# Some recent experiences with novel statistical approaches to dose finding

**Ekkehard Glimm**
Novartis Pharma AG, Basel, Switzerland

Joint MHRA/EFSPI meeting, London
March 30, 2010

# Outline

Introduction

Steps in MCPMod

Experience and statisticians' role

Conclusions

# Background

- Poor understanding of dose response (DR) for both efficacy and safety has been indicated by regulatory agencies and industry as a root cause of late phase attrition and post-marketing problems with approved drugs
- ICH E4: "Assessment of dose-response should be an integral component of drug development"
- ICH-E4: Purpose of dose-response information is to find the *Smallest dose with a discernible useful effect*

$\Rightarrow$ Need to develop designs and methods for efficient learning about DR, enabling better and faster decision making on dose selection and improved labeling

# MCPMod

## Motivation

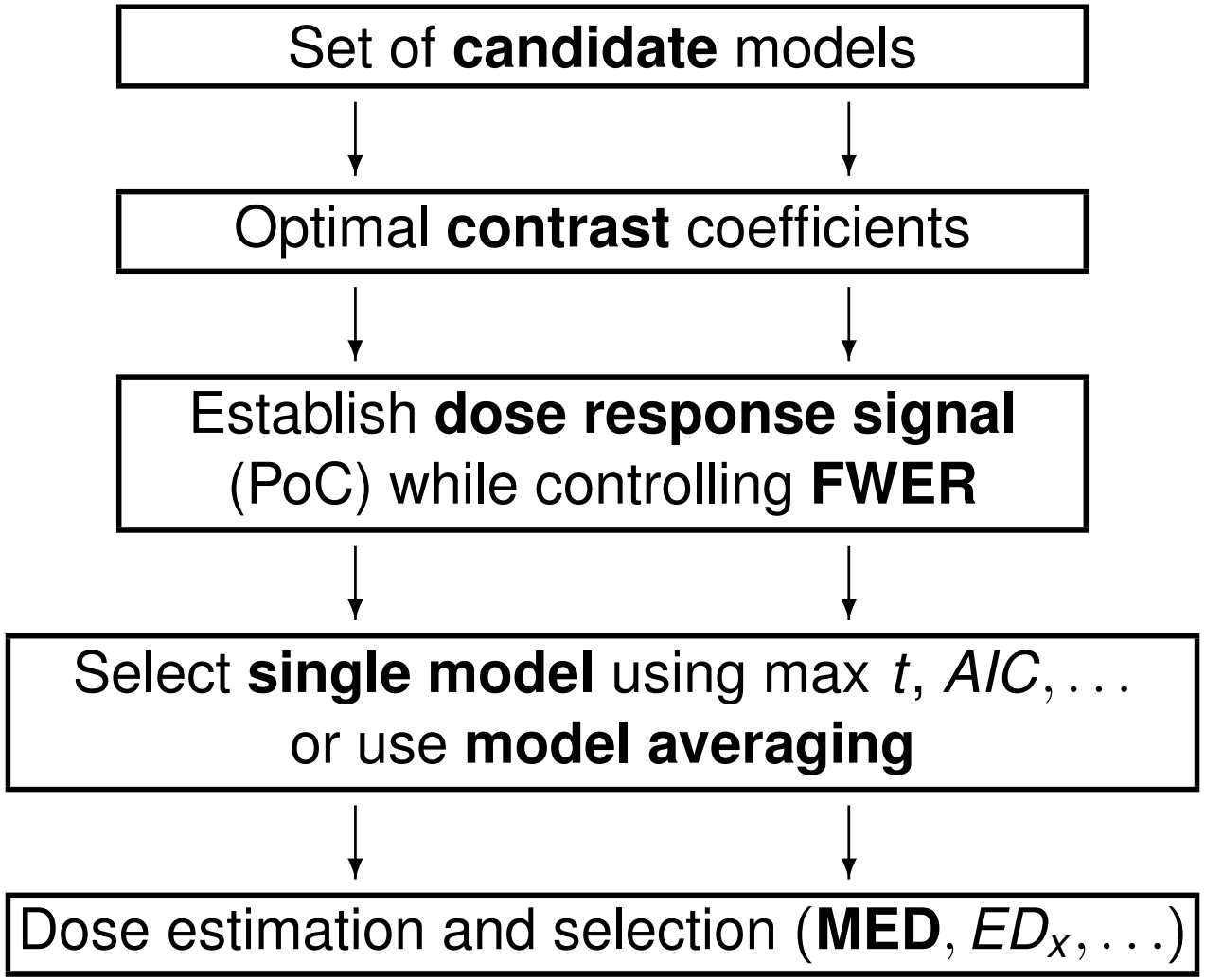
Two main **goals** in Phase II studies:

- **proof-of-concept** (PoC) – any evidence of dose response (i.e., treatment effect)
- **dose-selection** – which dose(s) to take into phase III?
- Main analysis strategies: *multiple comparisons (MCP) of contrasts between doses and modeling of dose response*

