Operating characteristics of clinical trials with borrowing from external data: one-arm and hybrid control arm trials

Annette Kopp-Schneider[(1)], Manuel Wiesenfarth[(1)], Vivienn Weru[(1)], Leonhard Held[(2)] and Silvia Calderazzo[(1)]

[(1)]Division of Biostatistics, German Cancer Research Center, Heidelberg, Germany

[(2)]Epidemiology, Biostatistics and Prevention Institute, University of Zurich, Switzerland

Motivation

- Sample sizes in precision medicine clinical trials often small
→ Borrowing from external sources to overcome associated challenges.

- Borrowing methods allow static or dynamic discounting of the amount of information transferred from external data (e.g., Viele et al 2014,…), often using Bayesian methods (e.g., Gravestock et al 2017, Neuenschwander et al 2010).

- Borrowing from external data by evaluating posterior distribution:
reject $H_0 \Leftrightarrow P(H_1 |$ current data; external data$) > 1 - \alpha$

- Frequentist operating characteristics (OCs) of hypothesis test when borrowing from external data are of interest: Type I error (TIE) rate, power.

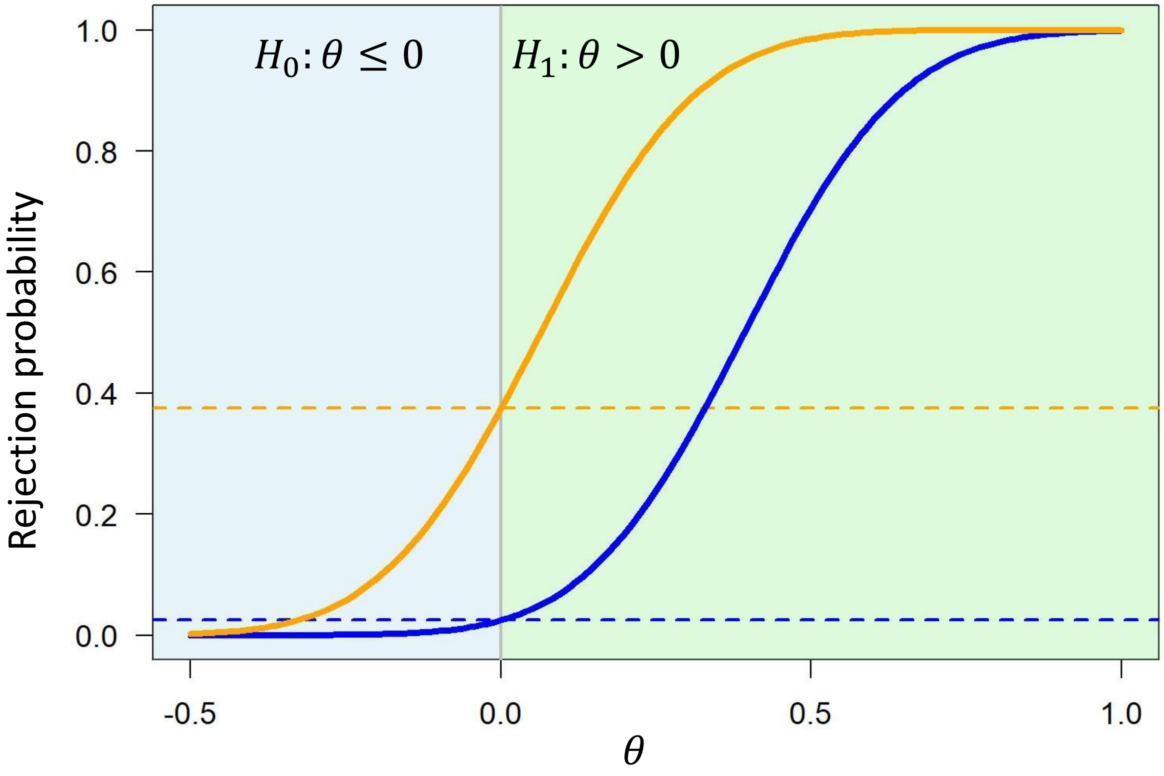
- Typically: External data is considered fixed → Evaluate "conditional" T1E and power.

- <u>Aim</u>: Discuss OCs for selected borrowing methods, and consider behavior for varying external data.

Example one-arm normal testing

- One-arm one-sided normal test situation: $H_0: \theta \leq 0$ vs. $H_1: \theta > 0$
- Current data $D \sim \mathcal{N}\left(\theta, {}^{1}/_{\sqrt{n}}\right)$, $n = 25$
- External data $D_E \sim \mathcal{N}\left(\theta_E, {}^{1}/_{\sqrt{n_E}}\right)$, $n_E = 20$
- Exemplary borrowing approach:
  Power Prior approach with fixed power parameter $\delta \in [0, 1]$:
  $\pi(\theta | d_E, \delta) \propto L(\theta; d_E)^{\delta} \pi(\theta)$ with $\delta = 0.5$
- Assume external data $\overline{d_E} = 1$
- $\alpha = 0.025$

Frequentist OCs w/o and w/ borrowing (fixed PP, $\delta = 0.5,\ \overline{d_E} = 1$)



