Discussion Meeting for MCP-Mod Qualification Opinion Request

Novartis

10 July 2013

EMA, London, UK

NOVARTIS

Attendees

Face to face:

- Dr. Frank Bretz Global Statistical Methodology Head, Novartis
- Dr. Björn Bornkamp Expert Statistical Methodologist, Novartis
- Dr. Geneviève Le Visage Regulatory Intelligence Head, Novartis

By telephone:

- Dr. José Pinheiro Senior Director, Janssen Research & Development

Agenda

1. Introduction:
   - Qualification request
   - Brief introduction to MCP-Mod
   - In-scope, out-scope
2. Answers to Issues 5 – 11 raised by the SAWP

Qualification request

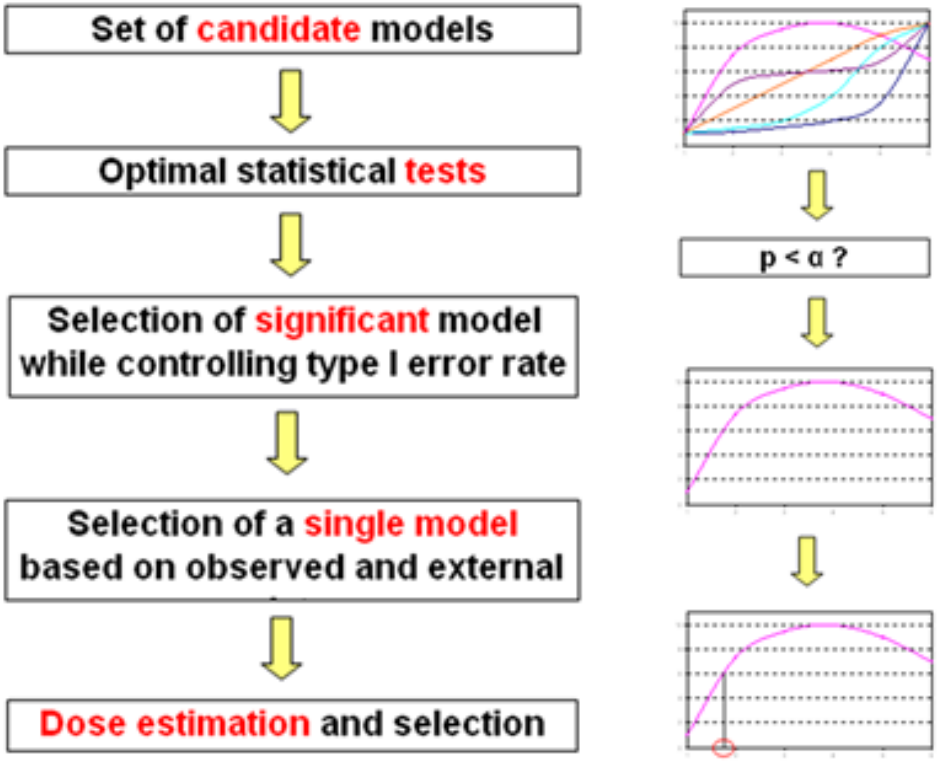
- Novartis is seeking qualification of MCP-Mod as an

Efficient statistical methodology for model-based design and analysis of Phase II dose finding studies under model uncertainty.

- The data supportive of this request consists of the following elements:
  - Worked examples, extensive simulations and real-life case studies to describe and quantify the performance
  - References from medical and statistical literature to illustrate applicability

Background on MCP-Mod methodology

- MCP-Mod stands for:
  Multiple Comparisons & Modelling
  - Combines testing and estimation
- Design stage
  - Pre-specification of candidate dose-response models
- Analysis stage: MCP-step
  - Statistical test for dose-response signal. Model-selection based on significant dose-response models
- Analysis stage: Mod-step
  - Dose-response and target dose estimation based on dose-response modelling



- Difference to traditional pairwise comparisons
  - Use of dose-response modelling
  - But, taking model uncertainty into account at *design* and *analysis* stage

NOVARTIS

In-scope

- Drug development stage
  - Phase II dose finding studies to support dose selection for Phase III
- Response
  - Univariate (efficacy or safety) variable (could be a binary, count, continuous or time-to-event endpoint). Observations could be cross-sectional (i.e. from a single time point) or longitudinal.
- Dose
  - Could be any univariate, continuous, quantitative measurement, as long as an ordering of the measurements is possible and the differences between measurements are interpretable
- Number of doses
  - For the MCP-step at least two distinct active doses are required
  - For the Mod-step, a minimum of three active doses required

Out-of-scope or limited experience

- Predictions from a surrogate / biomarker or short term readout to a clinical Phase III endpoint.
- Titration designs and dose escalation studies (e.g. to estimate the maximum tolerable doses using continual reassessment methods).
- Exposure-response analyses or PK-PD models are not the purpose of this request, per se.
- Regimen finding for biologics where there is no steady state.
- Application of MCP-Mod in confirmatory studies.
- Multivariate problems, e.g., joint modeling of efficacy and toxicity, the presence of two primary endpoints, or drug combination trials.

Answers to Issues 5 – 11

Issue 5

Selection of dose-range, number of doses and spaces of doses

- Can the procedure itself directly help with these choices?
  - Maximum dose: Based on information from previous trials
  - Optimal design theory and clinical trial simulations
    - Input: Anticipated dose-response shapes & trial objective(s)
    - Output: Number and location of doses and allocation ratios to the doses
  - In practice one might deviate from optimal designs
    - Logistical/manufacturing constraints, considerations beyond primary efficacy endpoint
- ... guidance for an optimal strategy for these pre-selection exercises?
  - Candidate models: Honest reflection of potential dose-response curves
    - Not too many shapes (decrease in efficiency), too few shapes (risk of biased results)
    - Often 3-7 dose-response models/shapes seem sufficient
  - Dose-range, number of doses, location of doses case-specific; rules of thumb:
    - >10-fold dose-range, 4-7 active doses, logarithmic dose-spacing

Issue 6

Considerations to optimise the choice of sample size

- Is this optimally based on the precision with which the dose-response curve can be characterised, which would also need to consider dose-spacing and number of doses?
  - Sample size calculations should reflect the study objectives
  - Estimating dose response (DR) is considerably harder than testing it
  - Sample sizes for dose finding studies, based on power to detect DR signal, are inappropriate for dose selection and DR estimation
- Is there a minimum level of information below which the relative benefits of an MCP-Mod approach are lost compared to a 'traditional' approach?
  - Particularly in situations with small sample sizes, borrowing strength through modelling is beneficial, although validation of assumptions becomes difficult
  - MCP-Mod requires at least two (three) active doses for the MCP (Mod) step
    - Traditional approaches don't perform well either for a small number of doses