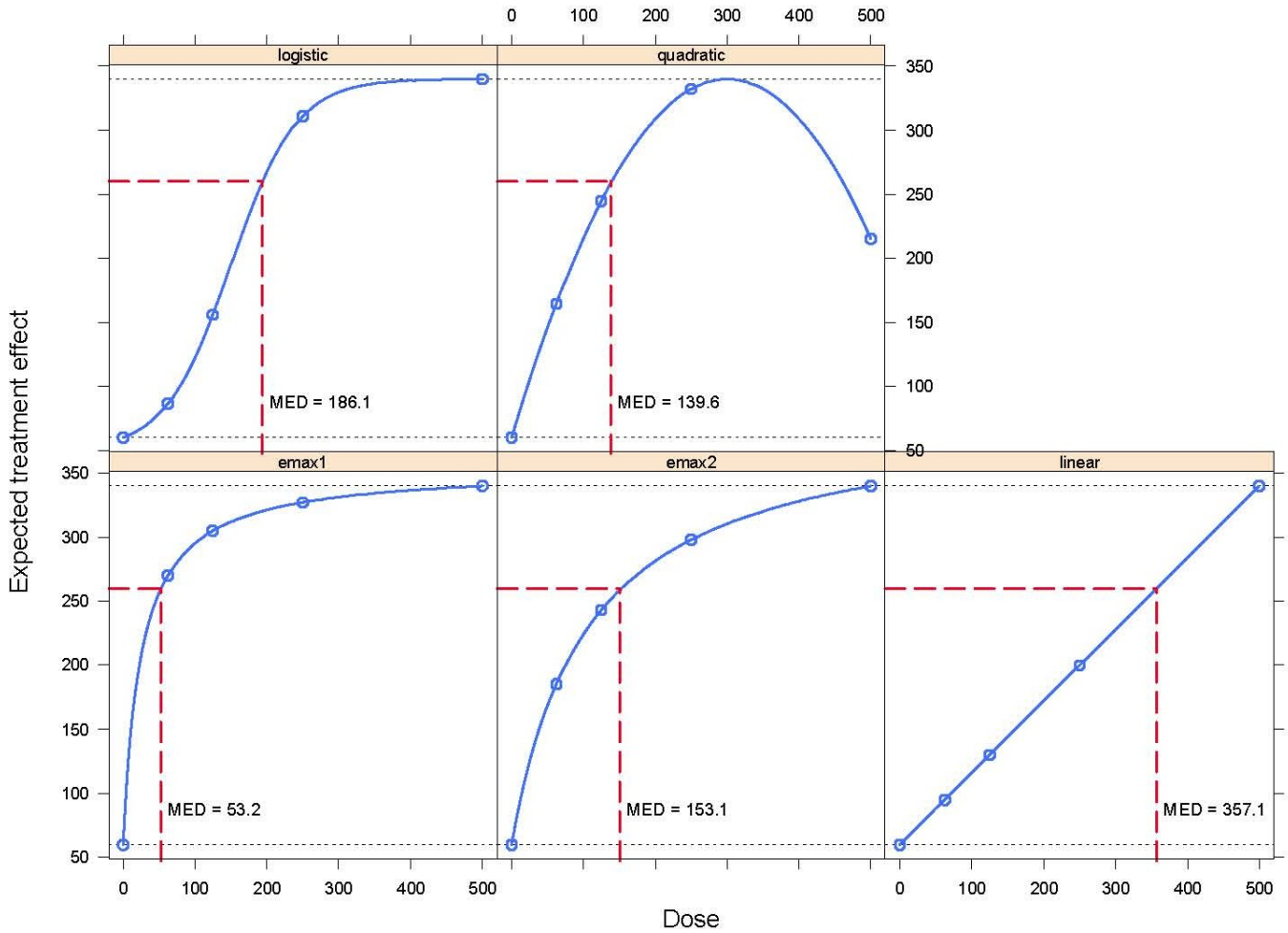
# MCPMod: A unified approach to dose finding

- Bretz et al. (2005) proposed **MCPMod** for testing PoC and estimating target doses in Phase II studies
- Combines advantages of MCP and modeling approaches:
  - Set of **candidate** DR models to account for **model uncertainty**
  - Test PoC using MCP based on optimal **model contrasts**
  - If PoC is established, model **selection** via information criteria (e.g., AIC, BIC) and/or other suitable criteria
  - Estimating DR and target doses (e.g., MED and MSD) via modeling, or model averaging

# MCPMod in a nutshell



# Impact of model uncertainty on target dose



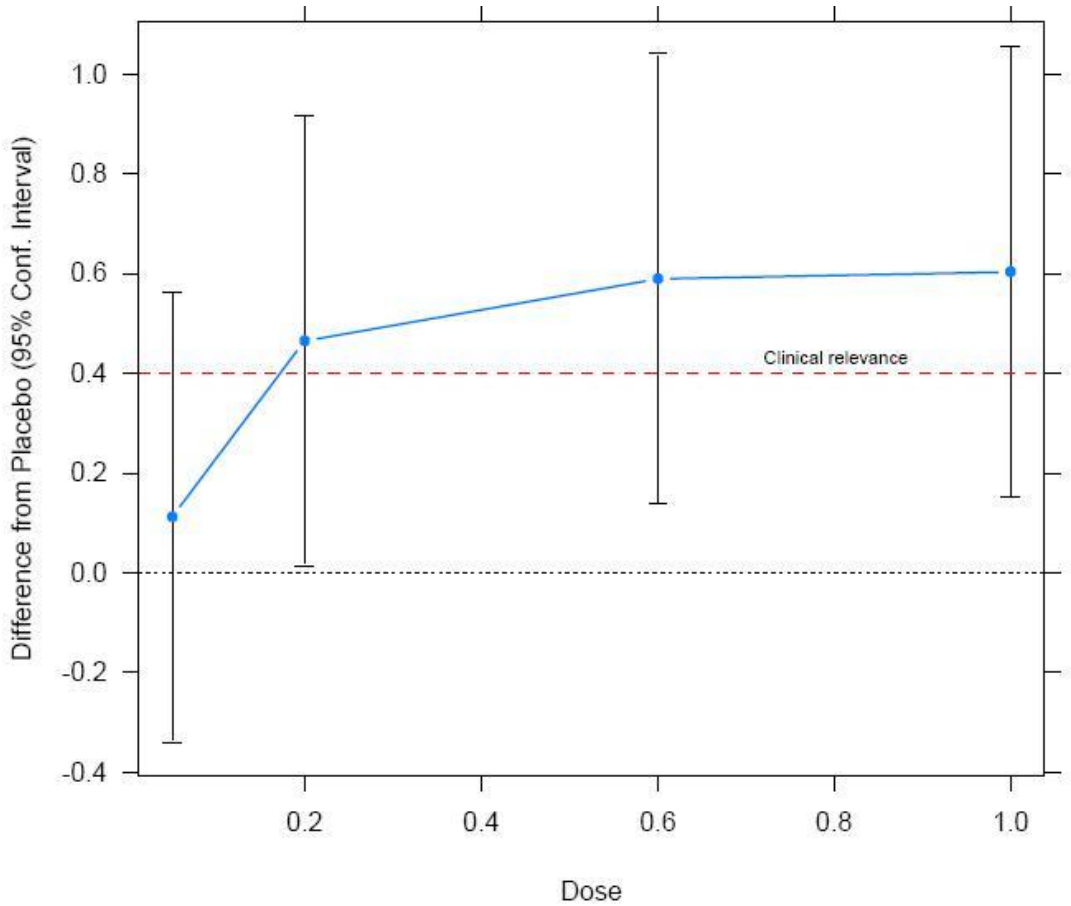
# Finding the right dose is not that simple