Test w/ borrowing → $\text{Power}_B(\theta)$

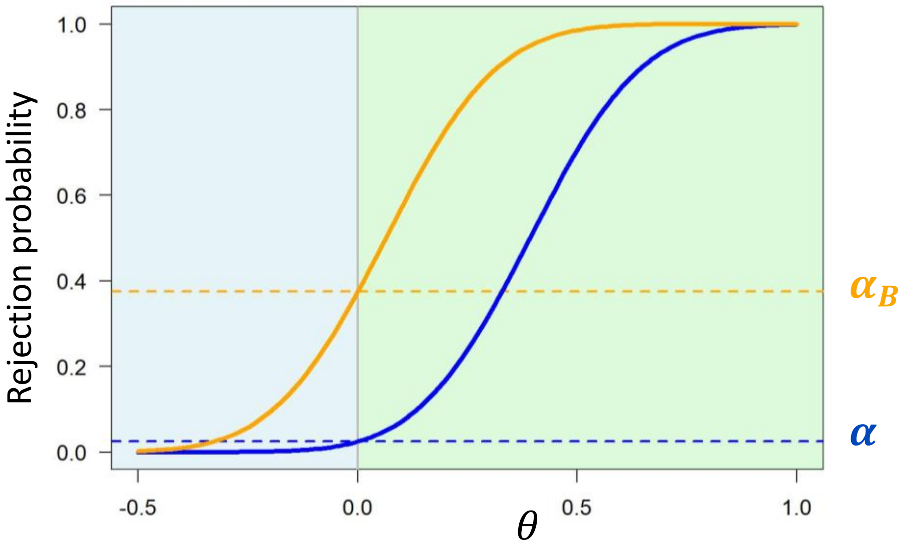
Test w/o borrowing → $\text{Power}^{(\alpha)}(\theta)$

α_B: T1E rate w/ borrowing

α: T1E rate w/o borrowing

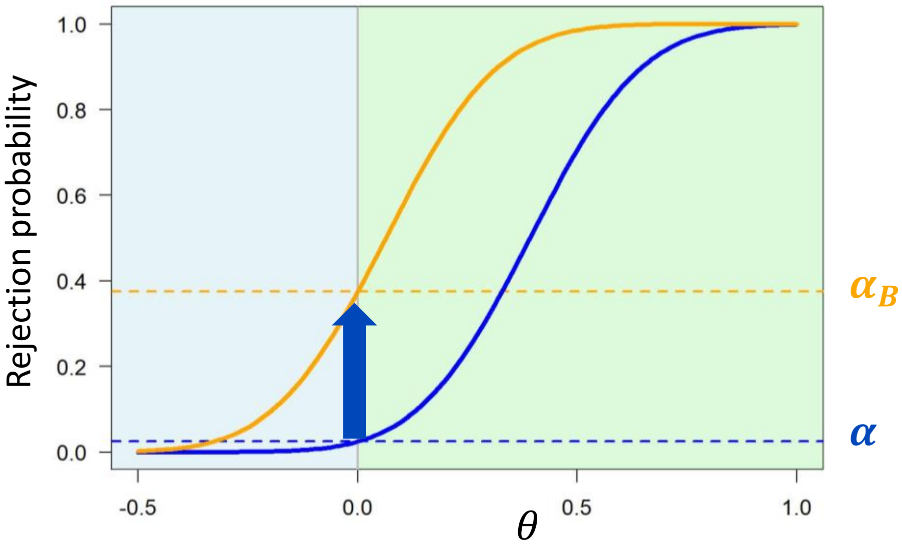
Problem

- Fair comparison of OC **w/** and **w/o** borrowing?



Problem

- Fair comparison of OC **w/** and **w/o** borrowing?



Solution

- Consider „**test calibrated to borrowing**“:

  test w/o borrowing, but T1E rate α_B instead of α

  → calibrated test and test w/ borrowing have same T1E rate ($= \alpha_B$)

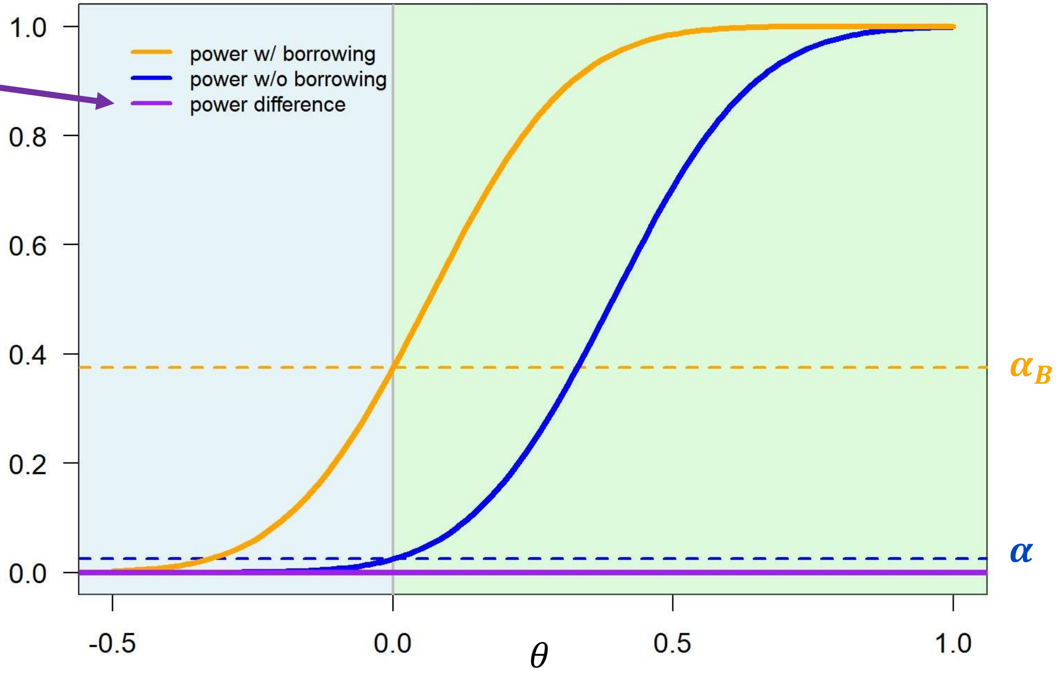
  → power for calibrated test: **Power$^{(\alpha_B)}(\theta)$**

