Considerations On Futility Rules For Adaptive Dose-Finding Designs

Silke Jörgens, Tobias Mielke

Futility – Where Do We Stop?

- Futility stopping in two-armed confirmatory trials:
  - Stop the trial if interim results are worse than a predefined threshold (e.g., in terms of first stage *p*-value)
- Futility stopping in multi-armed confirmatory trials:
  - Stop the trial if global interim results are worse than a predefined threshold (e.g., in terms of first stage *p*-value for global hypothesis test)
  - Stop the trial if no treatment arm fullfills selection criteria (in case of conditional selection $\rightarrow$ implied futility stop)
  - Stop specific treatment arms
- Futility stopping in dose-finding trials?

Goals for Interim Analyses in Dose-Finding Trials

- Goals for interim analyses in dose-finding trials are generally different from those in confirmatory trials
- There is generally no early stopping for efficacy but rather…
  - Expanding a study if there is evidence for a drug-related effect (e.g., top-down design)
  - Gathering information on the dose-response relation in order to expand recruitment on the most promising doses
- In either case, it is desirable if a futile development programme does not advance to the second stage

Futility Stopping in Dose-Finding Trials

- Futility Stopping in dose-finding trials can have a different implication than it has in confirmatory trials
- Instead of stopping one trial only, the entire development programme can be affected
- Reasons for futility stopping in such trials can be
  - No evidence against a flat dose-response relationship
  - Evidence against a commercially viable effect

Motivating Example

- Development of a new compound in the Osteoarthritis indication
- Two-stage design is planned:
  - For the first stage, placebo and up to four active doses are available
  - Second stage can add one active dose if no futility stopping after first stage
  - 1.1 on the target scale is considered a commercially viable effect

- Which futility rules to choose?
- Does the number of doses affect the futility stop?
- What about first stage sample size?

Approach 1: Pairwise Testing For Any Difference

- Derived from later stage multiple comparison trials with limited number of doses
- Apply a futility boundary of α_0 to the global intersection null hypothesis

$$H_0 = \cap H_{0i},\ H_{0i}: \mu_i \leq \mu_0$$

- Multiplicity correction needed (here, Dunnett considered throughout)
- What is the impact of the number of doses in the first stage?
  - E.g., comparison of top-down approach vs equal allocation to all available doses

Approach 2: Model-Based Testing for Any Difference

- Uses MCPMod approach:
  - Prespecify a number of candidate models
  - Derive optimal contrasts for these models in dependence of the planned allocation to the first stage doses
- Apply a futility boundary to the resulting multiple contrast test for the existence of a drug-related effect
- Multiplicity penalty is lower than in Approach 1

Approach 1 & Approach 2

- Both approaches control the probability of continuing development when there is no drug-related effect

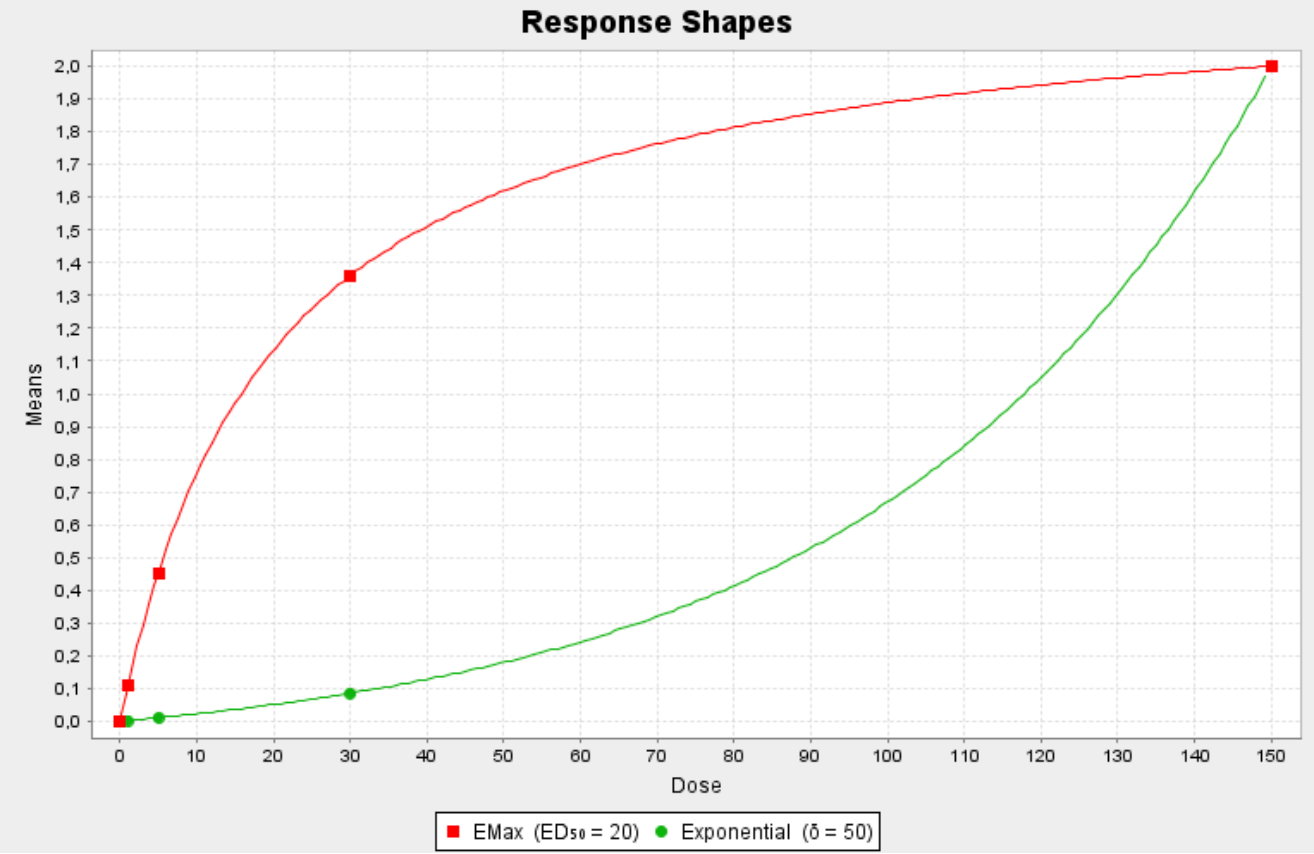
- Probability of continuing development when there is a true benefit is not controlled