Issue 7

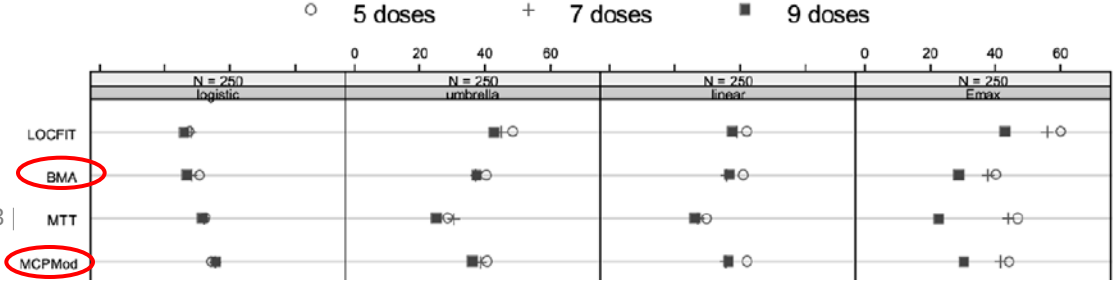
Rationale and the choice of nominal significance level

- Is control at the traditional 5% level optimal from a sponsor point of view?
  - Depends on the specific trial and context
    - Understanding the false positive rate is important for any decision procedure
    - What certainty does the company need in the decision to move forward?
    - What is being tested
      - dose response signal detection relative to placebo? active control?
  - One major focus of MCP-Mod is estimation of the dose-response curve
    - if sample size was calculated for estimation, power for signal detection will be high
- Under what circumstances might the data exhibit a dose-response of interest but the procedure fails to identify this?
  - Idea of MCP-Mod: Define „dose-responses of interest“ at the design stage
    - Design of the study (doses, sample size) can be chosen to be able to identify these
  - When a dose-response signal cannot be identified with MCP-Mod, the effect size of the drug is most likely smaller than anticipated

Issue 8a

Model selection: Using more than one model

- Is it plausible to select more than one model with which to continue development?
  - It is likely that model uncertainty will remain after completing Phase II
  - If uncertainty remains, more than one model might be kept for future use (especially if MCP-Mod used with model averaging)
- ... how a model averaging approach can improve over the use of a single model when multiple pre-selected models are found to be of interest?
  - Difference in interpretation
    - "a" single model vs. "weighted average" of >1 model
  - Average performance is rather similar; see e.g. plot of correct target dose interval probabilities from Bornkamp et al. (2007)



Issue 8b

Model selection: Challenges

- ... find that a model shape not included in the initial set of potential models actually perform better than a model in the initial set. Is it a realistic possibility? How can such situation be handled pragmatically in the framework of MCP-Mod?
  - For a reasonably broad candidate set often one model will be a good approximation
  - MCP-Mod just one component for the decision making in view of Phase III
- Describe the properties in situations when the selection of trial doses turn out to be flawed such that model selection is driven by a set of doses with zero or maximal effect?
  - Estimation of the increasing part of the dose-response curve (and target dose) challenging, inferences driven by the model assumptions
  - Important to quantify uncertainty (parameter estimates and models)
  - Response-adaptive designs may offer the opportunity to react accordingly

Issue 9

'interpolation' between doses and 'extrapolation' outside the dose range

- Discuss to what extent the procedure can support selection of a dose that has not been directly studied
  - Interpolation between doses is possible and encouraged
  - Extrapolation outside the dose range is discouraged
- Is inference restricted to the discrete set of doses used in the trial?
  - Traditional methods based on pairwise comparisons are not designed for extrapolation of information beyond the observed dose levels
  - MCP-Mod allows interpolation between doses under investigation
    - Recommend to always report uncertainty, e.g. on the "y-axis" (= effect estimates) or on the "x-axis" (= dose estimate)
    - Possibly accounting for multiplicity, e.g. use simultaneous confidence bands around dose response estimate instead of marginal confidence intervals at each dose

Issue 10

Increase in efficiency compared to traditional pairwise comparisons

- Does any increase in efficiency compared to traditional pairwise comparisons come at any cost to the developer, perhaps in terms of having less evidence to support for a particular dose level to take forward to Phase III?
  - Increase in efficiency by using modelling assumptions (i.e. prior information)
    - Testing and estimation gets optimized for realistic alternatives
    - Trade-off for unrealistic scenarios (e.g., zig-zag dose-response curve)
  - The dose to take forward to Phase III
    - Smoothing of dose-response estimates helps to safeguard against random highs (and lows), leading to a more robust planning for Phase III