- Evaluate power difference of test w/ borrowing and test calibrated to borrowing:

  Power$_B(\theta)$ $-$ **Power$^{(\alpha_B)}(\theta)$**

Comparing OCs w/ and w/o borrowing

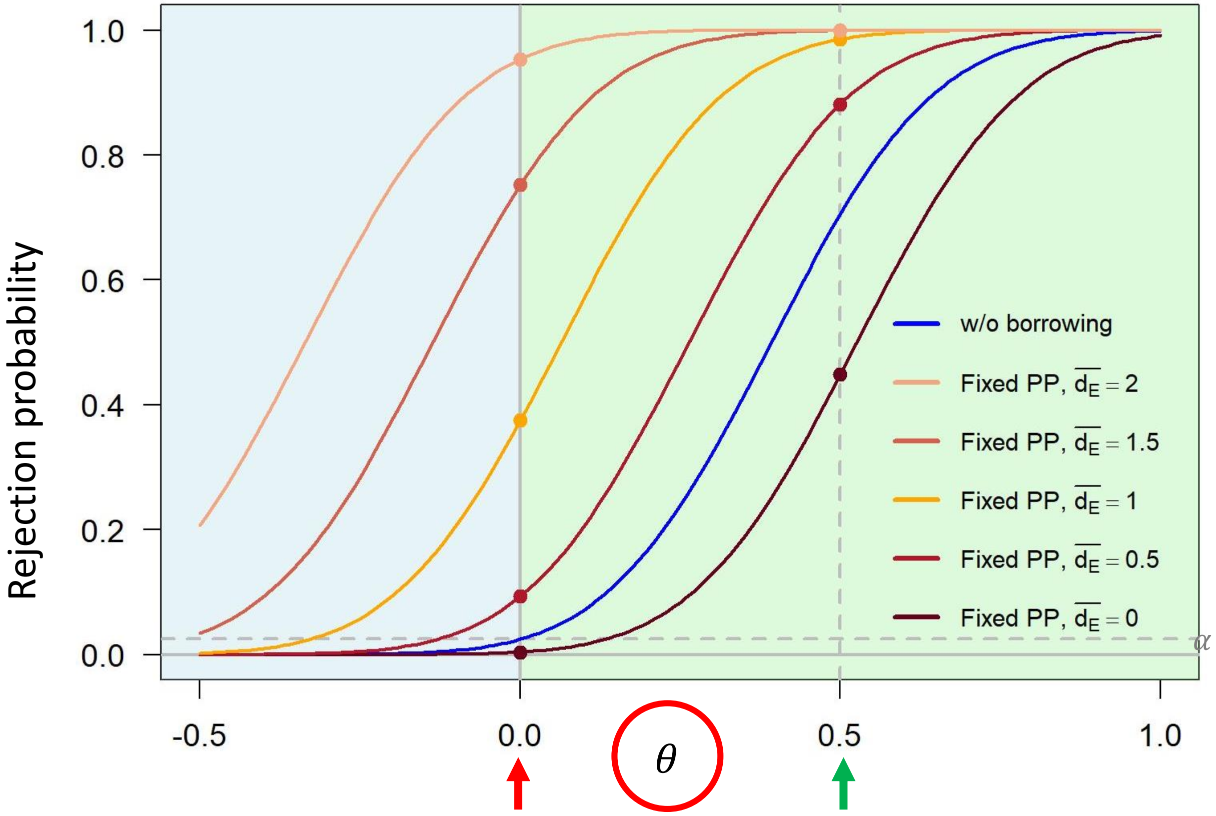
$\text{Power}_B(\theta)$ - $\text{Power}^{(\alpha_B)}(\theta)$



→ **Power difference = 0**: No power gain by borrowing.

Note: a uniformly most powerful test exists in this test situation → no test can have more power (Kopp-Schneider et al. 2020).