- Impacting the latter probability requires appropriate sample size for the first stage

Revisit Example

- Active doses for first stage are 1 mg, 5 mg, 30 mg, 150 mg
- Total sample size for first stage is set to 120 patients (including placebo patients)
- For simulations, we assume a maximum effect of 2 on the target scale (standard deviation: 4) and two different true dose-response shapes:
  - An EMax model with ED_{50} at 20 mg ($\eta(d) = e_{max} \times d / (d + 20)$)
  - An exponential model with δ=50 ($\eta(d) = e_{max} \times (\exp\{d/50\} - 1)$

Example Response Shapes



Example Revisited

– Consider the following possible allocations in the first stage:

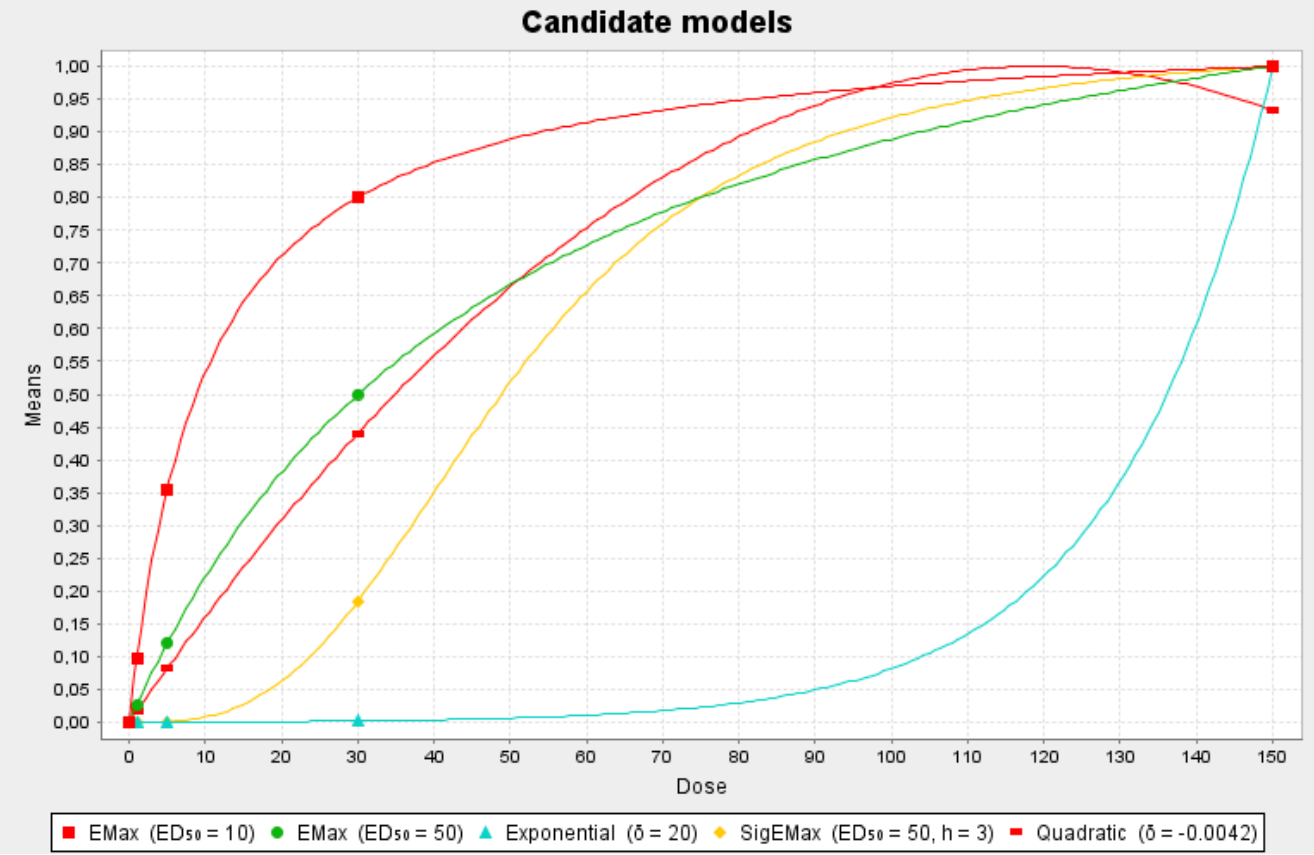
	PLA	1 mg	5 mg	30 mg	150 mg
D1	0.5	0	0	0	0.5
D2	0.2	0.2	0.2	0.2	0.2
D3	0.4	0	0	0.2	0.4

