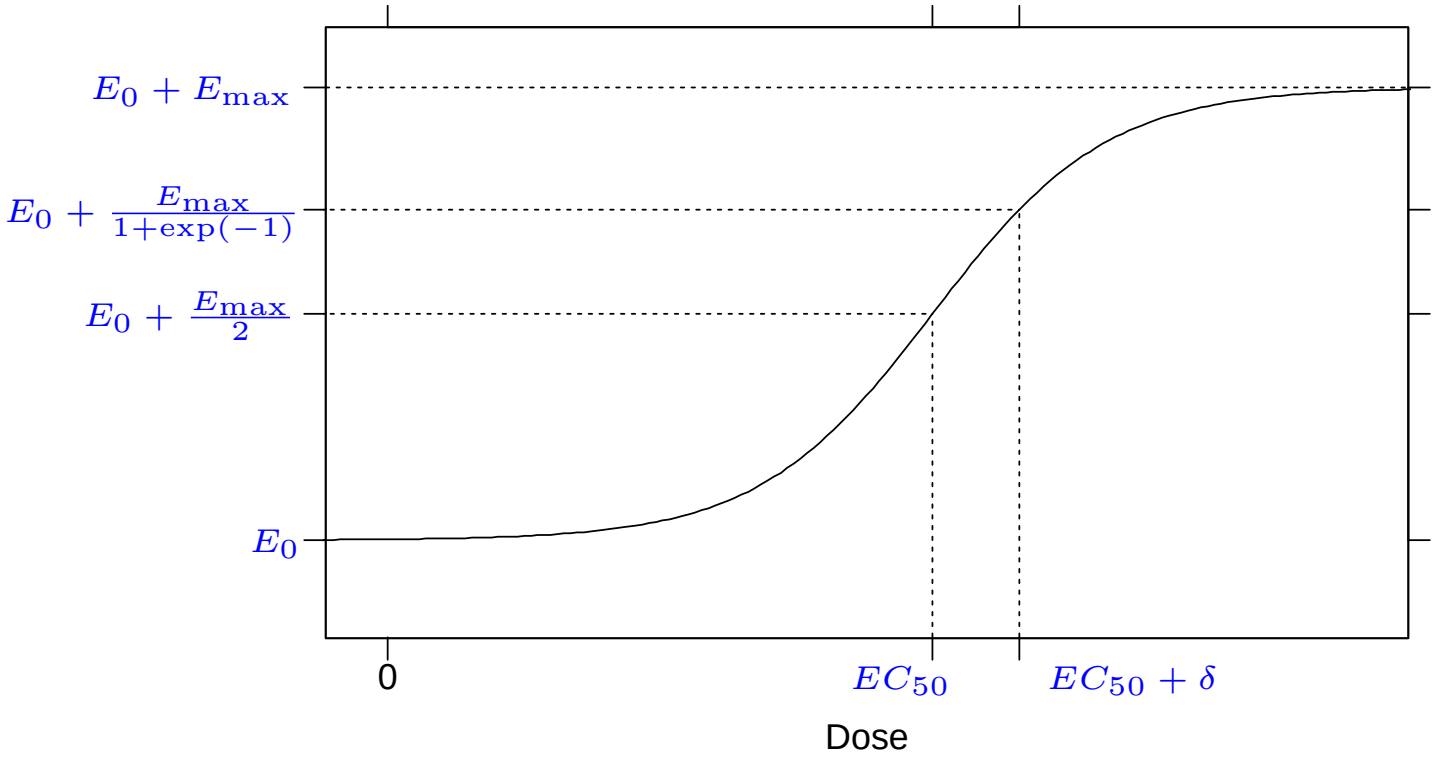
- $f(d, \boldsymbol{\theta}) = E_0 + \beta_1 d + \beta_2 d^2$
- $\beta_2 < 0 \Rightarrow$ umbrella (or inverted-U) shape; $\beta_2 > 0 \Rightarrow$ U-shape – will assume umbrella-shape
- Standardized form: $f^0(d, \boldsymbol{\theta}^0) = d + \delta d^2$, $\delta = \beta_2/|\beta_1|$
- Dose corresponding to max response: $d_{\max} = -1/2\delta$
- E_0 basal effect (at $d = 0$)
- Initial estimate for δ based on (d^*, p^*):

$$\widehat{\delta} = \begin{cases} -(1 - \sqrt{1 - p^*})/2d^*, & d^* < d_{\text{opt}} \\ -(1 + \sqrt{1 - p^*})/2d^*, & d^* \geq d_{\text{opt}} \end{cases}$$