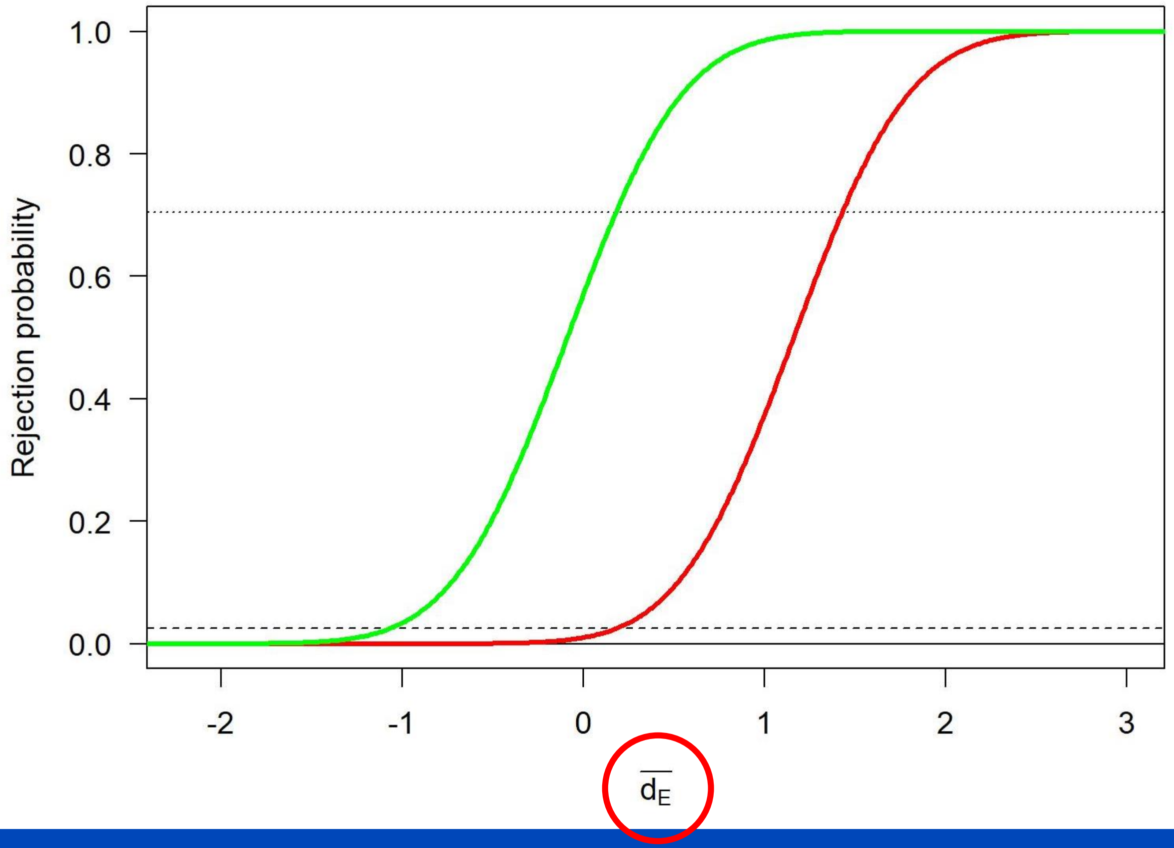
OCs for varying external data d_E (fixed PP, $\delta = 0.5$)



α_B: T1E rate

Power at $\theta = 0.5$

OCs for varying external data d_E (fixed PP, $\delta = 0.5$)



$\text{Power}_B(\overline{d_E})$ at $\theta = 0.5$

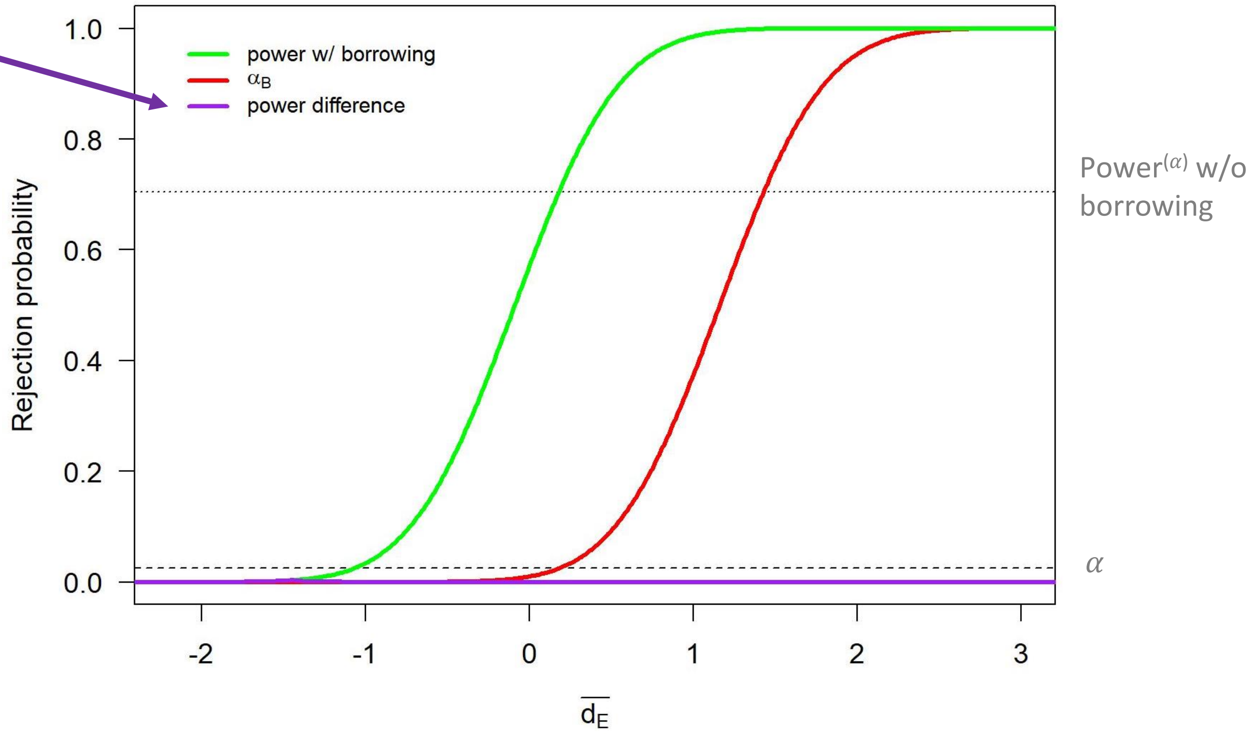
$\text{Power}^{(\alpha)}$ w/o borrowing

$\alpha_B(\overline{d_E})$: T1E rate w/ borrowing ($\theta = 0$)