Example Revisited

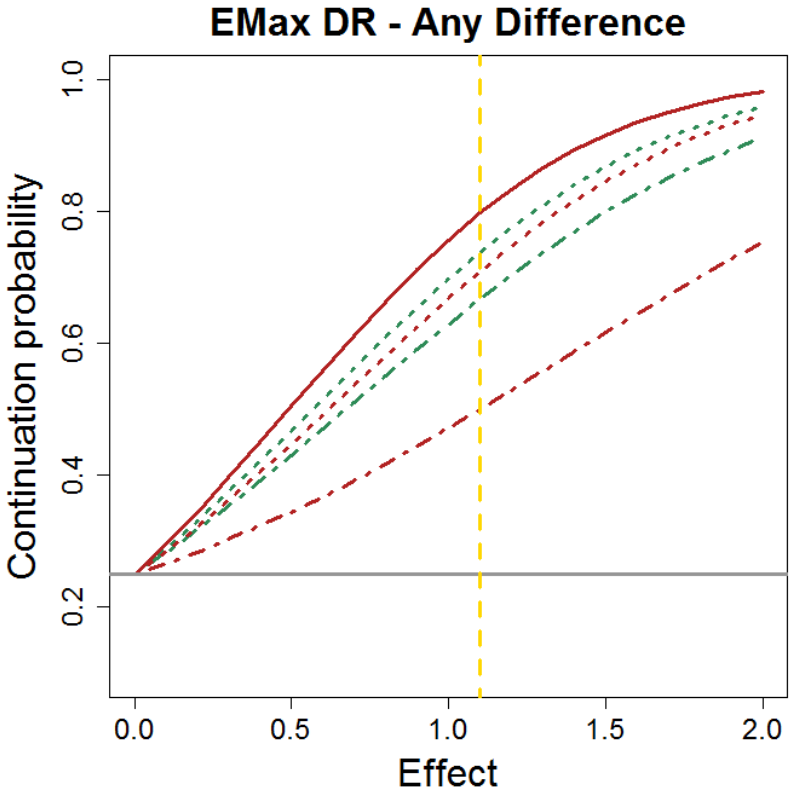
– For the model-based contrasts, assume the following candidate dose-response shapes:

Model	Parameter(s)
EMax	$ED_{50} = 10$
EMax	$ED_{50} = 50$
Sigmoid EMax	$ED_{50} = 50$, $h = 3$
Exponential	$\delta = 20$
Quadratic	$\delta = -1/240$

Candidate Models

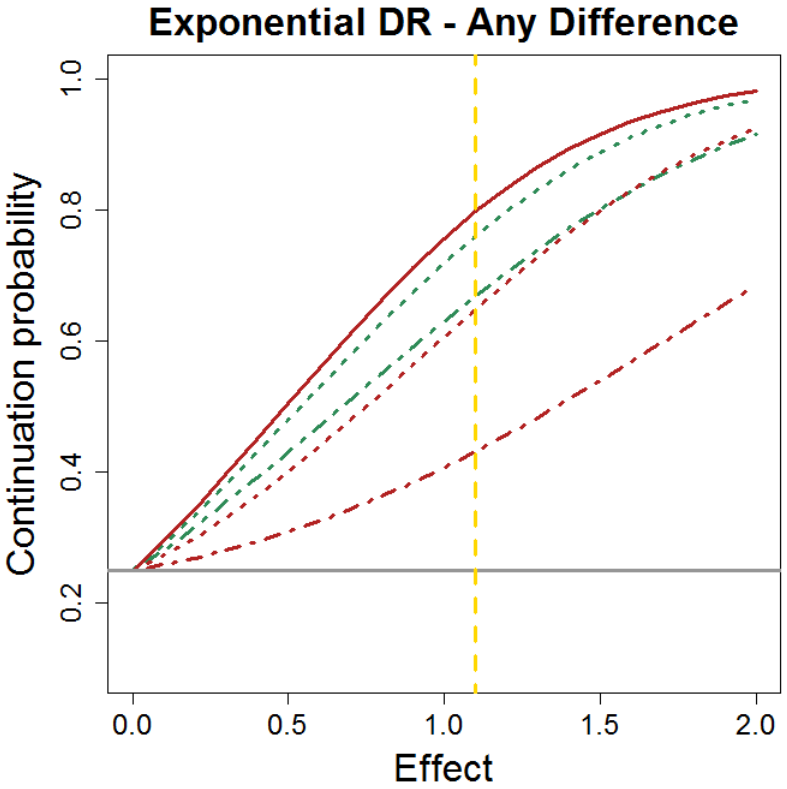


Approach 1 & 2: True Shape EMax



- Red: Pairwise comparisons
- Green: Model-based contrasts
- Solid: D1 (top down)
- Dash-dot: D2 (equal)
- Dotted: D3
- Grey reference line: α_0
- Yellow reference line: TV

Approach 1 & 2: True Shape Exponential



- Red: Pairwise comparisons
- Green: Model-based contrasts
- Solid: D1 (top down)
- Dash-dot: D2 (equal)
- Dotted: D3
- Grey reference line: α_0
- Yellow reference line: TV

Approach 3: Pairwise Testing for Target Value

- Complementary approach: Control probability of a development stop in case of true benefit

- Consider testing the following null hypothesis:

$$H_0: \{\mu_1 \geq \mu_0 + TV\} \cup \{\mu_2 \geq \mu_0 + TV\} \cup \ldots \cup \{\mu_k \geq \mu_0 + TV\}$$

- Stop for futility if H_0 can be rejected at level α_0 (implies that true effect for all doses is below TV)

- No multiplicity adjustment required

Approach 4: Model-Based Testing for Target Value

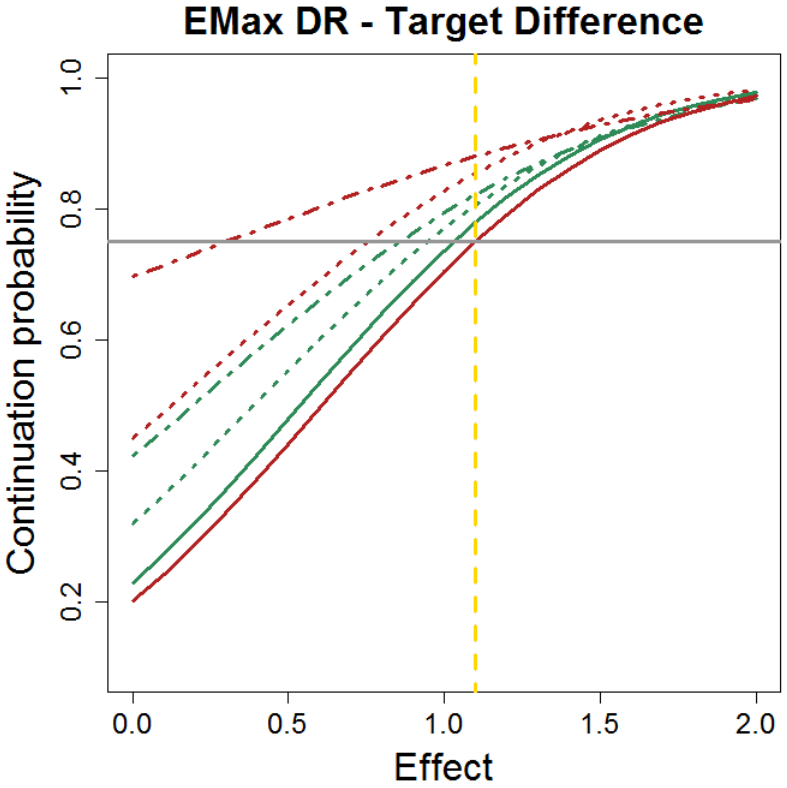
- The model-based approach uses the model-based estimators for the maximum treatment effect