Issue 11

Applicability ... without regard to therapeutic area or class of compound

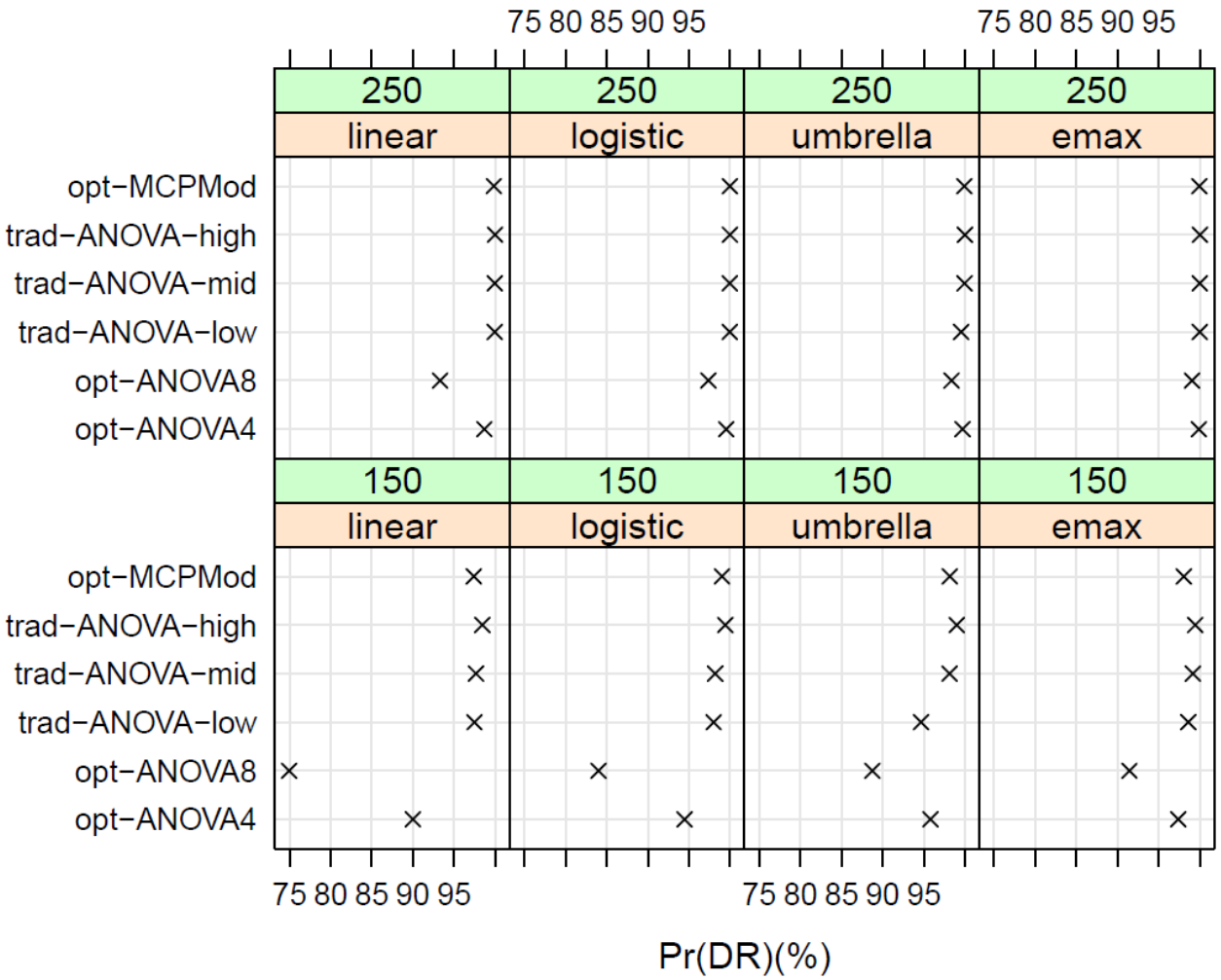
- MCP-Mod is applicable in any therapeutic area, since it essentially uses empirical dose-response models
- Discuss application in the context of dose selection that needs to consider both safety and efficacy
  - Any dose selection for Phase III requires safety / efficacy considerations
  - Need to understand safety / efficacy dose response relationships to estimate MED / MSD and thus the therapeutic window
  - Safety dose-response modelling less common, but MCP-Mod could be applied equally well
- Is there any quantitative approach to the synthesis of two univariate models, one for a key efficacy marker or parameter and one for safety?
  - One possibility is to derive a clinical utility index (CUI) that combines safety and efficacy information in one variable
  - In practice, derivation of CUI is quite hard
  - Limited experience at Novartis, but in principle MCP-Mod could be applied (unimodal shapes!)