- True shape of dose-response model is typically unknown
- Choice of a working model may have a substantial impact on dose selection
- Model selection using observed data needs to account for statistical uncertainty and associated multiplicity issues
  $\Longrightarrow$ Useful to have a unified approach combining the advantages of MCP and modeling: this is the goal of MCP-Mod

# Example: Analysis of a dose-finding study

- Randomized, double-blind parallel group Ph II trial with 100 patients equally allocated to placebo or one of four active doses: 0.05, 0.2, 0.6, or 1
- All doses well tolerated
- Normally distributed primary endpoint with equal variances across dose groups
- Planned PoC analysis: Fixed sequence test that preserves the Type I error (T1E) at 5% two-sided level
- Conclusion: Top three doses are significantly better than placebo.

# Example (cont.)

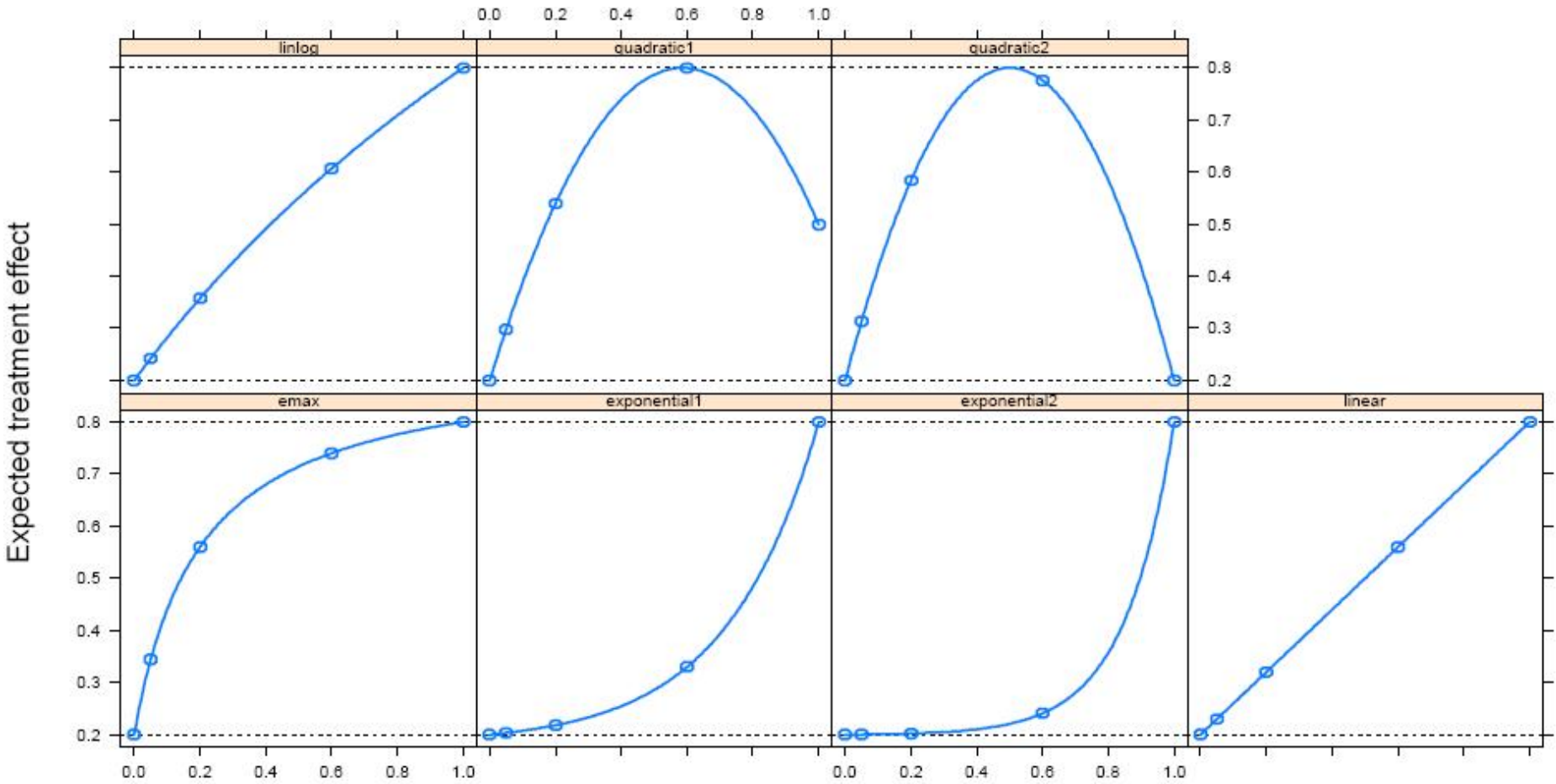



Which dose should be considered the MED?