Logistic model

- $f(d, \boldsymbol{\theta}) = E_0 + E_{\max} / \{1 + \exp[(EC_{50} - d)/\delta]\}$
  - E_0: basal effect (not placebo effect)
  - $E_{\max}$: max change from basal effect
  - EC_{50}: dose at 50% of max change
  - δ: controls rate of change
- Standardized form: $f^0(d, \boldsymbol{\theta}^0) = 1/\{1 + \exp[(EC_{50} - d)/\delta]\}$
  - represents percentage of maximum change

Logistic model (cont.'d)



Initial estimates require two points (d_1^*, p_1^*) and (d_2^*, p_2^*):

$$\widehat{\delta} = \frac{d_2^* - d_1^*}{\text{logit}(p_2^*) - \text{logit}(p_1^*)}, \quad \widehat{EC}_{50} = \frac{d_1^* \text{logit}(p_2^*) - d_2^* \text{logit}(p_1^*)}{\text{logit}(p_2^*) - \text{logit}(p_1^*)},$$

where $\text{logit}(p) = \log(p/(1-p))$

Estimating the MED

- For absolute clinically relevant difference Δ with respect to smallest dose d_1: $\text{MED} = \min_{d \in (d_1, d_k]} \{ f(d, \boldsymbol{\theta}) > f(d_1, \boldsymbol{\theta}) + \Delta \}$
- Let $p_d = f(d, \widehat{\boldsymbol{\theta}})$ denote predicted response at dose d, with corresponding confidence interval $[L_d, U_d]$
- Three different rules proposed for estimating MED:
  - $\widehat{\text{MED}}_1 = \min_{d \in (d_1, d_k]} \{ U_d > p_{d_1} + \Delta, L_d > p_{d_1} \}$
  - $\widehat{\text{MED}}_2 = \min_{d \in (d_1, d_k]} \{ p_d > p_{d_1} + \Delta, L_d > p_{d_1} \}$
  - $\widehat{\text{MED}}_3 = \min_{d \in (d_1, d_k]} \{ L_d > p_{d_1} + \Delta \}$
- By construction, $\widehat{\text{MED}}_1 \leq \widehat{\text{MED}}_2 \leq \widehat{\text{MED}}_3$
- Estimated MEDs may not exist for some, or all of the methods
- Different levels may be used for prediction bands (e.g., 60%, 80%) leading to possibly different MED estimates

Model fitting in phase II example

- Nonlinear models, such as $\mathrm{E}_{\max}$ and exponential, can be fitted in SAS with PROC NLIN; in S-PLUS or R with nls function
- Assume 50% of maximum effect attained at $d = 0.2$
- To get initial estimates for E_0 and $\mathrm{E}_{\max}$: conditional on

$$\widehat{EC}_{50} = 0.2(1 - 0.5)/0.5 = 0.2$$

  $\mathrm{E}_{\max}$ model is linear in $d/(d + 0.2)$:

```
> phaseII[1:2,]   # data
   response dose
1 -0.021950 0.05
2 -0.057114 0.00
> lm(response ~ I(dose/(dose+0.2)), phaseII)
Coefficients:
 (Intercept)  I(dose/(dose+0.2))
     0.28948             0.76854
```

$\Rightarrow \widehat{E}_0 = 0.29, \widehat{E}_{\max} = 0.77$

Model fitting in phase II example (cont.'d)

Fitting the E_{max} model:

```
> fmEmax <- nls(response ~ e0 + eMax*dose/(dose + ec50),
                data = phaseII,
                start = list(e0=0.29, eMax=0.77, ec50=0.2))
> summary(fmEmax)
. . .
Parameters:
       Value Std. Error t value
  e0 0.31092   0.090122  3.4500
eMax 0.82969   0.225388  3.6812
ec50 0.29685   0.275299  1.0783
. . .
```