Backup slides

Simulation Results Issue 1

Simulation Results

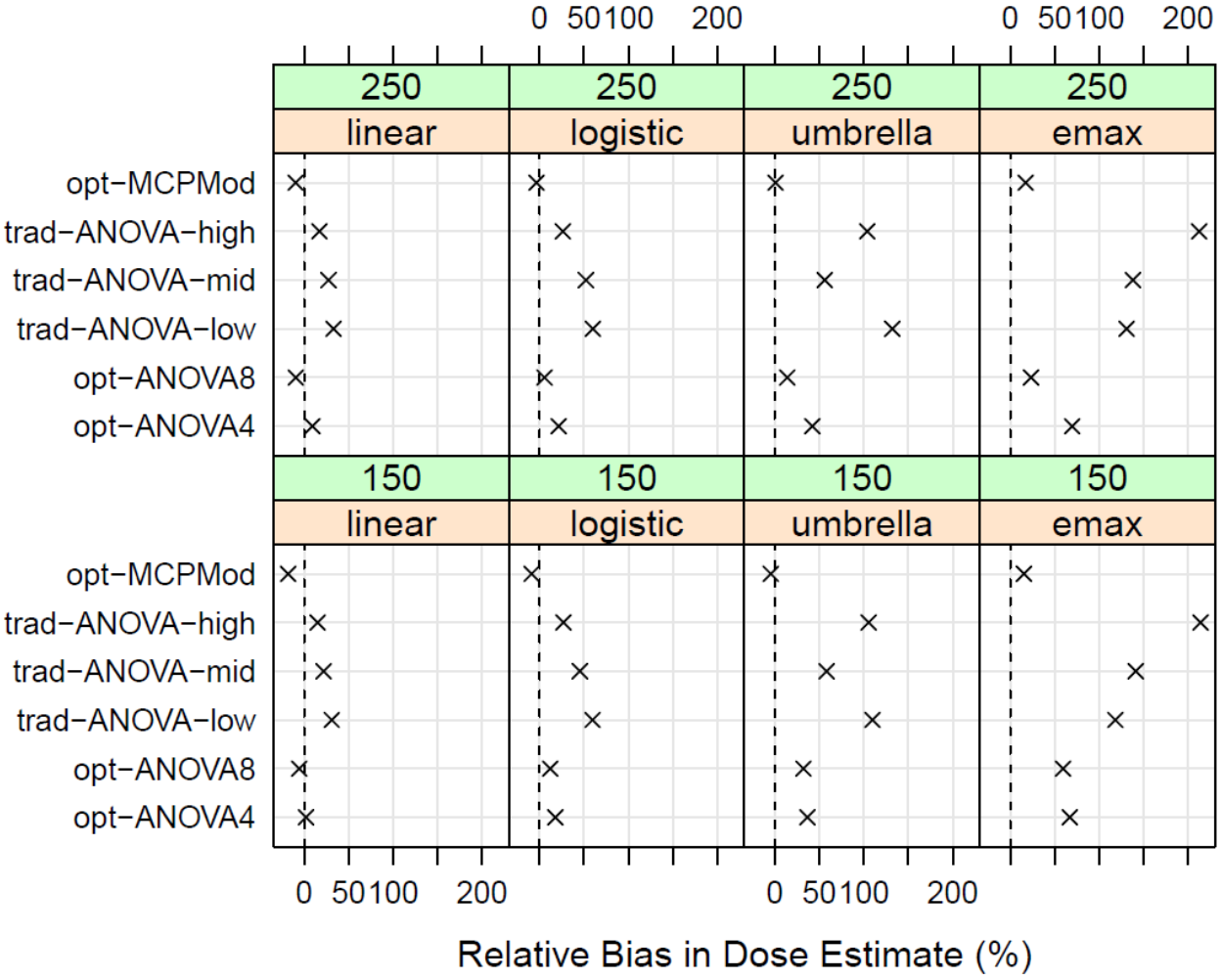
Power to detect dose-response



NOVARTIS

Simulation Results

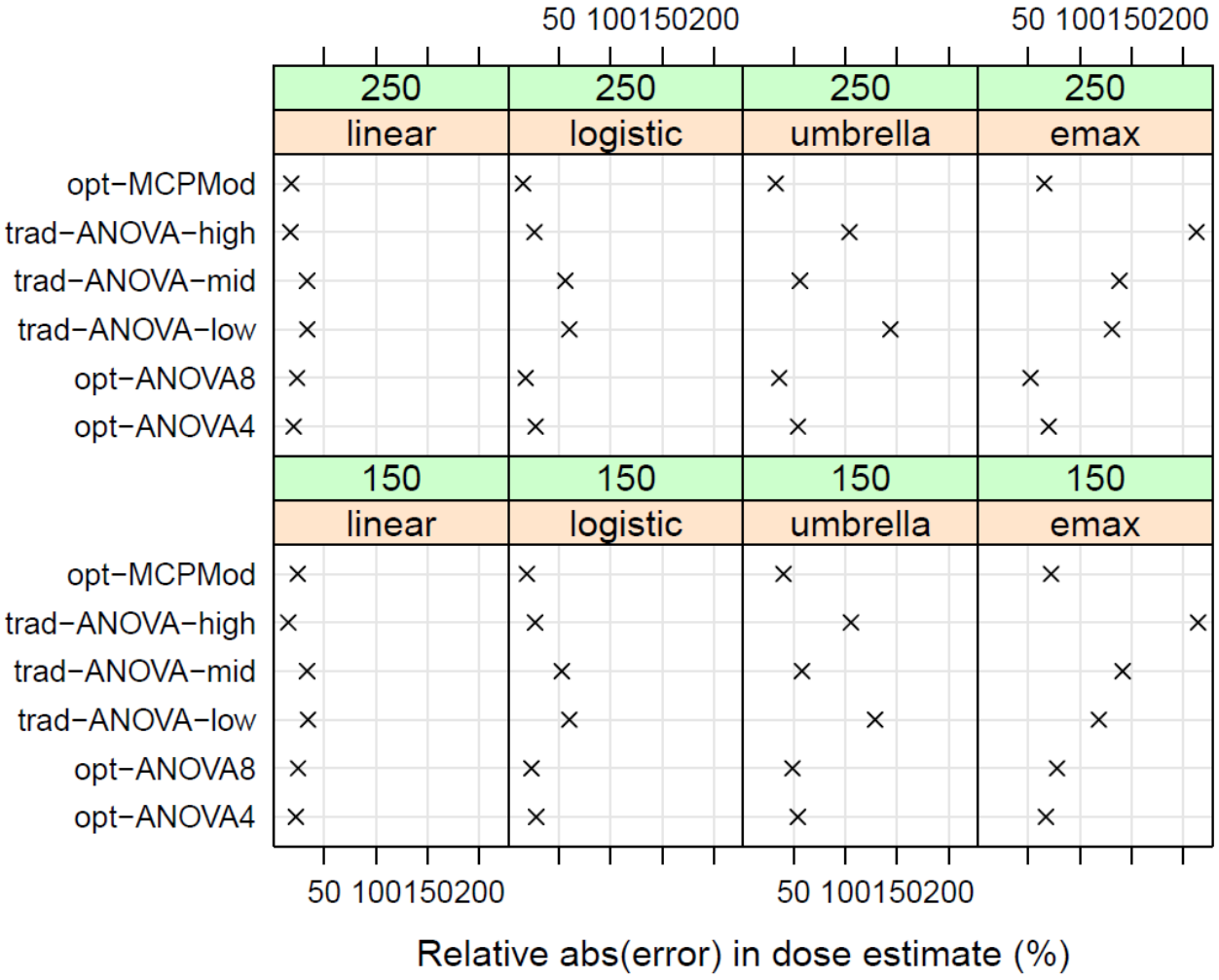
Relative Bias in dose estimate



NOVARTIS

Simulation Results

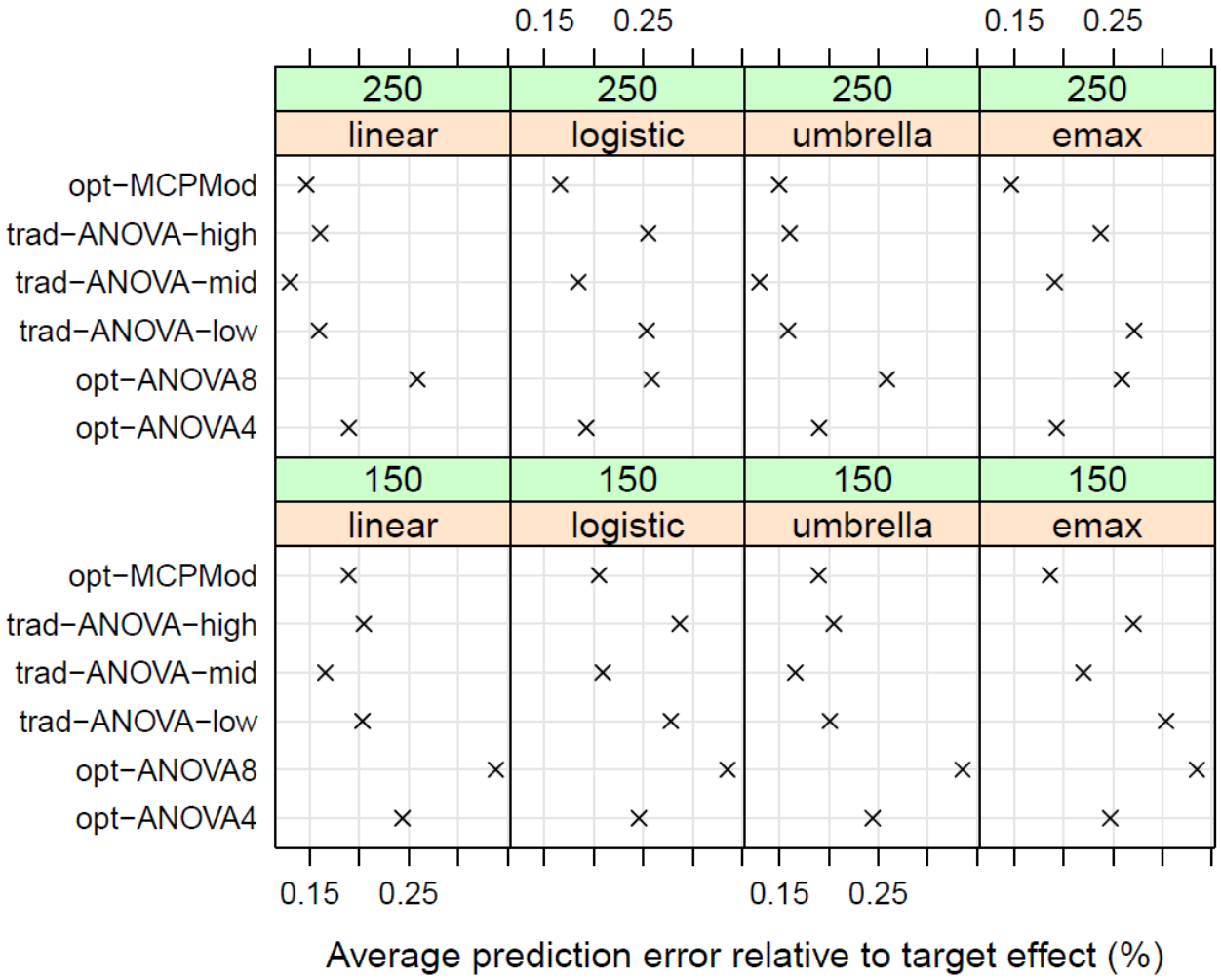
Relative absolute error in dose estimate



NOVARTIS

Simulation Results

Average prediction error in estimating the dose-response function