# Candidate models

- First step of MCP-Mod: Set of candidate models
- Propose several DR models at the planning stage to describe potential outcomes
- Model uncertainty directly acknowledged
- Requires close collaboration within clinical team
- Input based on available information (PK/PD data, historical data from similar compounds, mode of action, ...)

# Candidate models (Example)



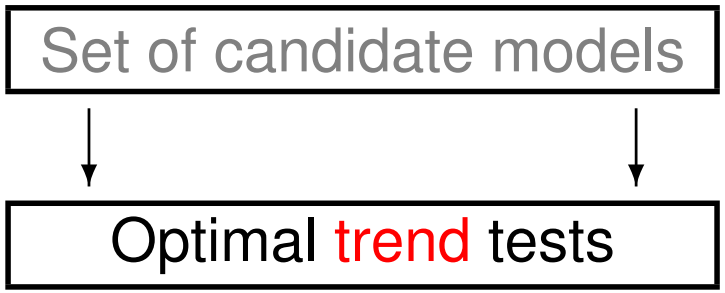
# Candidate models (Example cont.)

| Model | $f(d, \theta)$ |
|---|---|
| Linear | $E_0 + \delta d$ |
| $E_{max}$ | $E_0 + E_{max} d/(ED_{50} + d)$ |
| Linear log-dose | $E_0 + \delta \log(d + c)$ |
| Exponential 1 | $E_0 + E_1(\exp(d/\delta) - 1)$ |
| Exponential 2 | $E_0 + E_1(\exp(d/\tilde{\delta}) - 1)$ |
| Quadratic 1 | $E_0 + \beta_1 d + \beta_2 d^2$ |
| Quadratic 2 | $E_0 + \tilde{\beta}_1 d + \tilde{\beta}_2 d^2$ |

- Parameters in red need to be pre-specified by the clinical team: best guesses at planning stage

# Trend tests

- Second step of MCP-Mod:



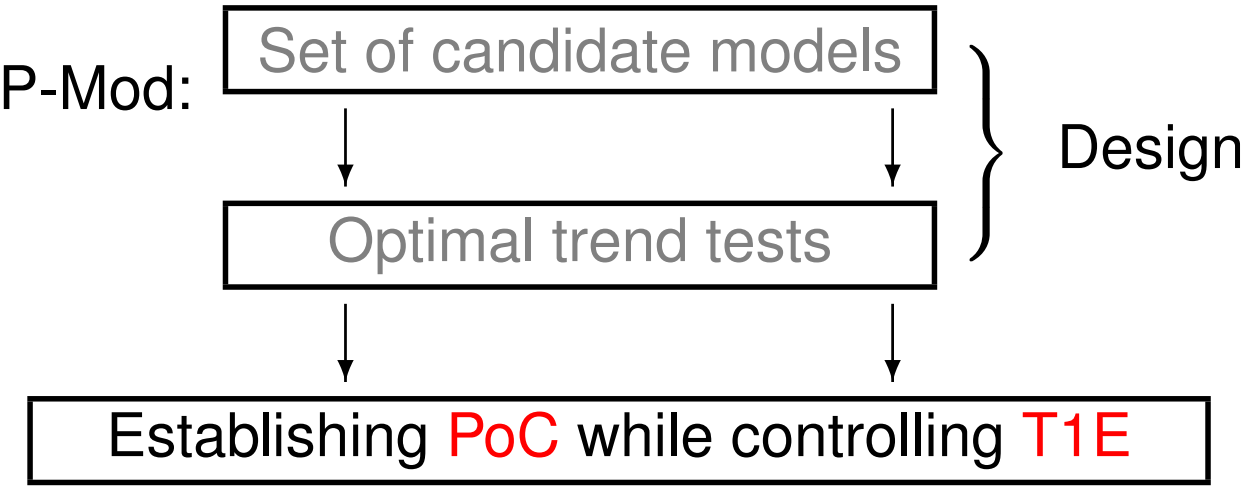

- Trend test: Each model will be tested using optimally chosen weights to detect a dose-related trend in effect.
- For each dose response model, the weights are chosen to maximize power in detecting a treatment effect for that model.
- That is, dose weights are calculated to maximize the difference between the flat model (= no DR) and each candidate model
- These weights depend on best guesses for model parameters available at the planning stage

# Remarks ...

- High correlations (e.g. 0.99 between linear and log-linear models) indicate similarity between candidate models
- Little multiplicity penalty, as fairly high, positive correlations are accounted for in calculating critical values, p-values, ...
- Degree of discrimination:
  models may be difficult to distinguish, but then resulting MEDs would also be similar
- Low number of models is recommended (typically $4 - 5$ models)
- Linear and $E_{max}$ models are often included in candidate set; other models (e.g., quadratic, logistic, exponential, ...) are included as needed

# Establishing proof-of-concept

- Third step of MCP-Mod:




- Analysis objectives:
  - Establish PoC
  - Select best model(s) out of candidate set
  - Estimation of target dose

# Trend test for a given candidate model

- Once data is observed, the response at each dose is multiplied by the corresponding weight for each candidate model and then summed up.
- Each of these sums will define a trend test and thus measure the "distance" of the observed data from the flat model (= no DR)
- If the observed DR curve is close to the candidate curve, this sum will be large.
- Final test statistic $T$ is obtained after standardizing this sum.