- $\widehat{\Delta}_g$ denotes the estimated maximum treatment effect for model g

- Then recommend stopping for futility if

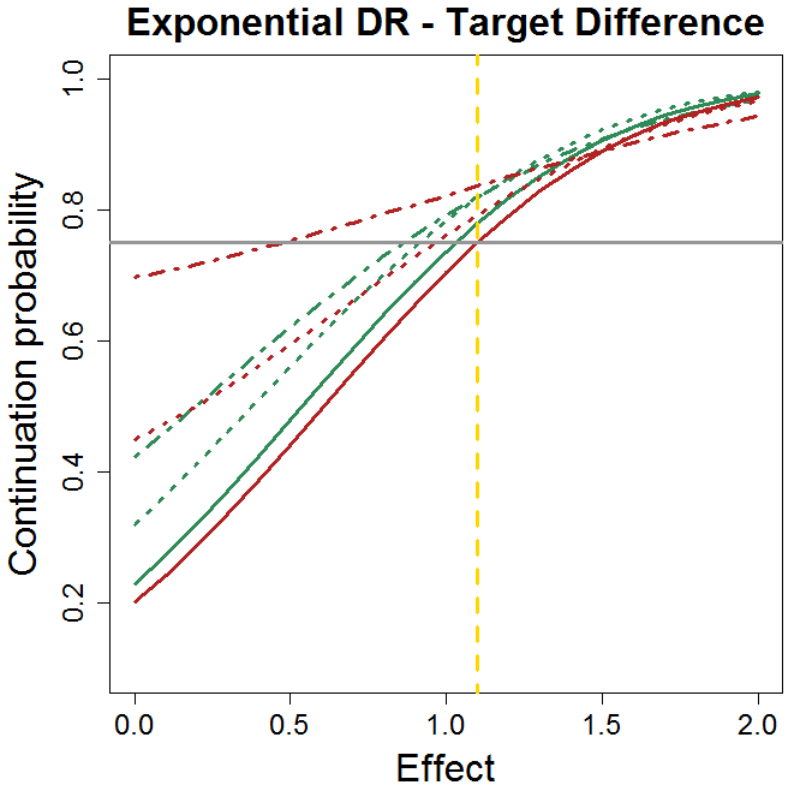
$$\min_{g=1,\ldots,5} \sqrt{N_1} \frac{TV - \widehat{\Delta}_g}{\sqrt{\sigma^2 {c_g}^T W^{-1} c_g}} > z_{1-\alpha_0}$$

Approach 3 & 4: True Shape EMax



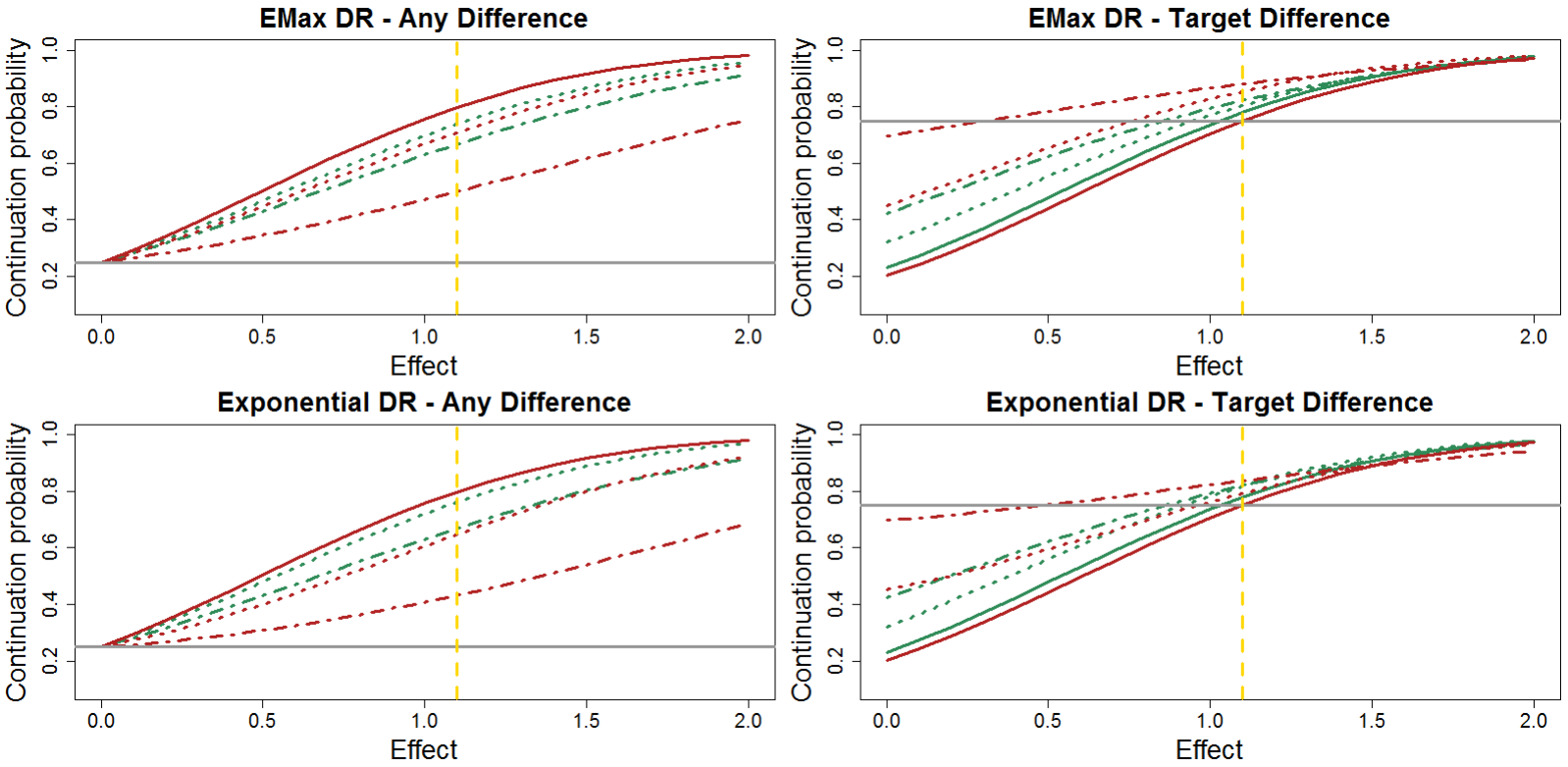
- Red: Pairwise comparisons
- Green: Model-based contrasts
- Solid: D1 (top down)
- Dash-dot: D2 (equal)
- Dotted: D3
- Grey reference line: α_0
- Yellow reference line: TV