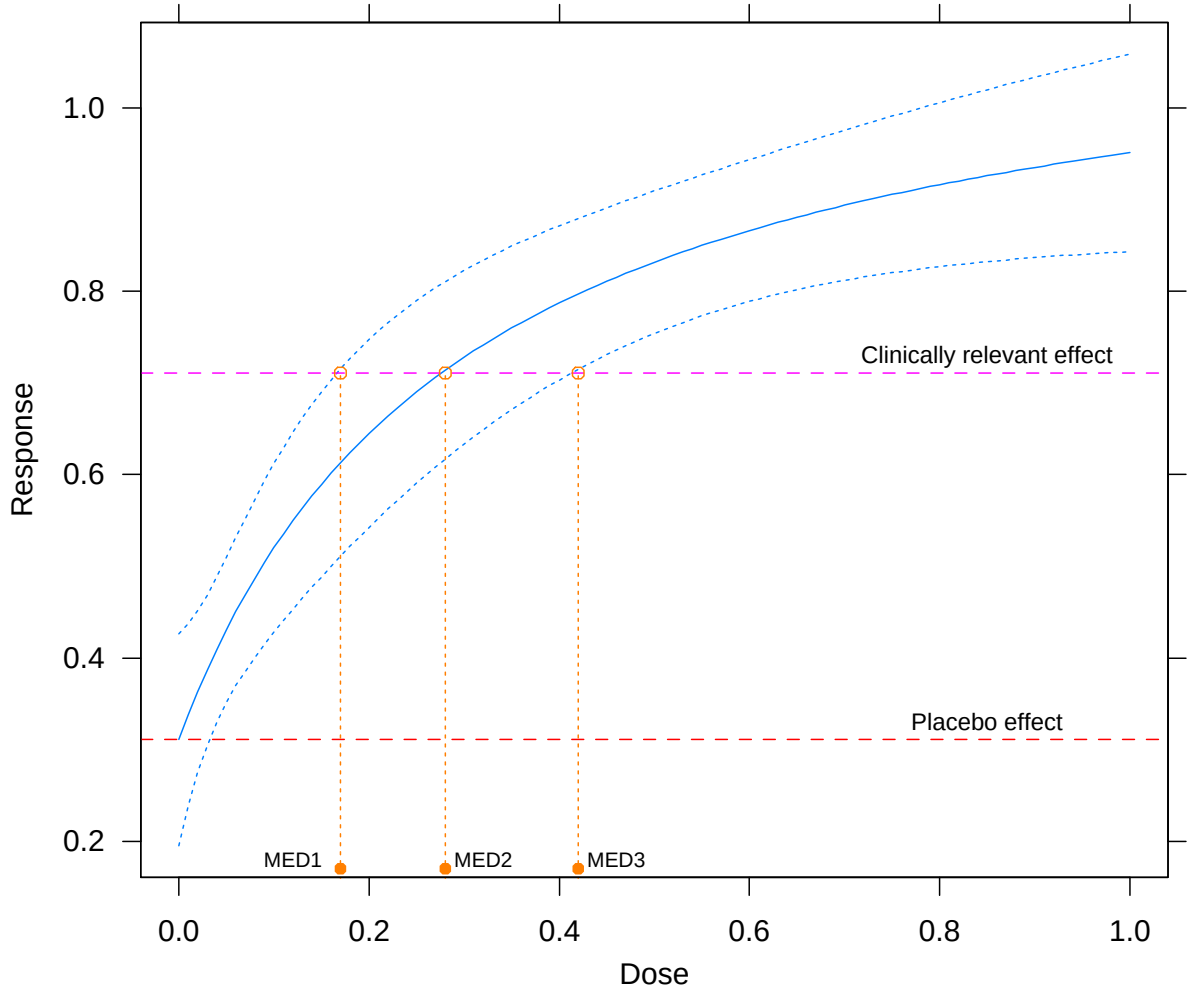
Dose selection and MED estimation in phase II example

Obtaining MED estimates using $\Delta = 0.4$ and 80% confidence level

```
> predEmax <- predict(fmEmax, list(dose = seq(0,1,0.01)),
                      se.fit = T)
> predEmaxDF <-
    data.frame(dose = seq(0,1,0.01), pred = predEmax$fit,
               ll = predEmax$fit - 1.282 * predEmax$se.fit,
               uu = predEmax$fit + 1.282 * predEmax$se.fit)
> predEmaxDF
   dose    pred      ll      uu
 1 0.00 0.31092 0.19538 0.42645   placebo effect
. . .
18 0.17 0.61304 0.51085 0.71524   MED1
. . .
29 0.28 0.71365 0.61692 0.81037   MED2
. . .
43 0.42 0.79703 0.71464 0.87942   MED3
. . .
```

Dose selection and MED estimation (cont.'d)