NOVARTIS

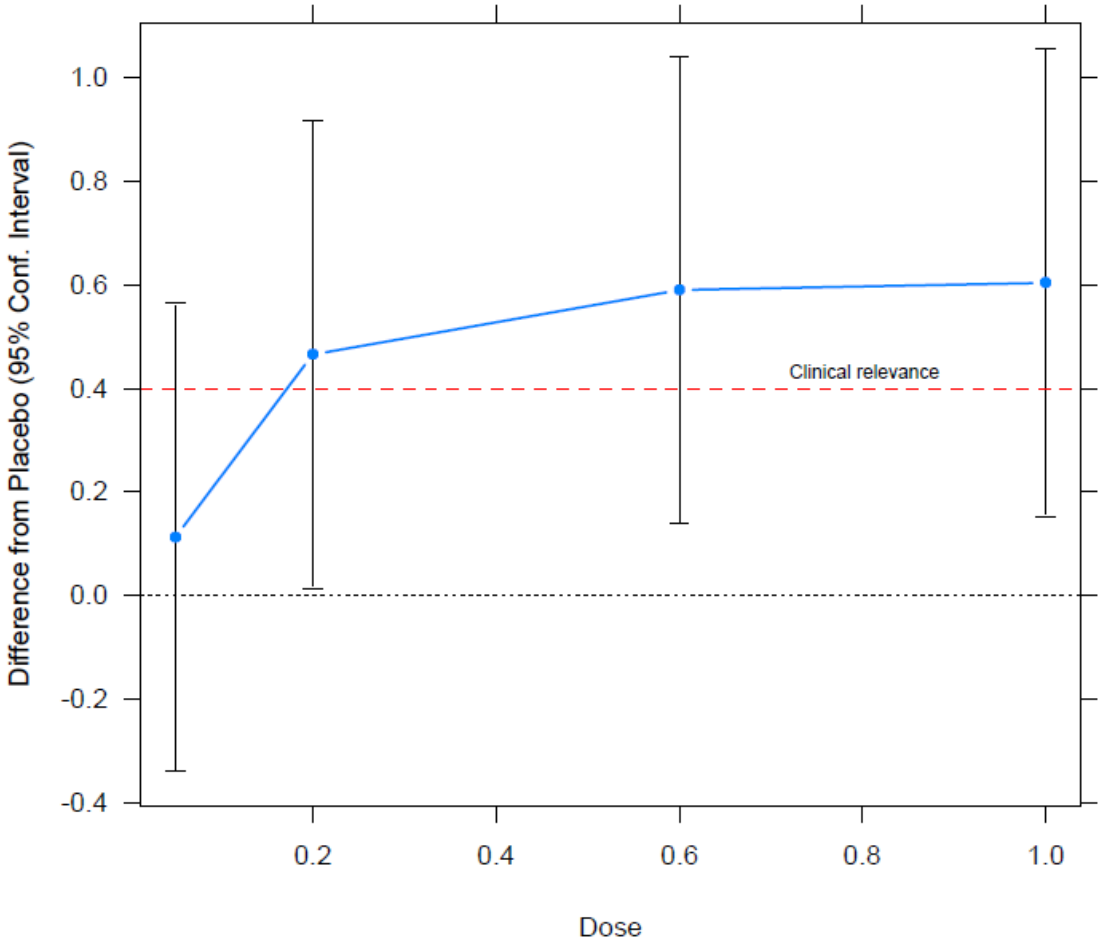
Case Example

Example

- 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
- Normally distributed, homoscedastic primary endpoint
- Planned analysis: Fixed sequence test that preserves type I error at 5% two-sided level
- Conclusion: Top three doses are significantly better than placebo.

Example

Which dose should be considered the MED?