α: T1E rate w/o borrowing

Comparing OCs w/ and w/o borrowing

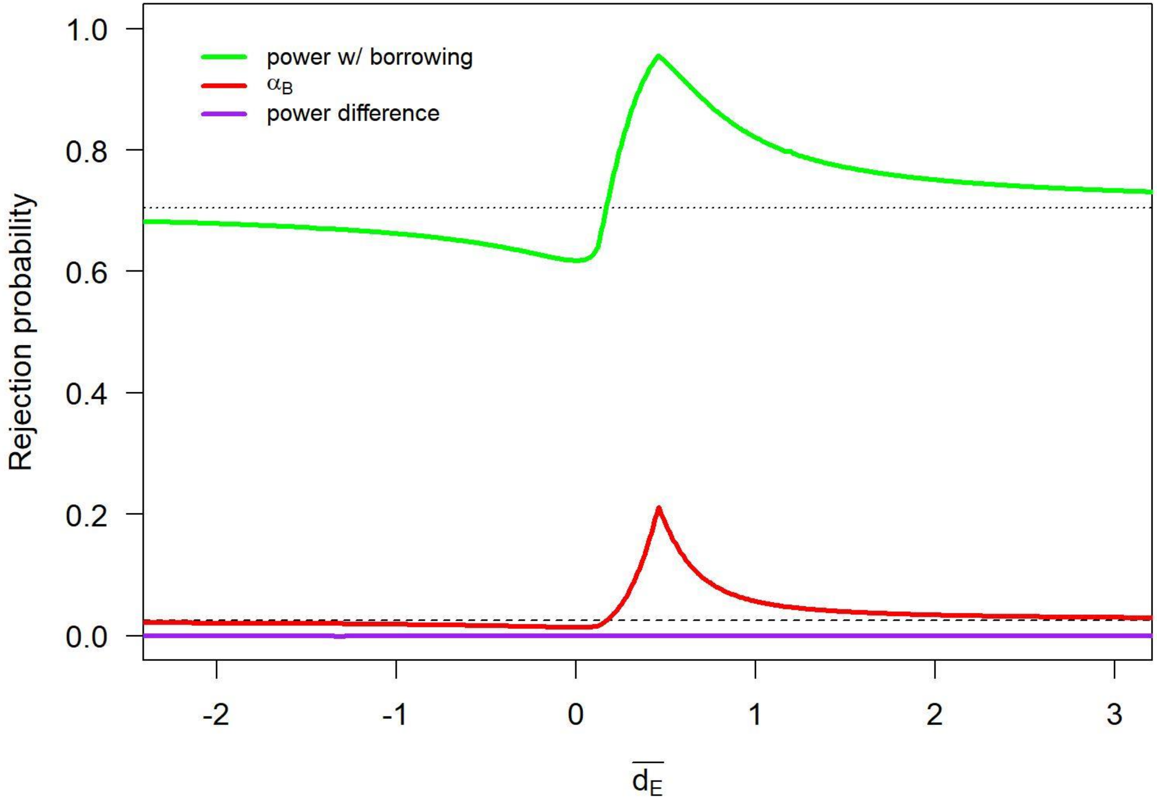
$\text{Power}_B(\theta = 0.5)$ - $\text{Power}^{(\alpha_B)}(\theta = 0.5)$



One-sided one-arm normal situation: Empirical Bayes Power Prior

- Power Prior approach $\pi(\theta|d_E, \delta) \propto L(\theta; d_E)^{\delta}\pi(\theta)$
- Now: Adapt $\delta = \delta(d; d_E)$ such that information is only borrowed for similar data
- Use Empirical Bayes approach for estimating $\hat{\delta}(d; d_E)$ (Gravestock, Held et al. 2017)