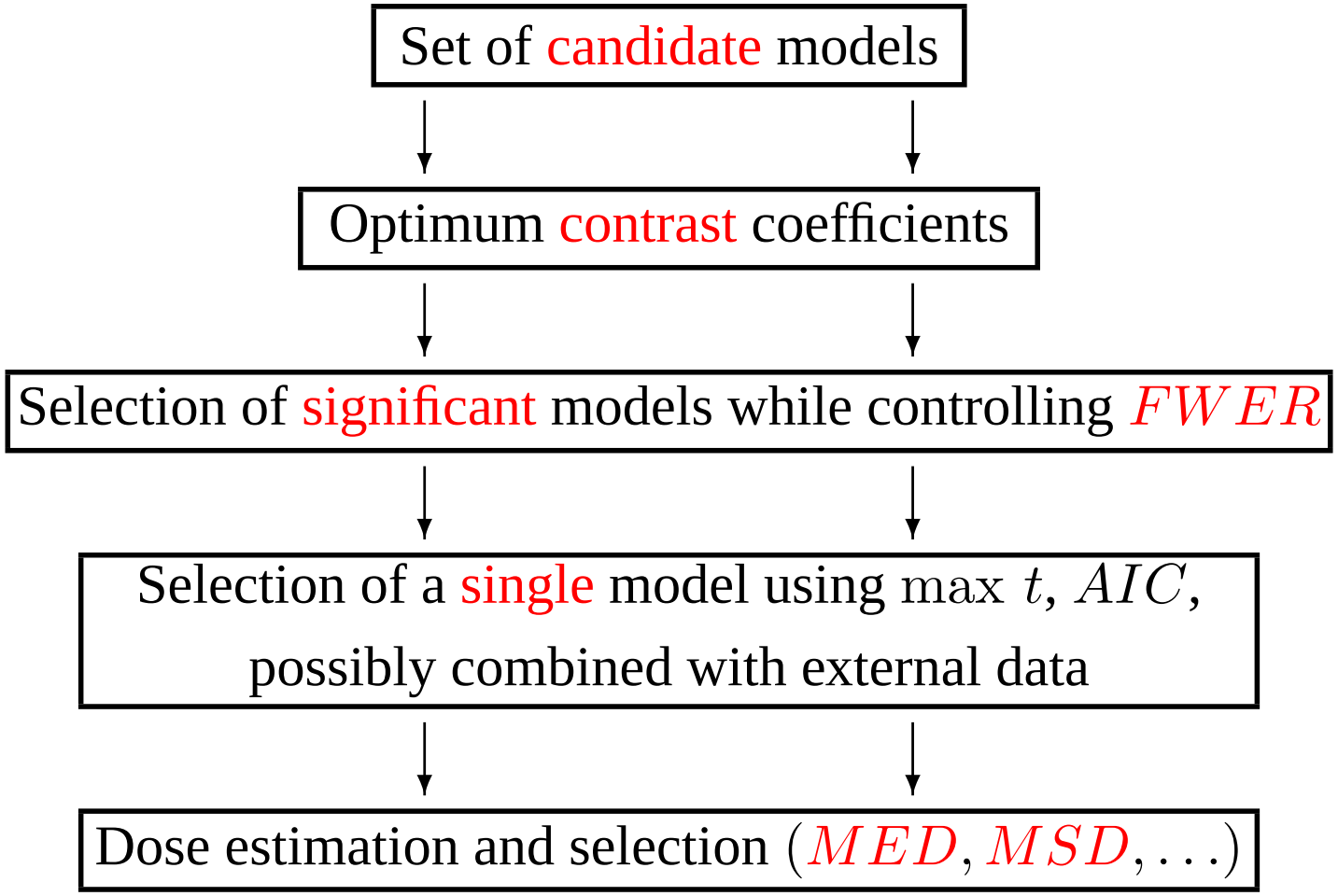
$\widehat{\text{MED}}_1 = 0.17$, $\widehat{\text{MED}}_2 = 0.28$, and $\widehat{\text{MED}}_3 = 0.42$



Model selection problem

- True dose-response model is typically unknown
- Choice of a working model may have a substantial impact on dose selection
- Current model selection approaches mostly do not take into account additional statistical uncertainty associated with the choice of the dose-response model
- How to combine MCP and modeling, using the advantages of both approaches?

MCP-Mod: a unified dose-finding approach



Model selection: Testing PoC

- Models in candidate set are tested using optimal contrasts to obtain t-statistics: $\boldsymbol{T} = \sqrt{n}\boldsymbol{C}'\overline{\boldsymbol{Y}}/s$, with $\boldsymbol{C} = [\boldsymbol{c}_1 \cdots \boldsymbol{c}_M]$
- Optimal contrast maximizes non-centrality parameter of associated t-statistics: $\tau = \boldsymbol{c}'\boldsymbol{\mu}/\sigma\sqrt{\sum_{i=1}^{k} c_i^2/n_i}$
- Under balanced allocation, optimal contrast maximizes $\boldsymbol{c}'\boldsymbol{\mu}$: $\boldsymbol{c}_{\text{opt}} = (\boldsymbol{\mu} - \bar{\mu}\mathbf{1})/\|\boldsymbol{\mu} - \bar{\mu}\mathbf{1}\|$ – location and scale invariant
- Under unbalanced allocation, $\boldsymbol{c}_{\text{opt}}$ is obtained via numerical optimization
- $\boldsymbol{c}_{\text{opt}}$ depends on prior estimates for standardized model

Model selection: Testing PoC (cont.'d)

- t-statistics for candidate models are jointly distributed as multivariate-t with correlation matrix determined by the model contrasts
- Critical value q for individual tests derived from multivariate t-distribution: controls FWER in strong sense
- If $\max \boldsymbol{T} \leq q$, PoC is not established (no apparent dose-response relationship)