Approach 3 & 4: True Shape Exponential



- Red: Pairwise comparisons
- Green: Model-based contrasts
- Solid: D1 (top down)
- Dash-dot: D2 (equal)
- Dotted: D3
- Grey reference line: α_0
- Yellow reference line: TV

All Together…



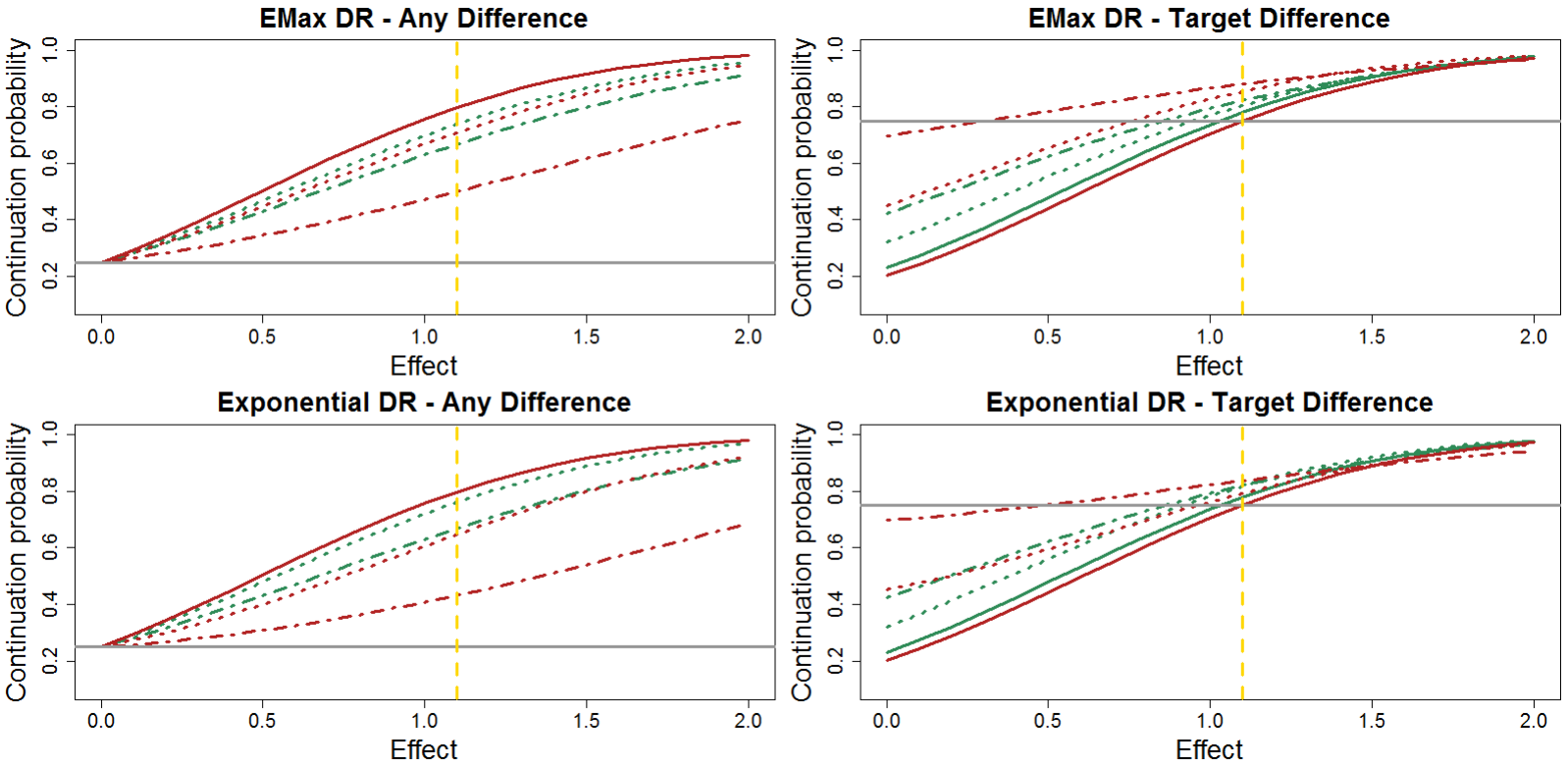
Interim Summary

- Hard requirements are fulfilled in both cases:
  - Stopping probability is controlled at α_0 in the Any Difference case
  - Continuation probability is controlled at 1- α_0 in the Target Value case