One-sided one-arm normal situation: Empirical Bayes Power Prior

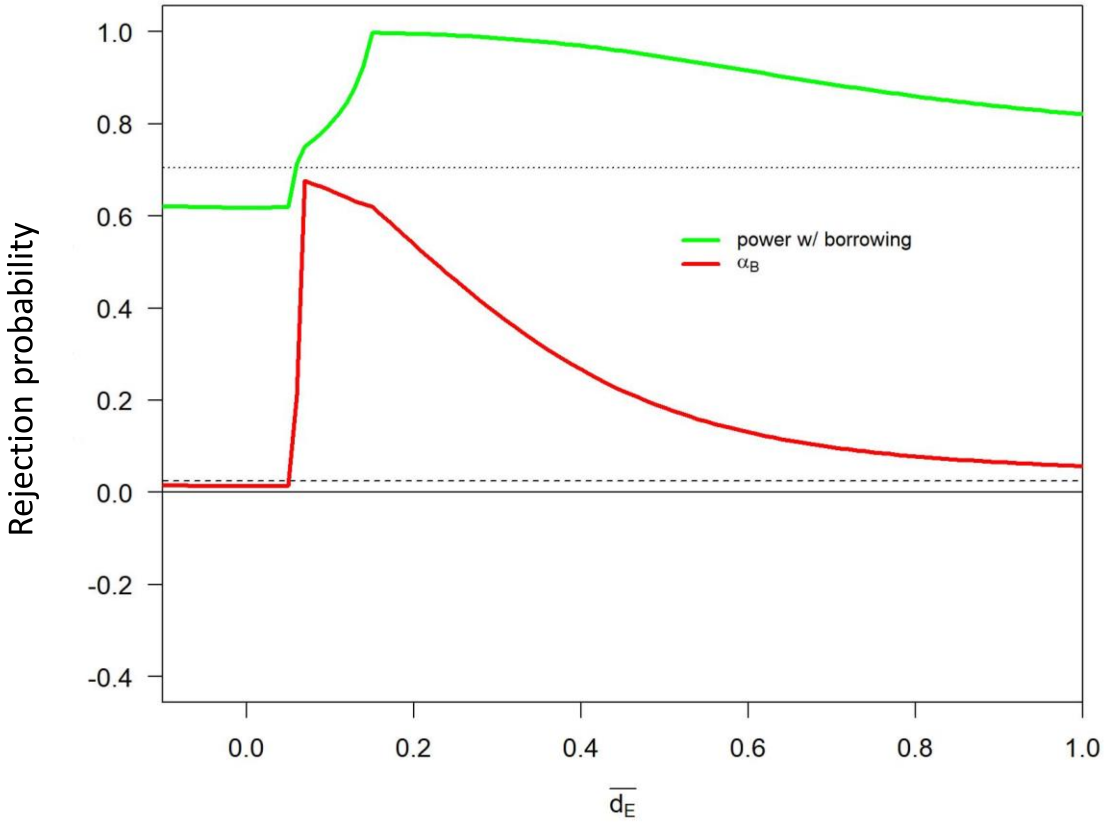


Power$^{(\alpha)}$ w/o borrowing

α

Note: Empirical Bayes Power Prior can be coerced to result in power loss

„Extreme borrowing“: external data sample size $n_E = 1000$

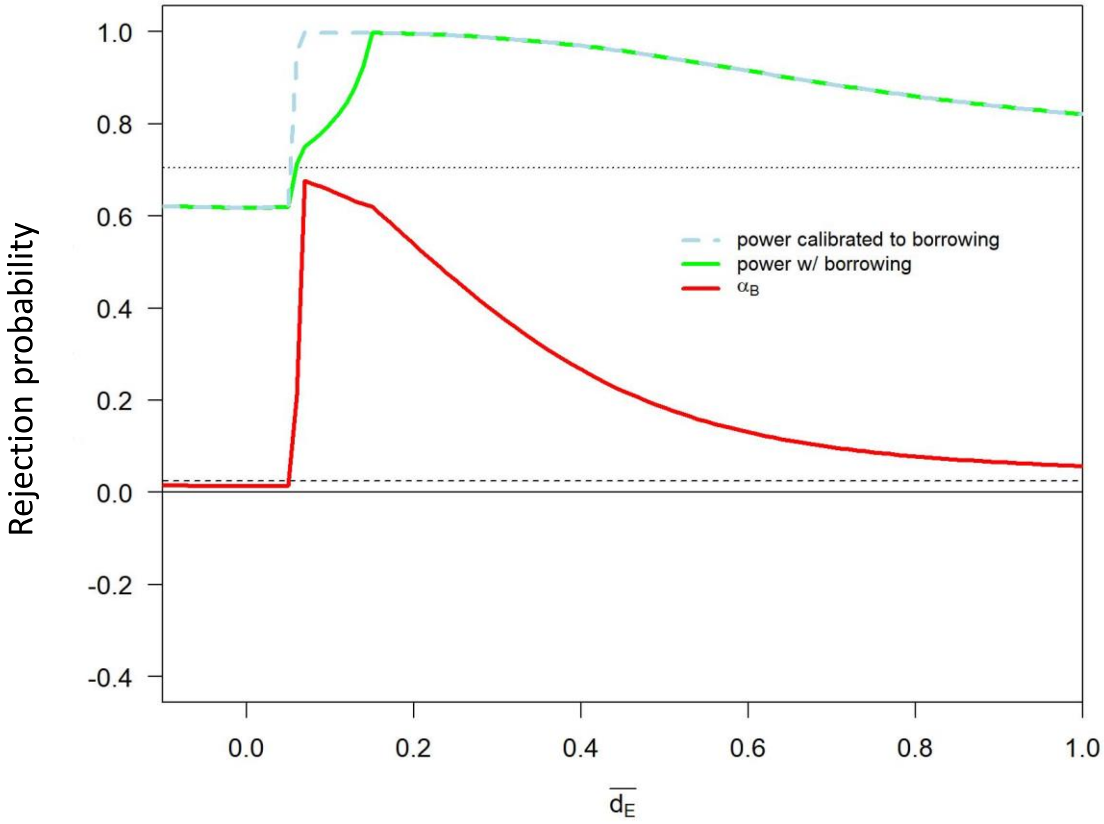


Power$^{(\alpha)}$ w/o borrowing

α

Note: Empirical Bayes Power Prior can be coerced to result in power loss

„Extreme borrowing“: external data sample size $n_E = 1000$