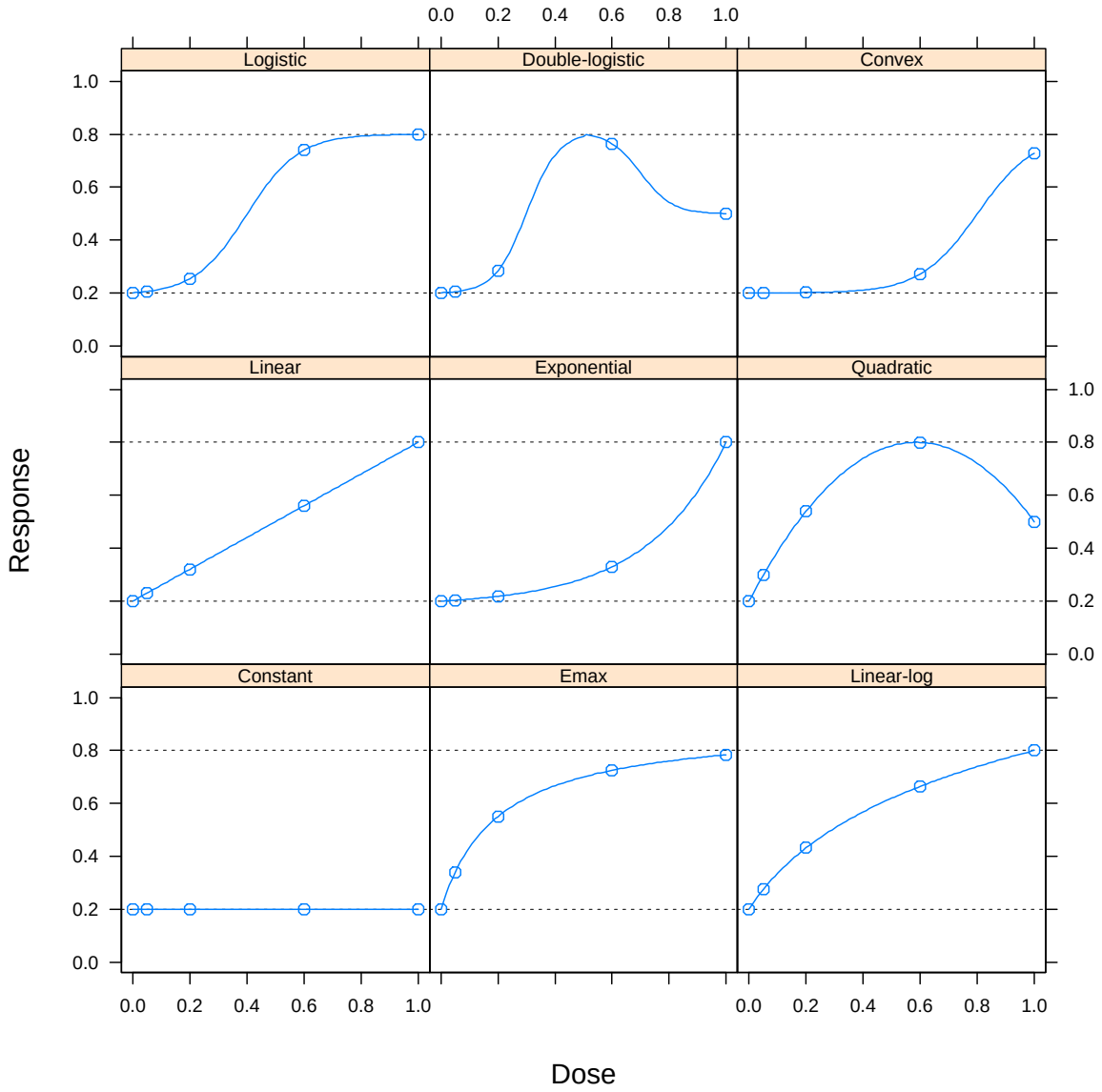
Model selection

- Models with $T_m > q$ are kept for possible use in dose selection
- Different criteria may be used to choose dose-response model among those passing the PoC filter
- Once PoC is established, a most adequate dose-response model is selected among those indicated as significant by the PoC tests, and target doses of interest are estimated using the fitted model
- Different criteria can be used to choose the dose-response model: e.g., max t-statistic, min AIC or min BIC