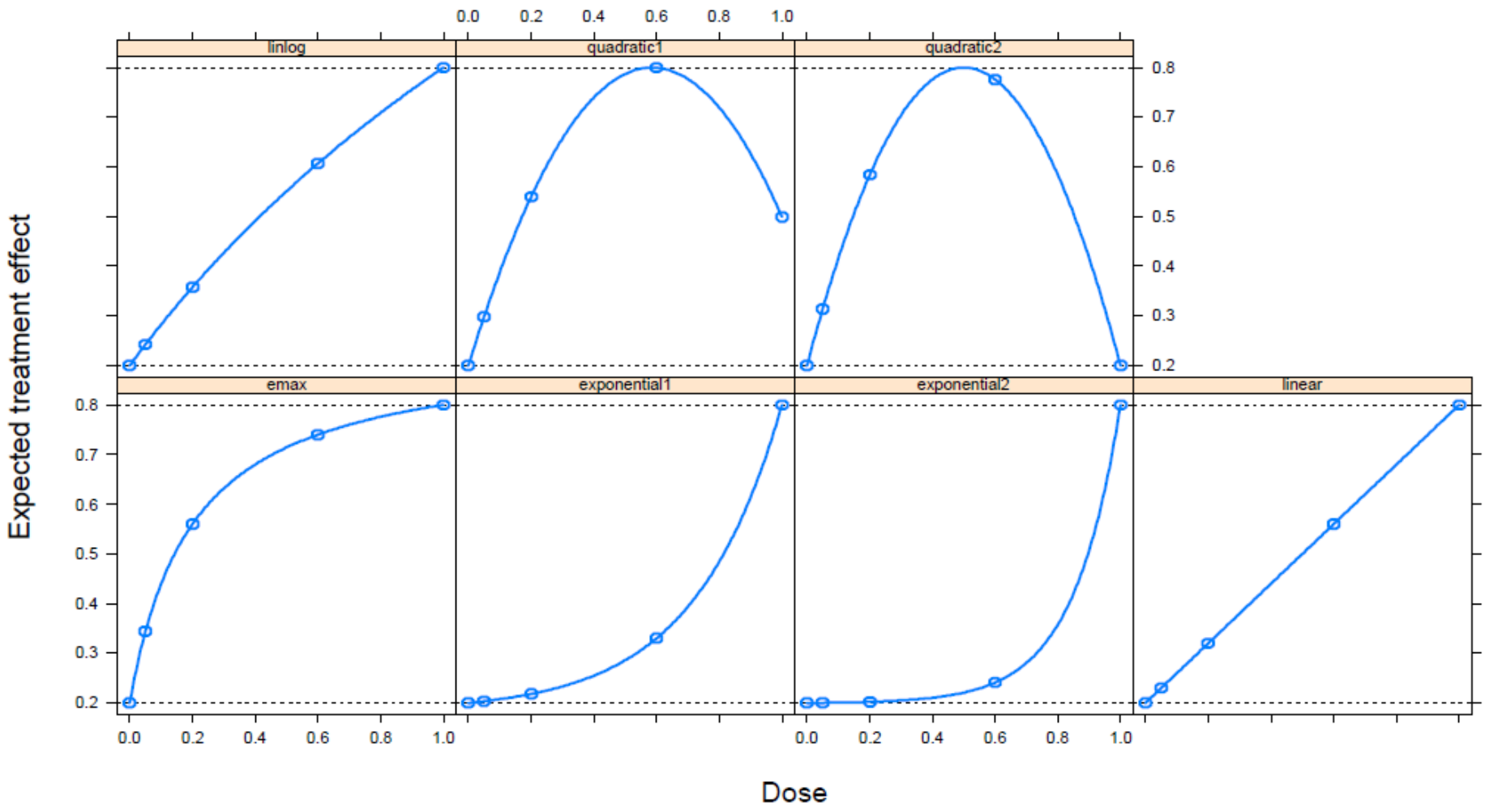


NOVARTIS

Example

- First step of MCP-Mod: Set of candidate models
- Propose M dose-response models at planning stage to describe potential outcomes
- Model uncertainty directly acknowledged
- Requires strong collaboration with clinical team
  - Input based on available information (PK data, historical data)

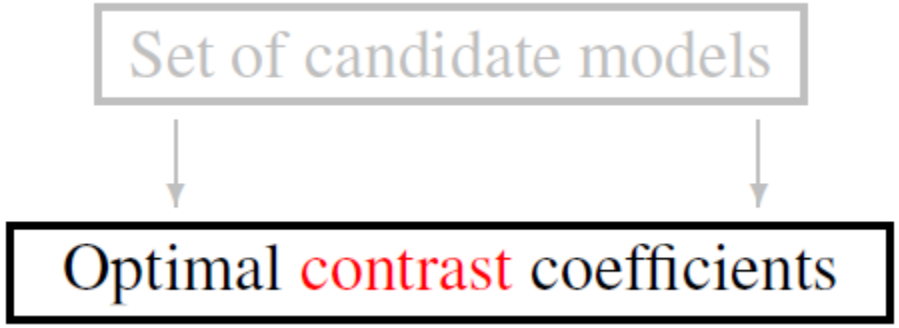
Example



NOVARTIS

Example

- Second step of MCP-Mod:



- Each model will be tested using a contrast test with optimally chosen weights
- For each dose response model, contrast weights are chosen to maximize power in detecting that model if it is true

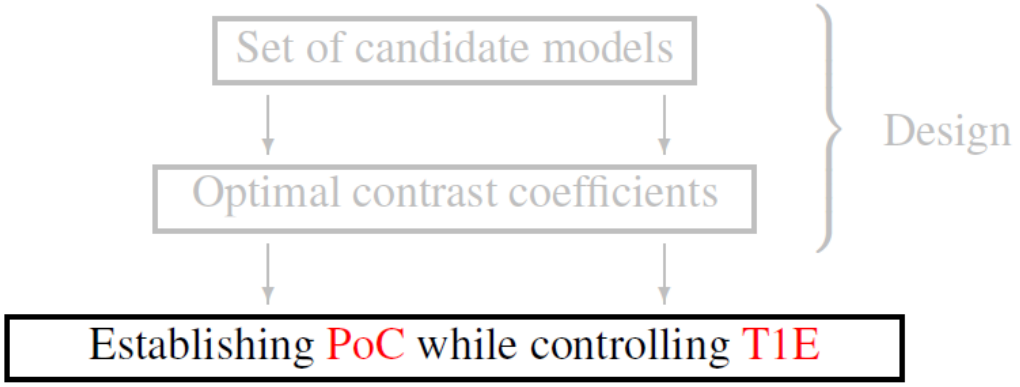
Example

Optimal contrasts:

	Candidate Models						
Dose	Linear	Emax	Linlog	Exp 1	Exp 2	Quad 1	Quad 2
0	-0.44	-0.64	-0.47	-0.29	-0.24	-0.57	-0.42
0.05	-0.38	-0.36	-0.39	-0.29	-0.24	-0.36	-0.20
0.2	-0.20	0.06	-0.16	-0.26	-0.24	0.16	0.33
0.6	0.27	0.41	0.32	-0.04	-0.17	0.71	0.71
1	0.74	0.53	0.70	0.87	0.89	0.07	-0.42

Example

- Third step of MCP-Mod:



- Each model will be tested using a contrast test with optimally chosen weights
- For each dose response model, contrast weights are chosen to maximize power in detecting that model if it is true

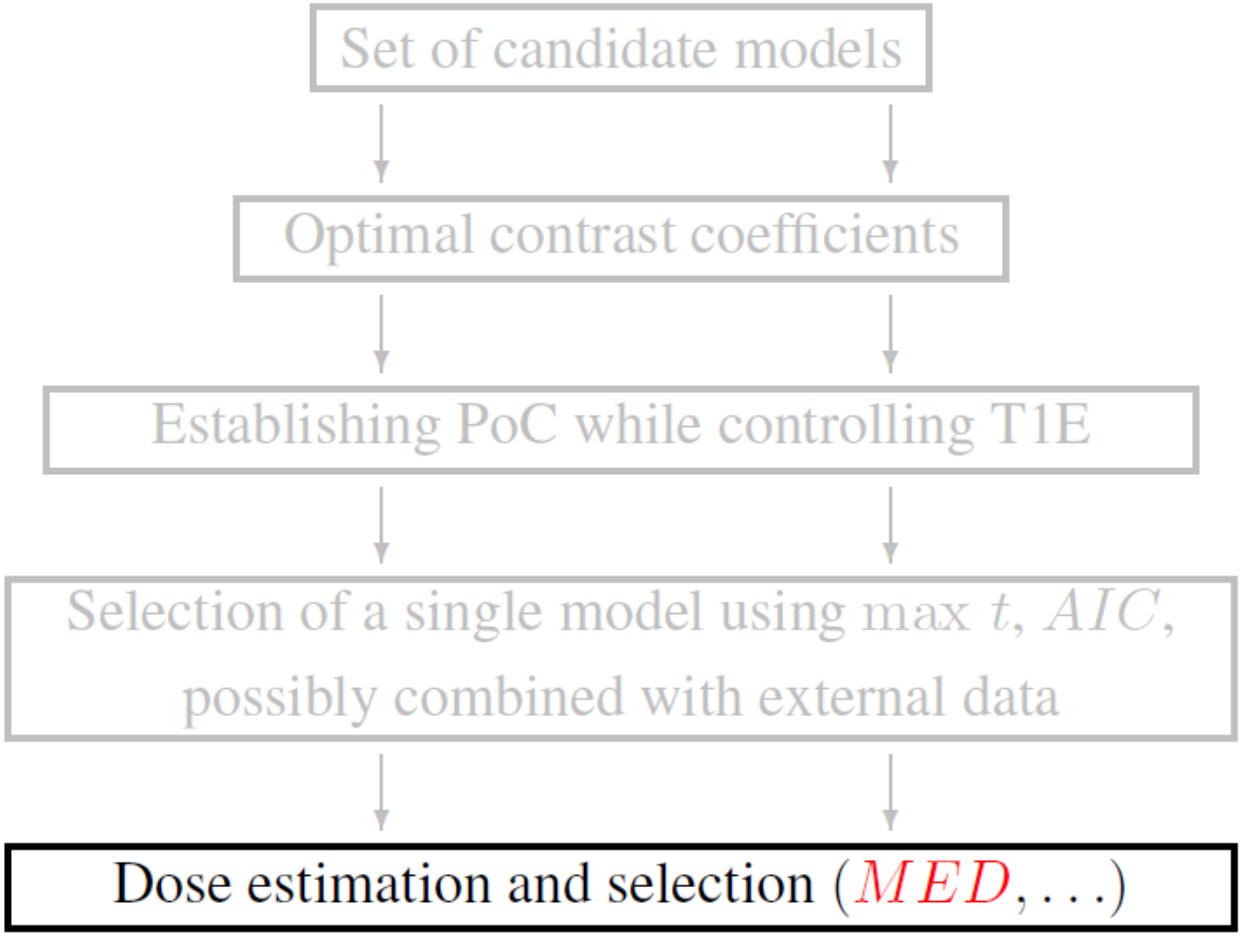
Example

- Significant result is established if the maximum contrast test statistics (across all models) is larger than the critical value, i.e.
  - max $T_m > crit_{1-\alpha}$
- All models with $T_m > crit_{1-\alpha}$ are kept for possible use in dose-response modeling
- If max $T_m < crit_{1-\alpha}$ no significant dose-response
- Here $crit_{0.95} = 2.15$