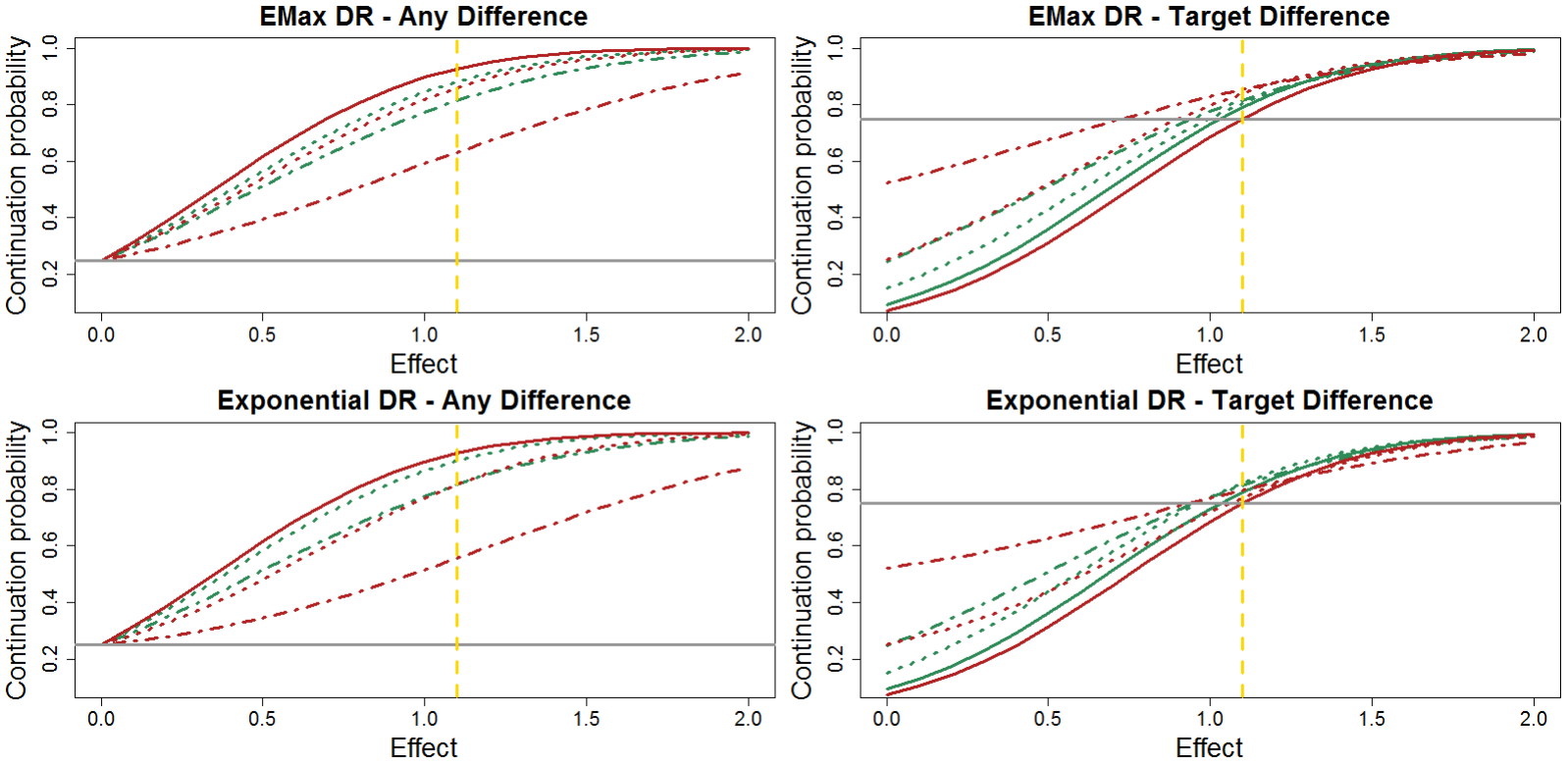
- With the considered sample size, curves for Target Value case are too smooth

- For investigation of sample size impact, consider sample sizes of 120 (as before), 240 and 600

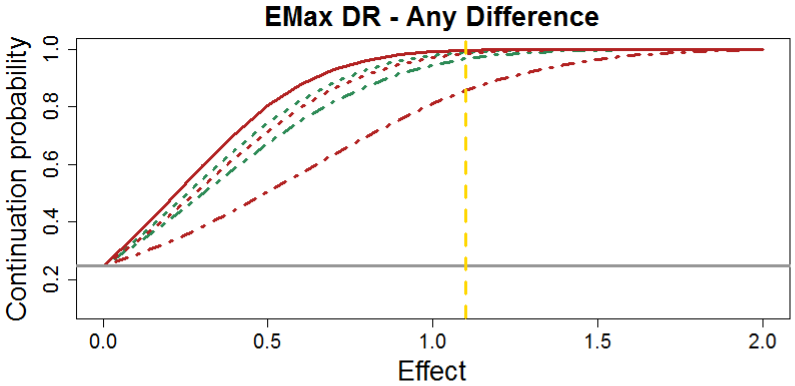
Total N 120

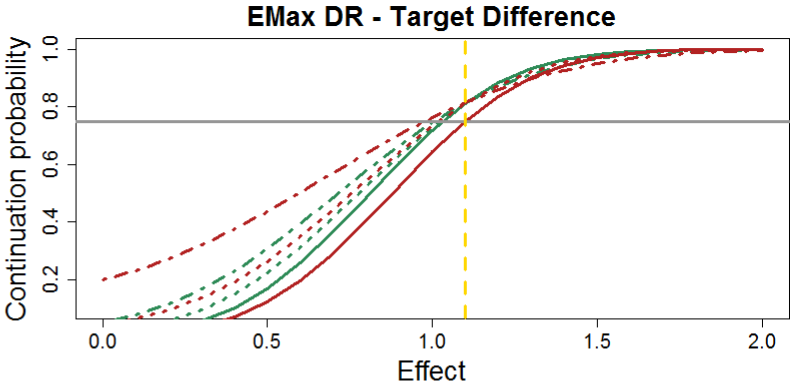

EMax DR - Any Difference
EMax DR - Target Difference
Exponential DR - Any Difference
Exponential DR - Target Difference
Continuation probability
0.2
0.4
0.6
0.8
1.0
0.0
0.5
1.0
1.5
2.0
Effect