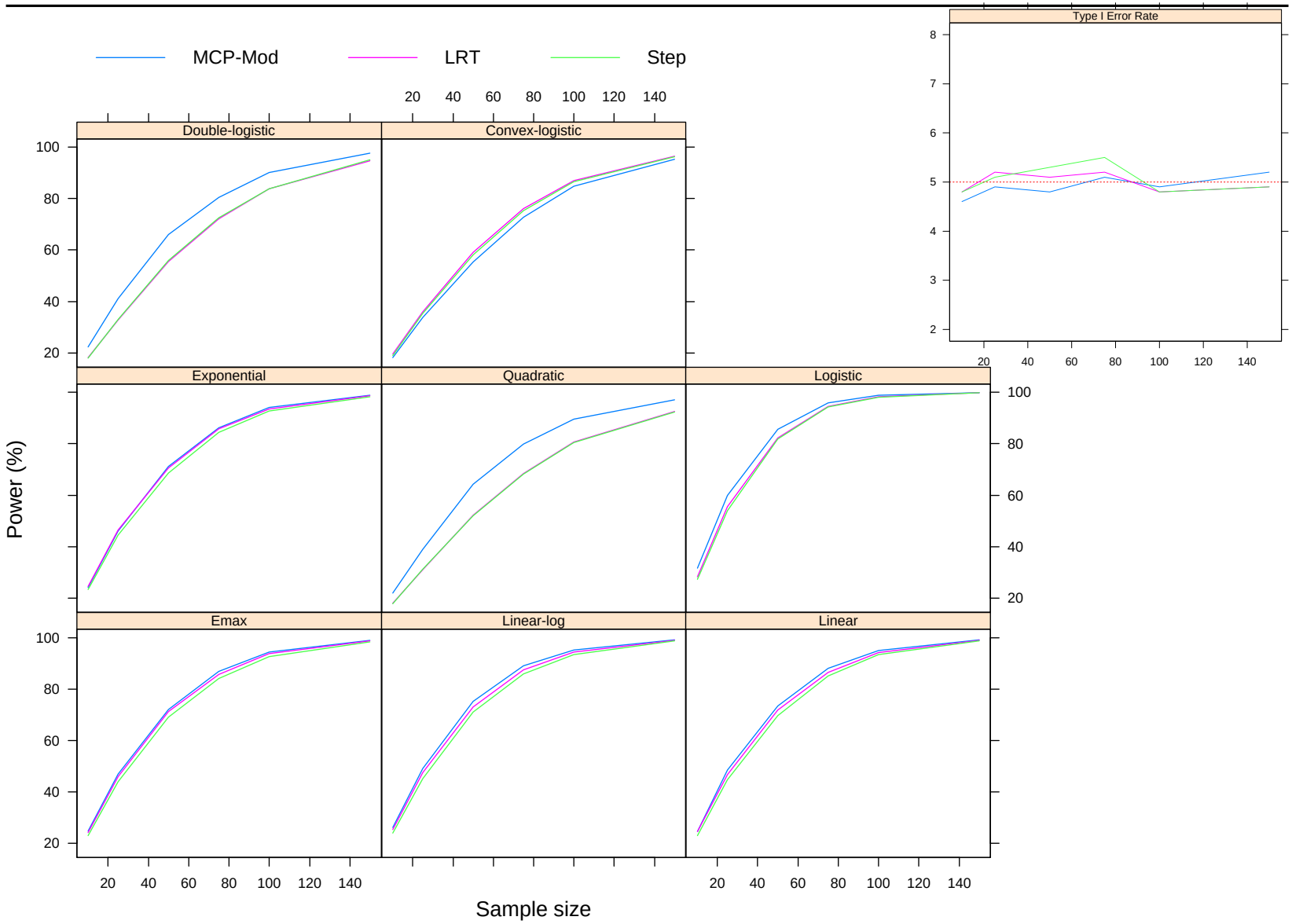
Simulation study

- Objective: to investigate performance of MCP-Mod method with respect to PoC identification and dose selection (MED)
- Design:
  - Same doses and assumptions as in phase II example
  - Parallel groups with equal sample sizes per treatment: $n = 10, 25, 50, 75, 100$ and 150
  - Standard deviations for response: $\sigma = 1.478$ for PoC evaluation and $\sigma = 0.65$ (same as in phase II example) for dose selection evaluation
  - Nine different generating dose-response models
  - 10,000 simulated trials for each sample size $\times$ model comb.
- Likelihood ratio test (LRT) and step contrasts included for comparison with MCP-Mod in PoC evaluation