# Establishing proof-of-concept

- Establish PoC if at least one trend test is significant.
- PoC is established if the maximum test statistics (across all $M$ models) is larger than the critical value, i.e.

$$\max\{T_1, \ldots, T_M\} > q_{1-\alpha}$$

- Critical value $q_{1-\alpha}$ is calculated such that T1E is controlled, where $\alpha$ is specified in the protocol
- All models with $T > q_{1-\alpha}$ are kept for possible use in dose-response modeling
- If $\max\{T_1, \ldots, T_M\} \leq q_{1-\alpha}$, PoC cannot be established

# Establishing proof-of-concept (Example)

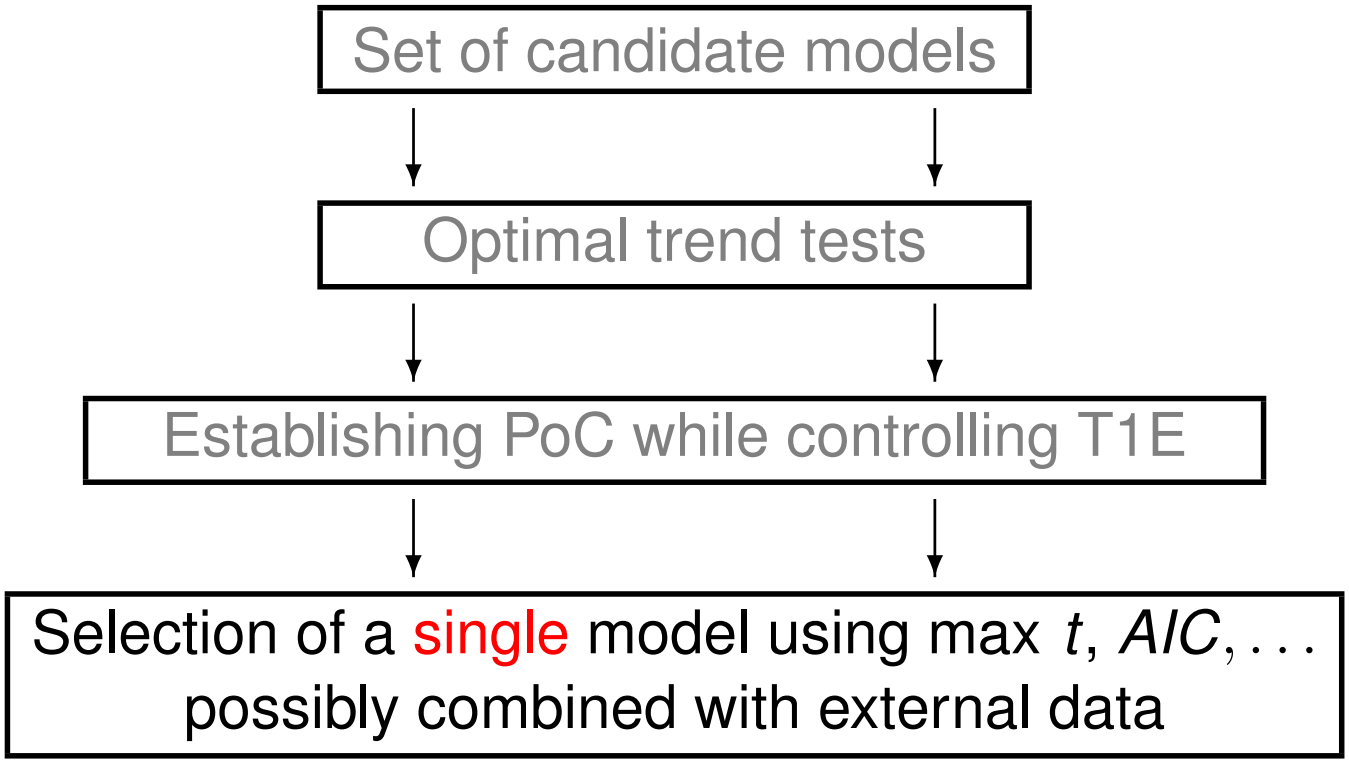
- 5% one-sided critical value for $T_m$'s: $q = 2.15$

| Model | est. | s.e. | $t$-value ($T_m$) | $P$-value | adj. $P$-value |
|---|---|---|---|---|---|
| Emax | 0.55 | 0.159 | 3.46 | 0.0004 | 0.001 |
| Linlog | 0.49 | 0.159 | 3.11 | 0.0012 | 0.004 |
| Quad 1 | 0.49 | 0.159 | 3.10 | 0.0013 | 0.004 |
| Linear | 0.47 | 0.159 | 2.97 | 0.0019 | 0.006 |
| Exp 1 | 0.35 | 0.159 | 2.22 | 0.0145 | 0.044 |
| Exp 2 | 0.30 | 0.159 | 1.90 | 0.0304 | 0.086 |
| Quad 2 | 0.29 | 0.159 | 1.85 | 0.0337 | 0.094 |

- PoC established: $\max_m T_m = 3.46 > 2.15$

# Model Selection

- Fourth Step of MCP-Mod:

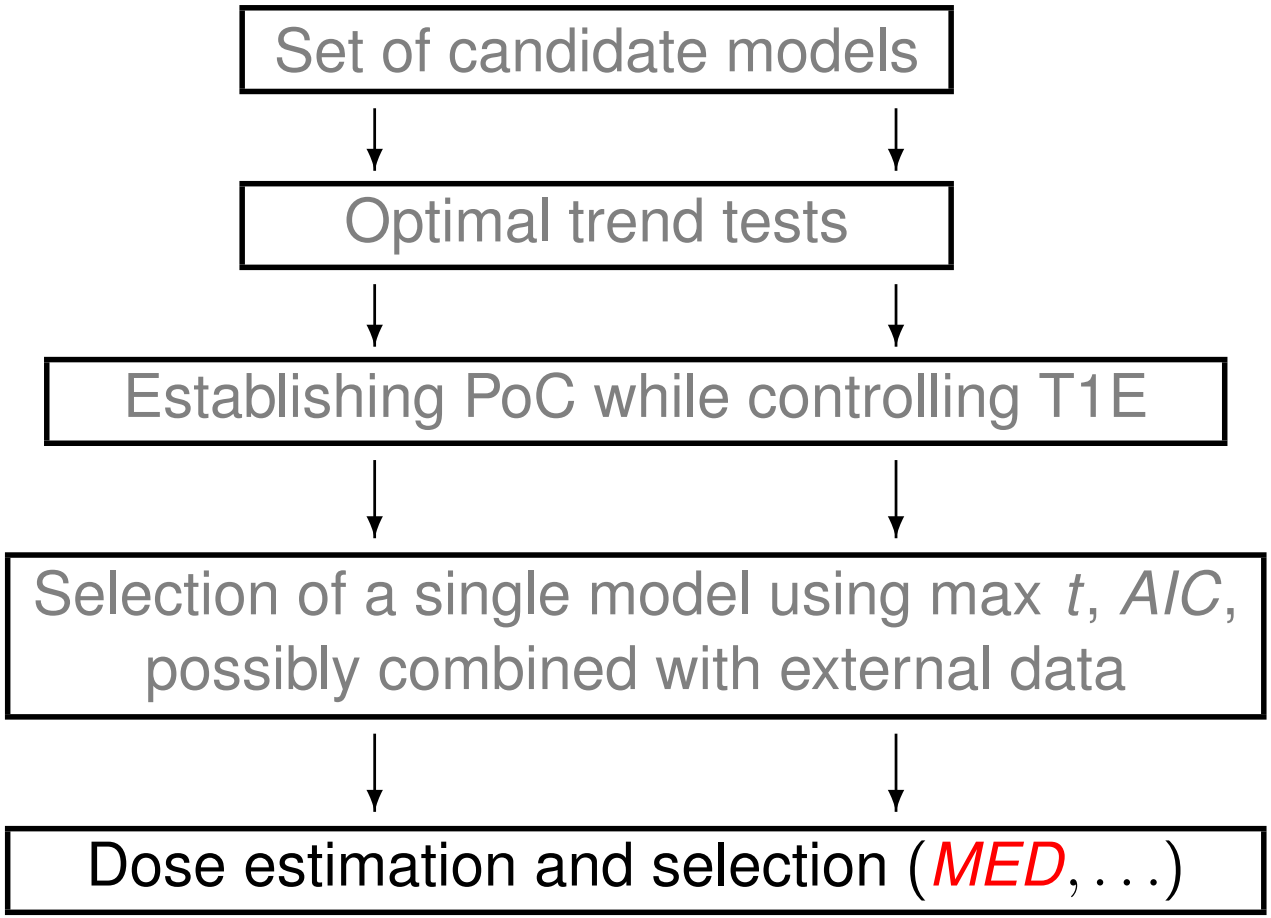


# Model selection

- Once PoC is established , a most adequate dose-response model is selected among those indicated as significant
- Different criteria may be used to choose among models passing the PoC filter, e.g., max $t$-statistic, min AIC or min BIC
- Target doses of interest are estimated using the selected model
- Alternative: Take weighted average of target doses derived from all significant models