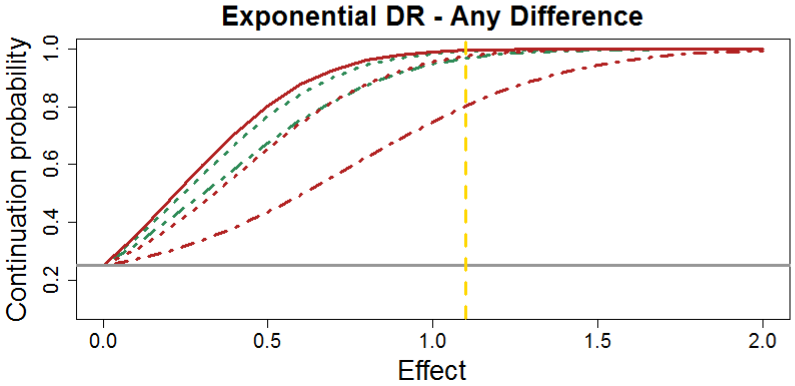
Total N 240

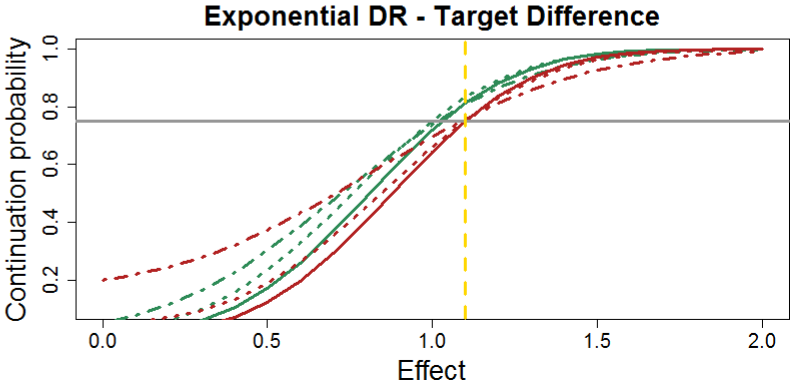

EMax DR - Any Difference
EMax DR - Target Difference
Exponential DR - Any Difference
Exponential DR - Target Difference
Continuation probability
0.2
0.4
0.6
0.8
1.0
0.0
0.5
1.0
1.5
2.0
Effect

Total N 600


EMax DR - Any Difference
Continuation probability
0.2
0.4
0.6
0.8
1.0
0.0
0.5
1.0
1.5
2.0
Effect


EMax DR - Target Difference
Continuation probability
0.2
0.4
0.6
0.8
1.0
0.0
0.5
1.0
1.5
2.0
Effect


Exponential DR - Any Difference
Continuation probability
0.2
0.4
0.6
0.8
1.0
0.0
0.5
1.0
1.5
2.0
Effect


Exponential DR - Target Difference
Continuation probability
0.2
0.4
0.6
0.8
1.0
0.0
0.5
1.0
1.5
2.0
Effect

Summary

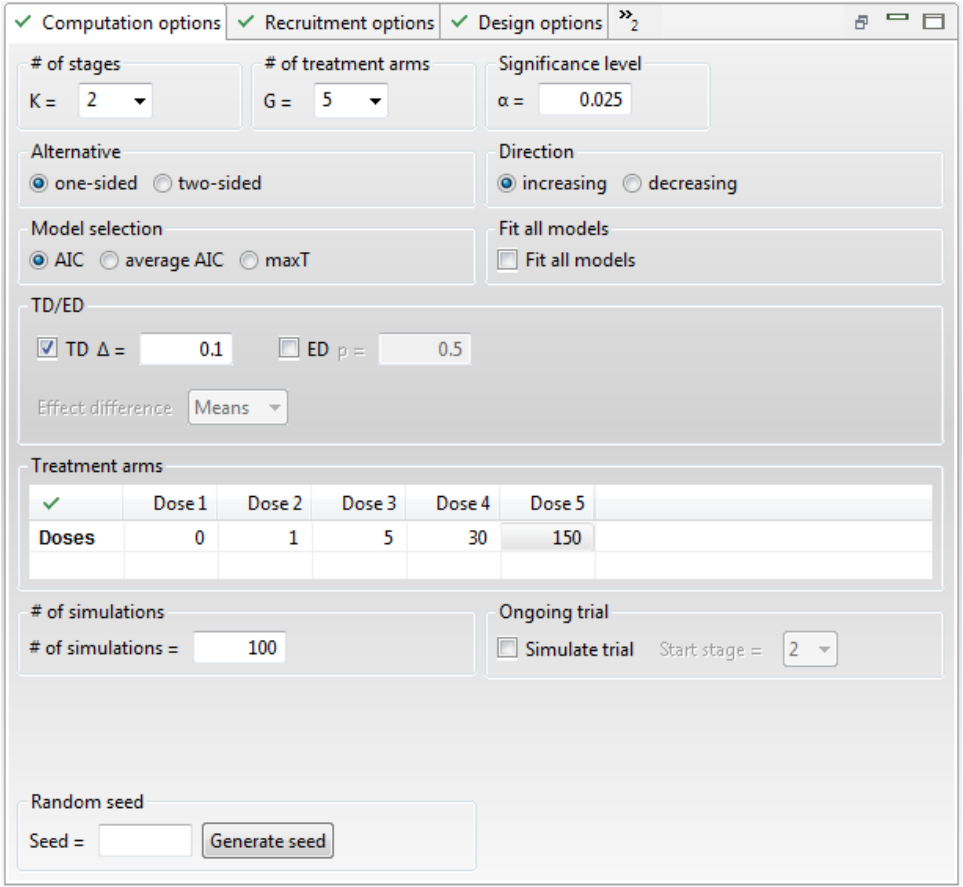
- Sensitivity of futility stopping boundaries depend on number of doses tested and sample size
- Pairwise comparison procedures suffer from multiplicity penalty (Any Difference) and from lack of power for individual hypotheses (Target Value)
- The „Target Value“ approach can improve identification of promising developments
- Model based techniques can improve discrimation between promising and futile developments, especially in the case of many doses
- Increasing the first stage sample size increases the continuation probability uniformly in the Any Difference case but increases the selectivity in the Target Value case