Simulation study: generating dose-response models

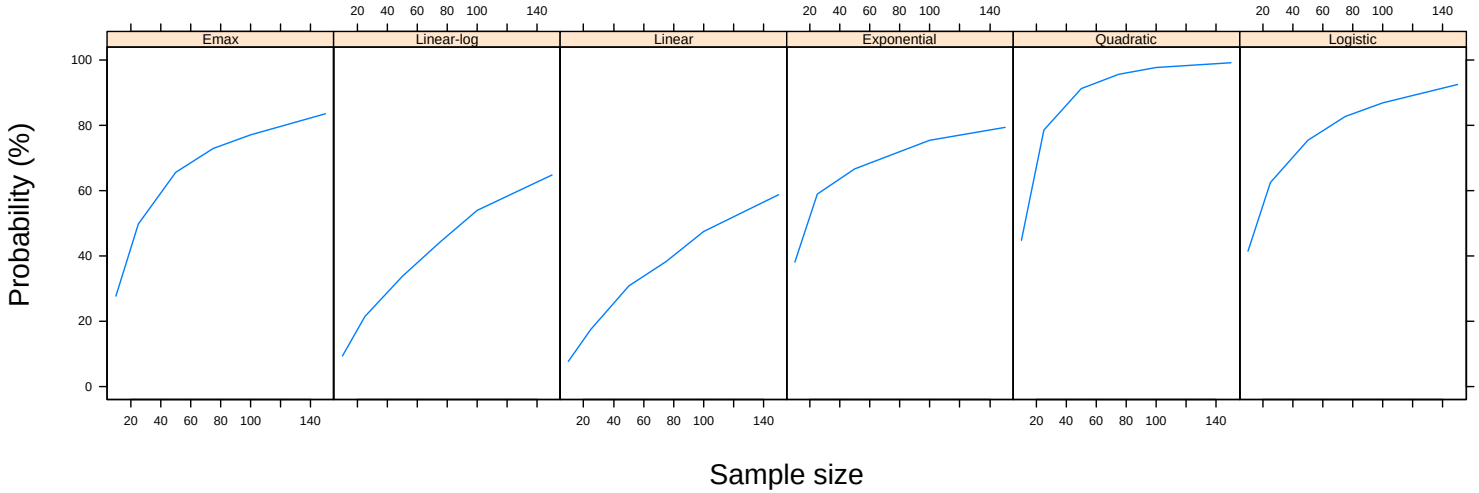


Simulation results: PoC power and Type I error rate

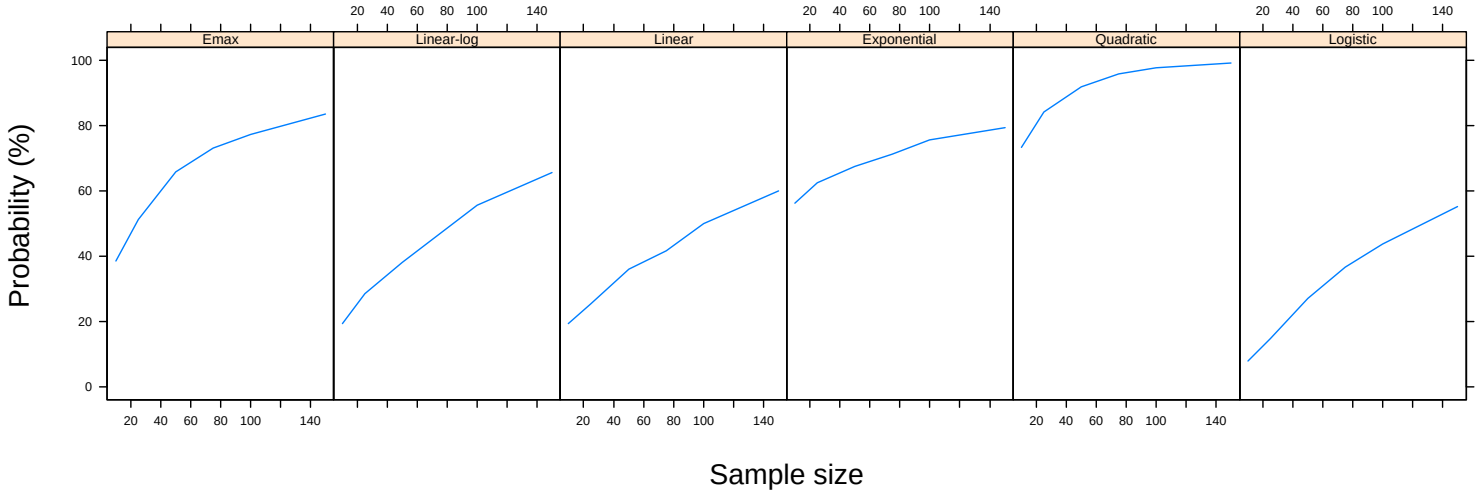


PoC step model vs. dose selection model

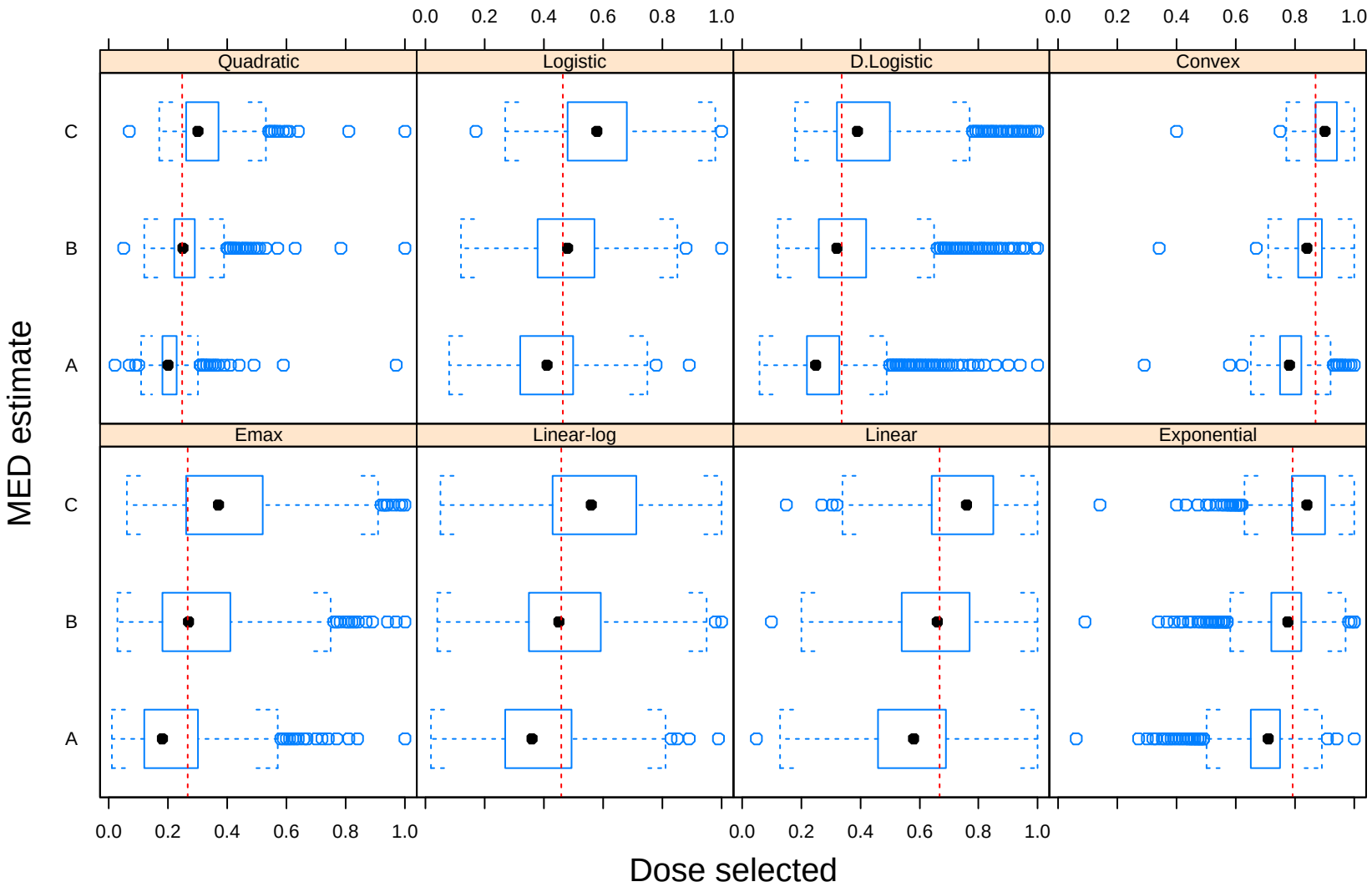
Probabilities of selecting generating model in PoC step



Probabilities of using generating model for dose selection



MED estimation: n = 75, conf. band level = 80%



Results from simulation study

PoC power simulation results

- MCP-Mod has PoC power comparable to MCP methods (LRT, step contrasts) in the case of monotonic shapes and better than MCP methods for non-monotonic shapes
- All methods give good control over the FWER, staying close to the nominal 5% level for all sample sizes

MED estimation results

- Dose selection is a more difficult problem than establishing PoC: sample sizes that have sufficient power for PoC do not give enough precision for estimation of MED
- $\widehat{\text{MED}}_2$ tends to have better performance in estimating target doses: $\widehat{\text{MED}}_1$ tends to underestimate, $\widehat{\text{MED}}_3$ to overestimate
- Precision and bias of dose selection depend on dose-response shape: performance tends to improve with amount of curvature

Conclusions

- Described a unified approach for analysis of dose finding studies: testing PoC and estimating target doses, thus essentially combining independent Ph IIa and Ph IIb studies into a single Ph II study
- Proposed method, MCP-Mod, combines advantages of MCP and modeling approaches while including learning and confirmatory in a single study
- Advantage of MCP-Mod is its greater flexibility in searching for and identifying target doses
- Extensions of MCP-Mod currently under investigation: longitudinal data, binary outcomes, robust designs, two-stage designs, sensitivity analysis