# Dose estimation

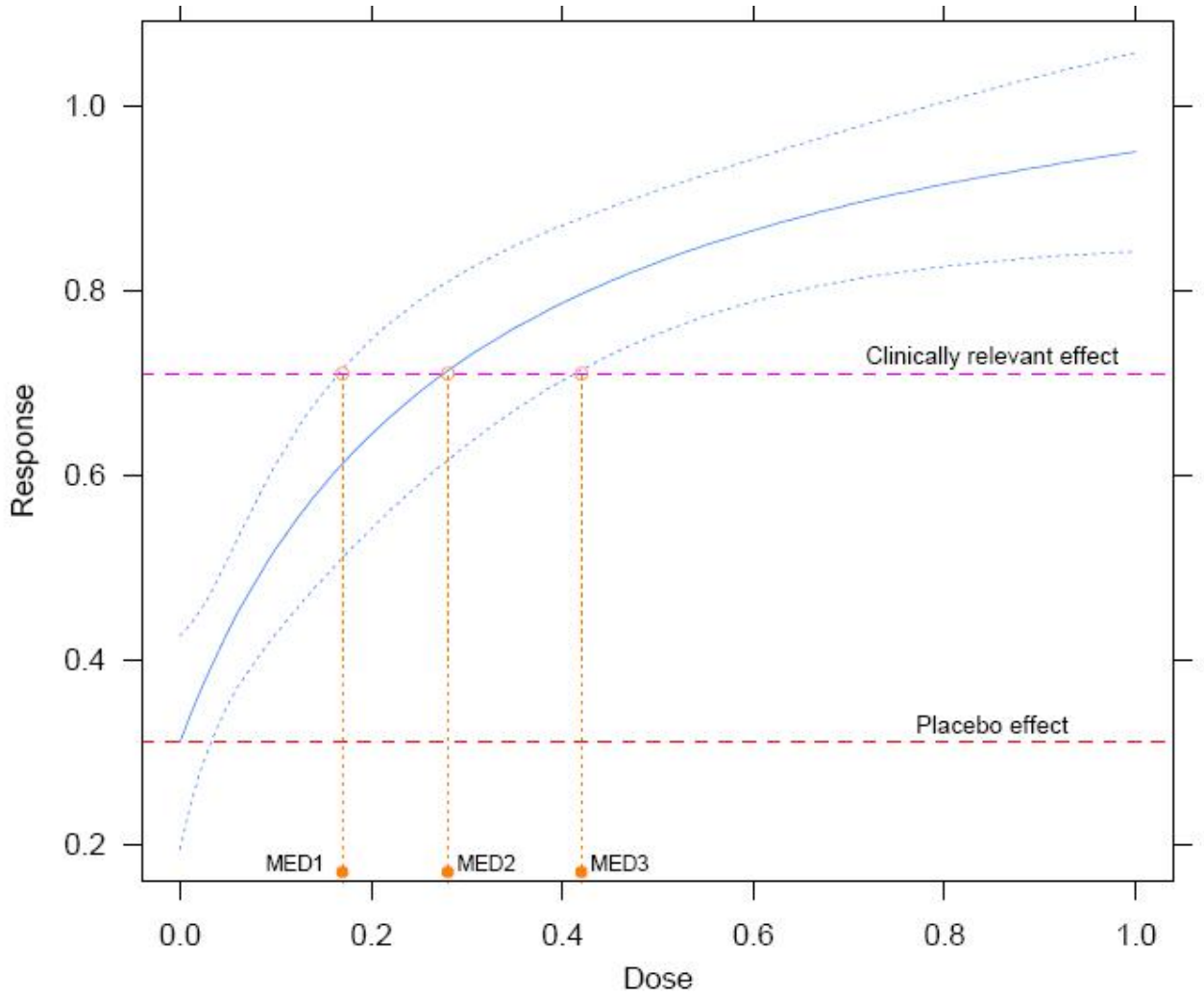
- Fifth and last step of MCP-Mod



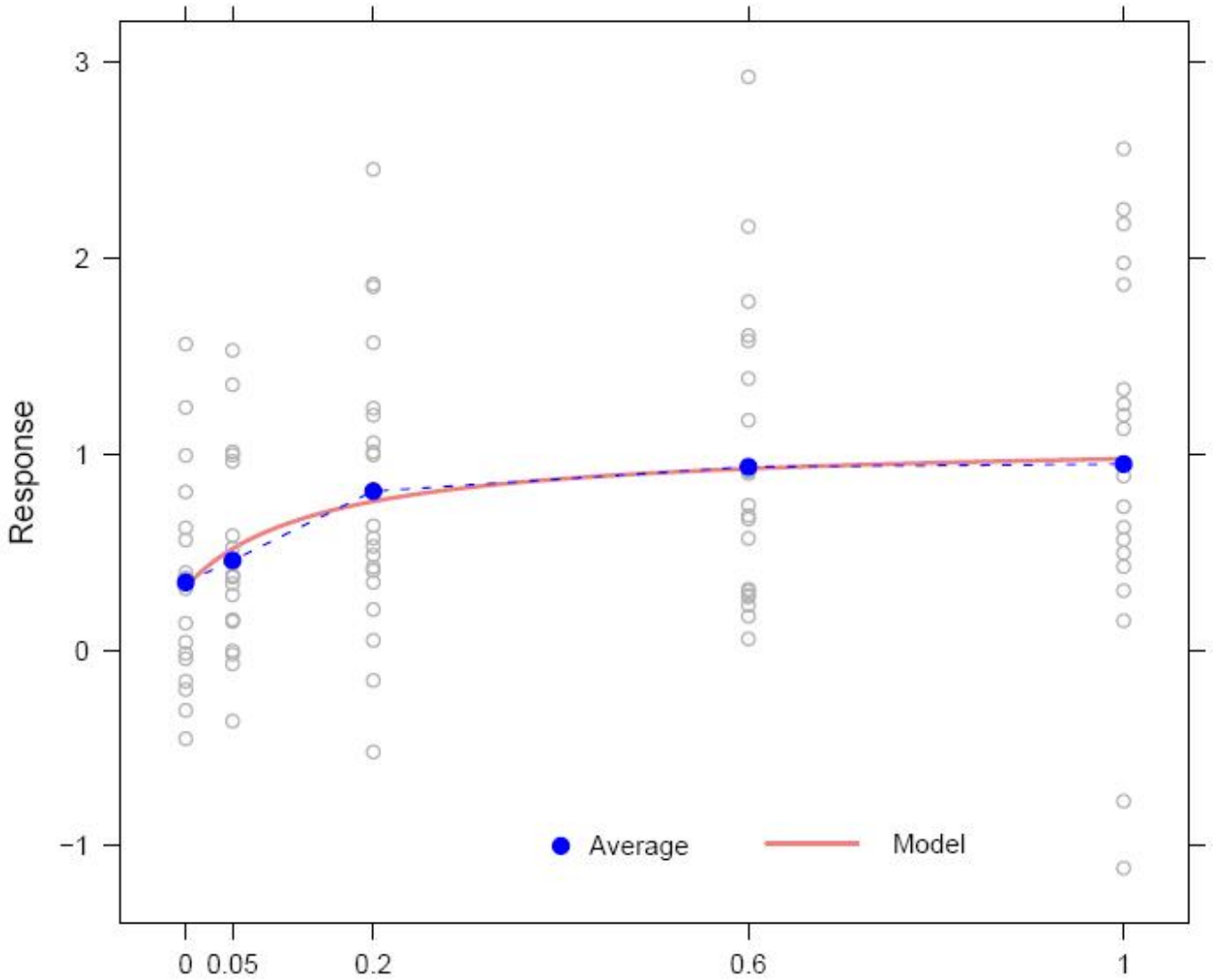
# Dose estimation

- From guidelines/clinicians: Absolute clinically relevant difference $\Delta$ with respect to placebo
- Relevant target dose: MED, i.e. smallest dose with a discernible useful effect (ICH E4)
- Different rules proposed for estimating MED available
- Estimated MEDs may not exist for some, or all of the methods

# Dose estimation



# Dose estimation (Example) – full data



# Experience

## Statisticians

- Many Novartis studies ($>$30) from various therapeutic areas have used this approach
- Project statisticians have been very supportive
- The approach gives statisticians an important role in dose-finding discussions in general
- Many extensions of the original methodology originated from statisticians' requests
- Several TA statisticians very actively contributed to software development (R and S+ packages available, some SAS implementations)

# Extensions

- Comparison with an active control
- non-normal endpoints (binary, time-to-event)
- covariates: baseline, PK-characteristics (e.g. AUC)
- optimal and response-adaptive allocation
- bootstrapping confidence intervals to account for model selection uncertainty
- more robust fitting algorithms, different regimens, time profiles, ...

# Experience

## Collection of relevant information

- The need to fix some aspects (dose levels, candidate profiles, guesstimates, ...) forces close collaboration and initiates a thought process
- Candidate models are often a set of standard models (linear, Emax) plus some from brainstorming of the entire clinical team
- Guesstimates are from publications and/or inferred by statisticians from discussion with clinicians
- Dose group allocation also a mix of statistical optimality and clinical considerations

# Experience

## Clinicians

- also receptive, but not as enthusiastic as statisticians
- If there were reservations then
  - Transition problems between early and full development: ED used to very small PoC trials
  - Some clinicians seem to be content with identifying a "dose that works", no interest in MED or the entire dose-response profile (and for some drugs, they are right)
  - Some clinicians are used to the ANOVA approach (esp. the Dunnett test)
- Points that require some explanation
  - Many clinicians (at least in pharma) think in terms of statistical significance of dose groups $\rightarrow$ explain the need to replace by concepts like MED if there is an underlying dose-response model

# Common pitfalls and limitations in dose finding

DF should be an integral component of drug development, but:

- Model uncertainty is rarely acknowledged, but has severe consequences: model selection problems, biased estimates, overfitting, overconfident conclusions, etc.
- Traditional modeling approaches are often not appropriate, if the class of considered working models is too narrow
- Traditional hypotheses tests (e.g. Dunnett test) are not appropriate for dose estimation
- Neither of these approaches acknowledges that DF studies have multiple objectives
- Estimating DR is considerably harder than testing for it
- Current sample sizes for DF studies, based on power to detect DR, are inappropriate for dose selection and DR estimation

# Conclusions and recommendations

- Model-based dose-finding designs should be used routinely in drug development, as they can lead to substantial gains in performance over traditional DF methods
- Sample size calculations for Phase II studies should take into account desired precision of estimated target dose and possibly also estimated DR (current methods are not appropriate)
- When resulting sample size is not feasible, should consider selecting two or three doses for confirmatory phase to increase likelihood of including "correct" dose – adaptive designs could be used in confirmatory phase for greater efficiency (e.g., dropping less efficient doses earlier)

# References on MCP-Mod


- Bornkamp, B. et al. (2007) Innovative approaches for designing and analyzing adaptive dose-ranging trials – White Paper from the PhRMA working group on "Adaptive Dose-Ranging Studies" (with discussion). *Journal of Biopharmaceutical Statistics*, **17**(6), 965–995.
- Bornkamp, B., Pinheiro, J.C., and Bretz, F. (2009) `MCPMod` - An R Package for the Design and Analysis of Dose-Finding Studies. *Journal of Statistical Software*, **29**(7), 1–23.
- Bretz, F., Pinheiro, J.C., and Branson, M. (2004) On a hybrid method in dose finding studies. *Methods of Information in Medicine*, **43**(5), 457–461.
- Bretz, F., Pinheiro, J.C., and Branson, M. (2005) Combining multiple comparisons and modeling techniques in dose-response studies. *Biometrics*, **61**(3), 738–748.
- Bretz, F., Hsu, J., Pinheiro, J.C., and Liu, Y. (2008) Dose Finding – A Challenge in Statistics. *Biometrical Journal*, **50**(4), 480-504.
- Dette, H., Bretz, F., Pepelyshev, A., and Pinheiro, J.C. (2008) Optimal Designs for Dose Finding Studies. *Journal of the American Statistical Association* **103**(483), 1225–1237.
- Pinheiro, J.C., Bretz, F., and Branson, M. (2006) Analysis of dose-response studies: Modeling approaches. In: *Dose Finding in Drug Development*. Ting, N. (ed.), Springer, New York, pp. 146-171.
- Pinheiro, J.C., Bornkamp, B., and Bretz, F. (2006) Design and analysis of dose finding studies combining multiple comparisons and modeling procedures. *Journal of Biopharmaceutical Statistics*, **16**(5), 639–656.