Software Solution

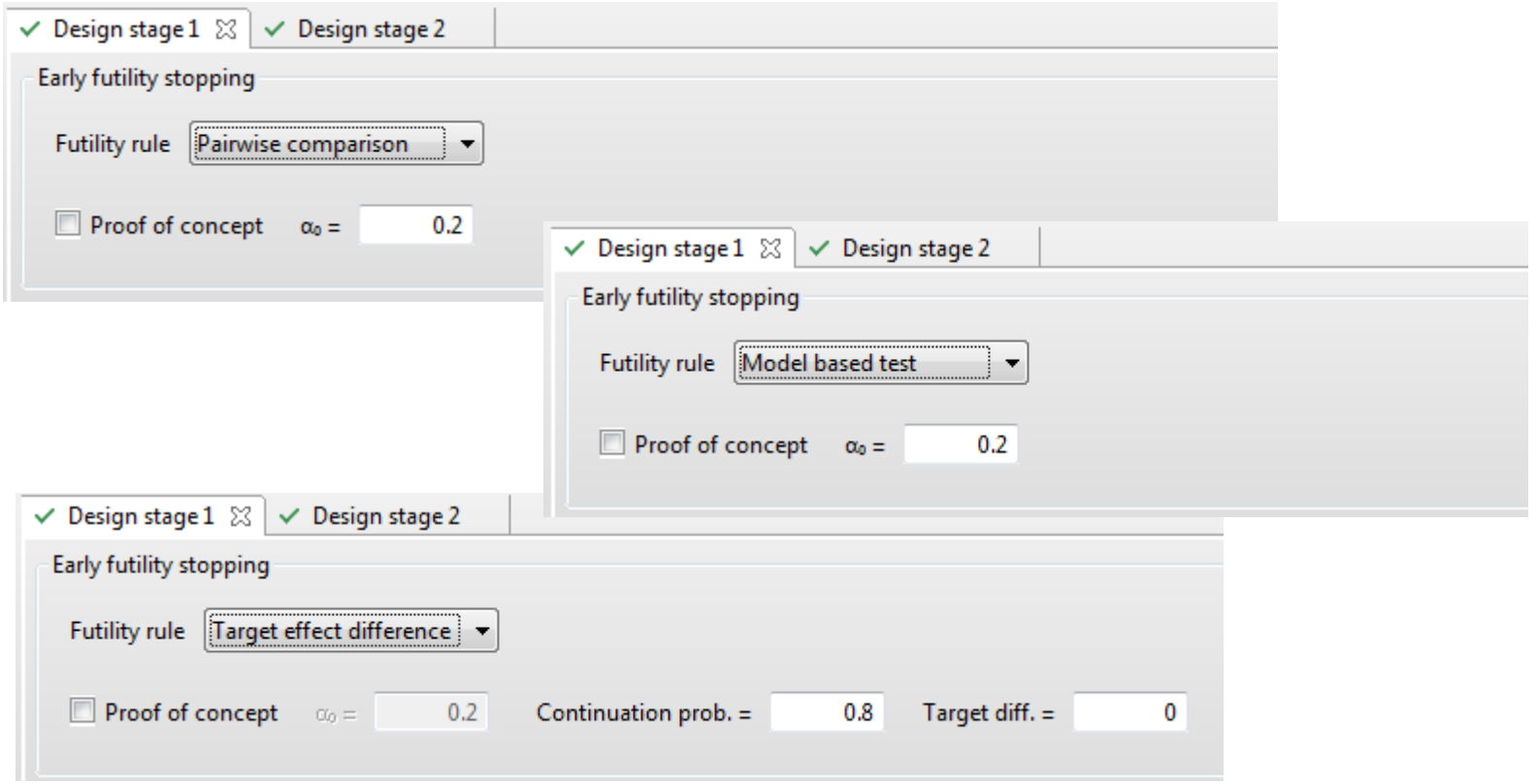
- ADDPLAN DF 4.0 partly incorporates these methods:
  - Multi stage designs available in simulation module for MCPMod and Nonlinear Regression

- Options for futility stopping are
  - Pairwise comparisons at α_0
  - Model-based test at α_0
  - Test for target effect difference at α_0

Multi Stage Designs



Design Options for Stage 1



References

- ADDPLAN, INC., an Aptiv Solutions Company. ADDPLAN DF - Adaptive Designs, Plans and Analysie - Dose Finding Module. Aptiv Solutions, Cologne, Germany 2014
- Berger, R. and Hsu, J., Bioequivalence Trials, Intersection-Union Tests and Equivalence Confidence Sets. Statistical Science, (11): 283-319, 1996
- Bornkamp, B., Pinheiro, J. and Bretz, F., MCPMod: An R Package for the Design and Analysis of Dose Finding Studies. Journal of Statistical Software, (29):1-23, 2010
- Lalonde, R., Kowalski, KG., Hutmacher, MM., Ewy, W., Nichols, DJ., Milligan, PA., Corrigan, BW., Lockwood, PA., Marshall, SA., Benincose, LJ., Tensfeld, TG., Parivar, K., Mantea, M., Glue, P., Koide, H., Miller, R., Model-based drug development. Clinical pharmacology and therapeutics, (82):21-32, 2007
- Mielke, T., and Dragalin, V., Two-stage designs in dose-finding, to appear in Handbook of methods for designing and monitoring dose finding trials (eds. A Iasonosa, J. O'Quigley, B. Bornkamp and A. Hartford), CRC Press, Taylor and Francis, 2017

Thank you!

iconplc.com