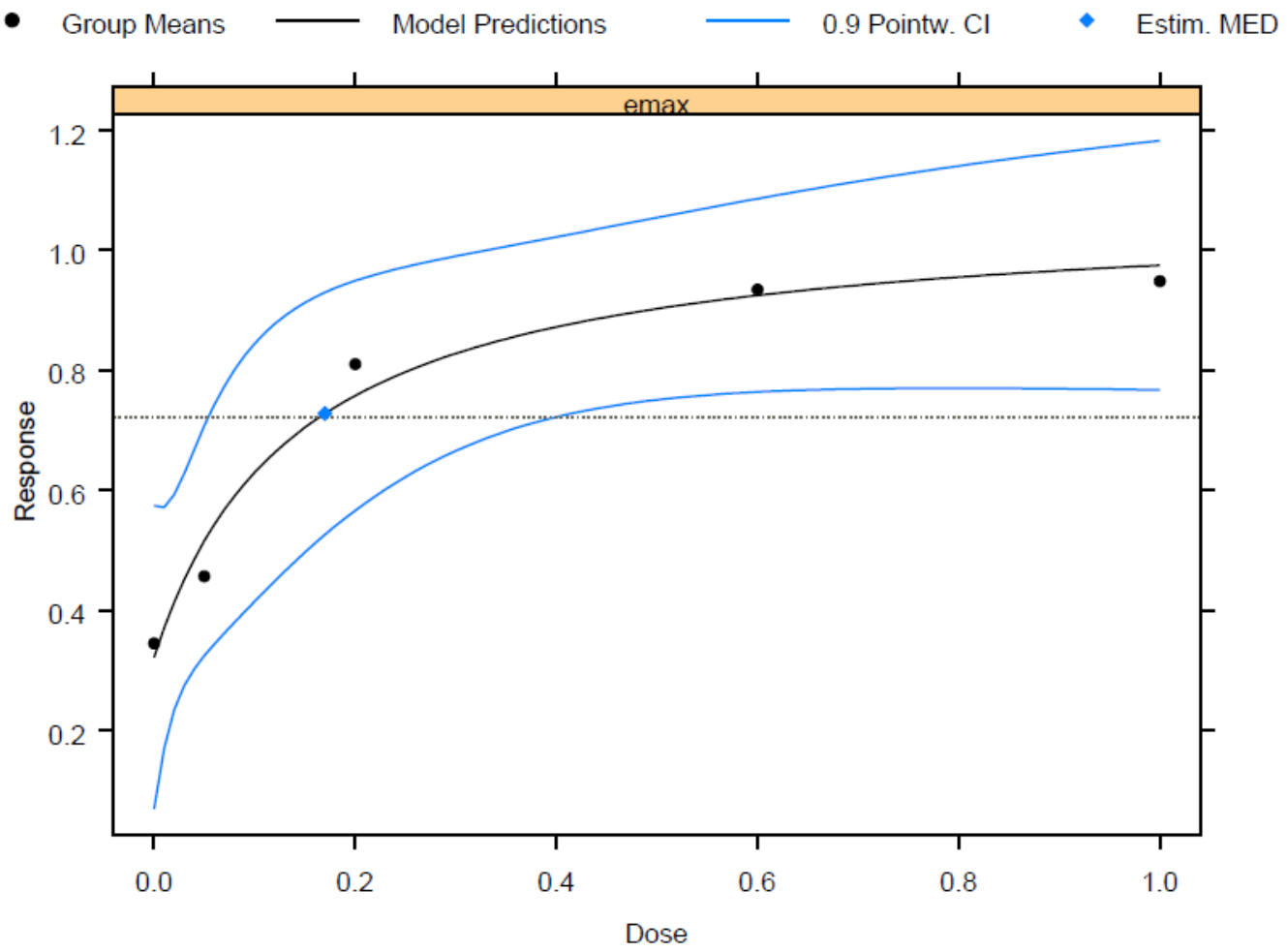
Example

Contrast	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

Example

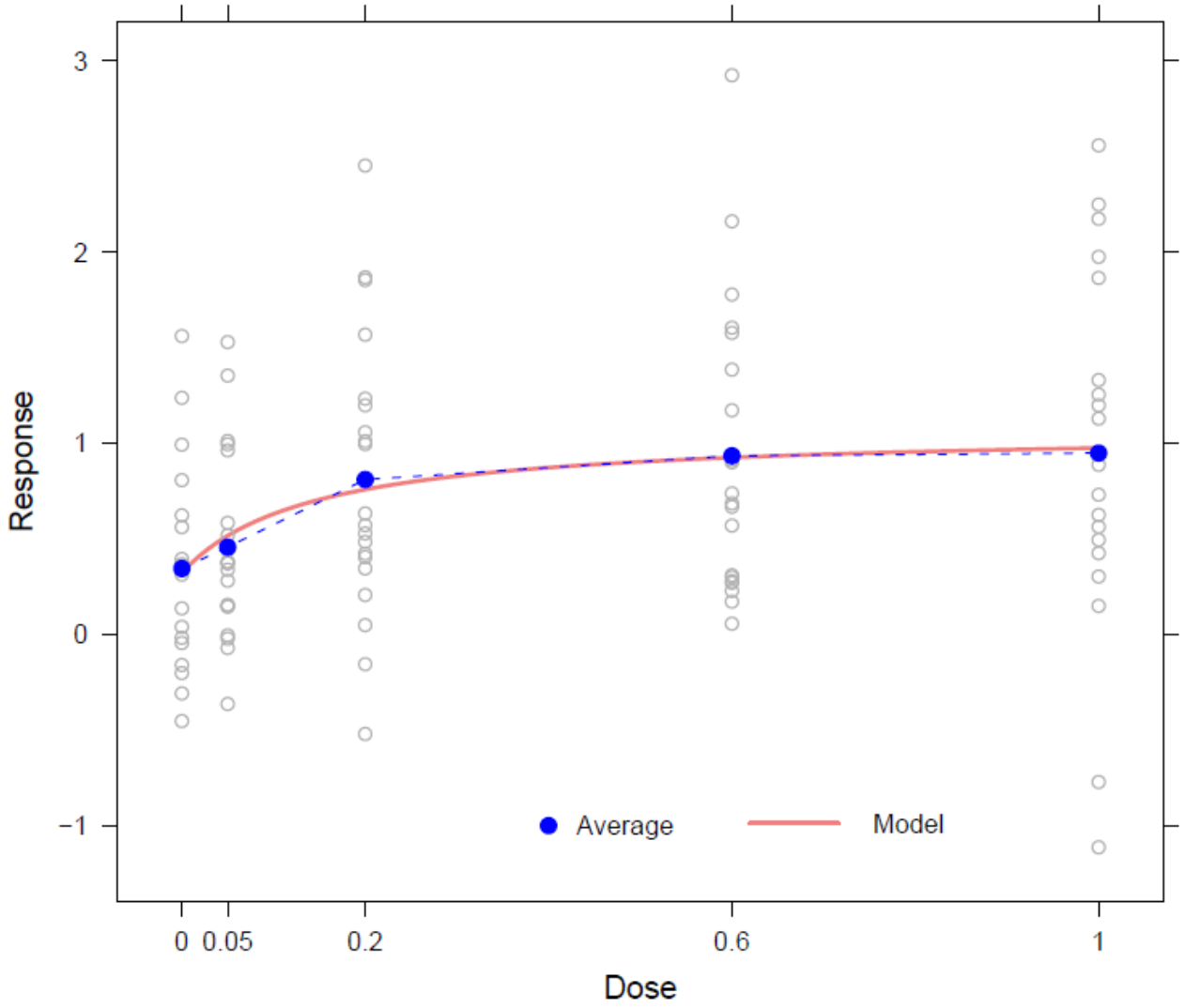

Set of candidate models
Optimal contrast coefficients
Establishing PoC while controlling T1E
Selection of a single model using max t, AIC, possibly combined with external data
Dose estimation and selection (MED, . . .)

Example



NOVARTIS

Example



NOVARTIS