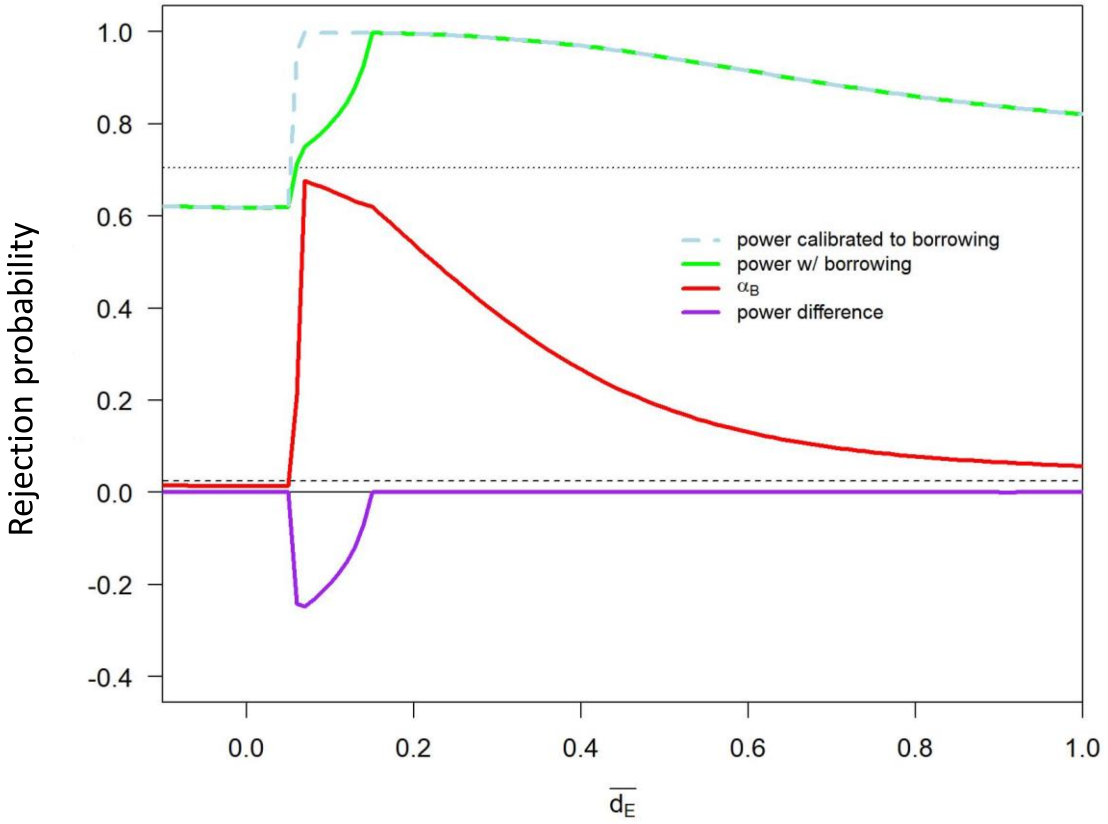


Power$^{(\alpha)}$ w/o borrowing

α

Note: Empirical Bayes Power Prior can be coerced to result in power loss

„Extreme borrowing“: external data sample size $n_E = 1000$



Power$^{(\alpha)}$ w/o borrowing

α

Note: Empirical Bayes Power Prior can be coerced to result in power loss

Reason?

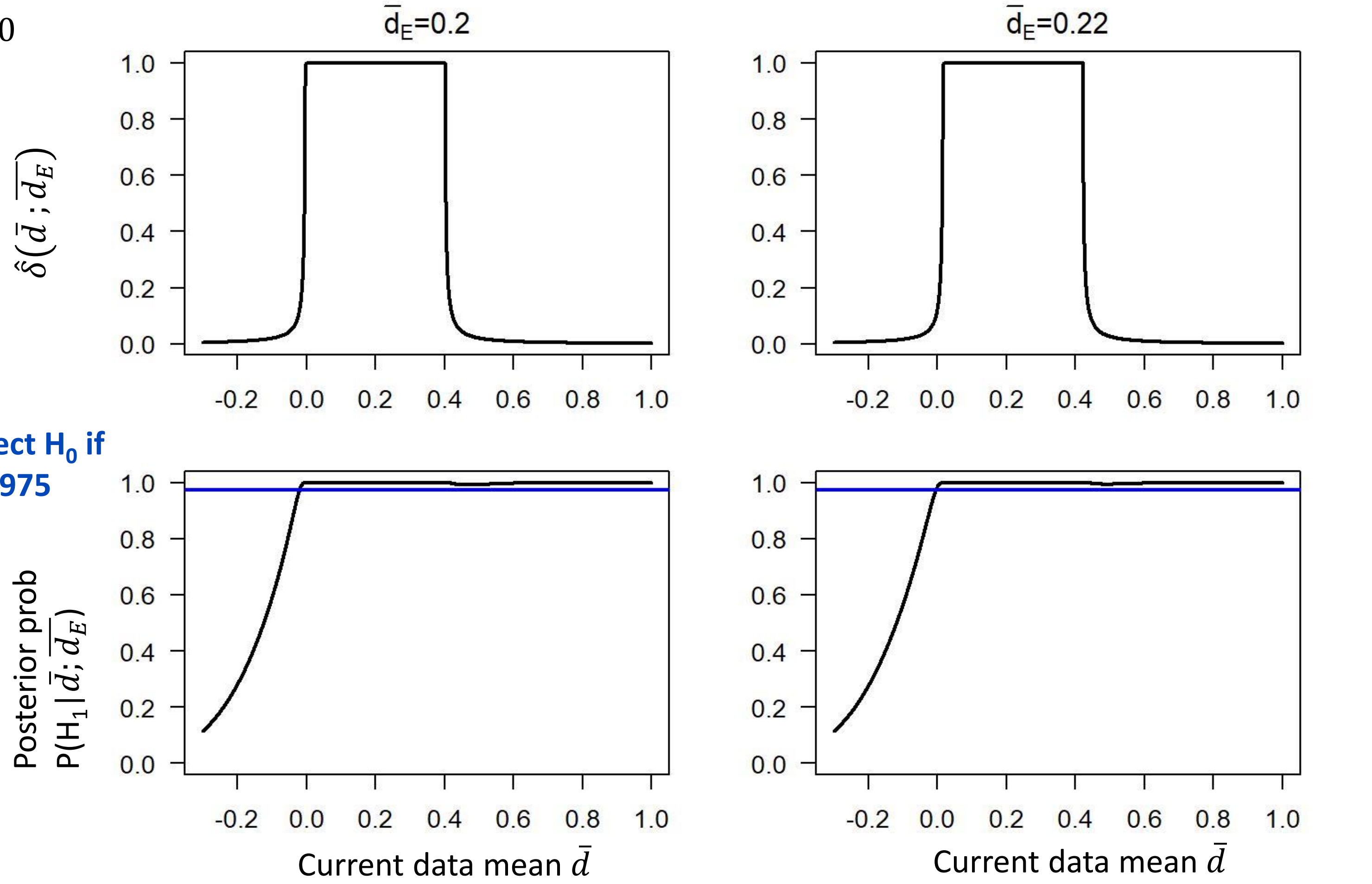
„Heavy“ external data ($n_E = 1000$) leads to posterior probability

P(H_1| current data d; external data d_E)

that is non-monotone in d.

Remember: reject H_0 ⇔ P(H_1| current data d; external data d_E) $> 1 - \alpha$

$n_E = 1000$



Reject H_0 if > 0.975

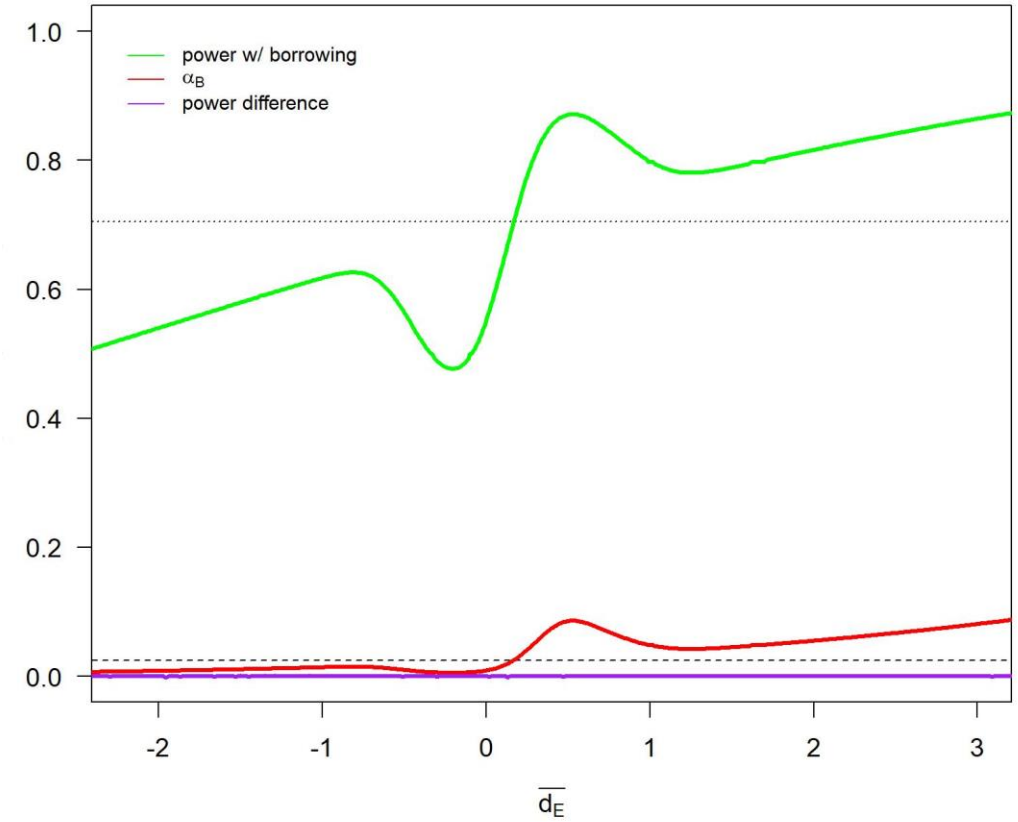
One-sided one-arm normal situation: Robust Mixture Prior

- Prior is mixture of informative and robust component („unit prior“) (Neuenschwander et al. 2010):

$$w \cdot \mathcal{N}\left(\overline{d_E}, {}^{1}/_{\sqrt{n_E}}\right) + (1 - w) \cdot \mathcal{N}\left(\overline{d_E}, 1\right)$$

- Here: use, e.g., $w = 0.5$

One-arm: Robust Mixture Prior



Power$^{(\alpha)}$ w/o borrowing

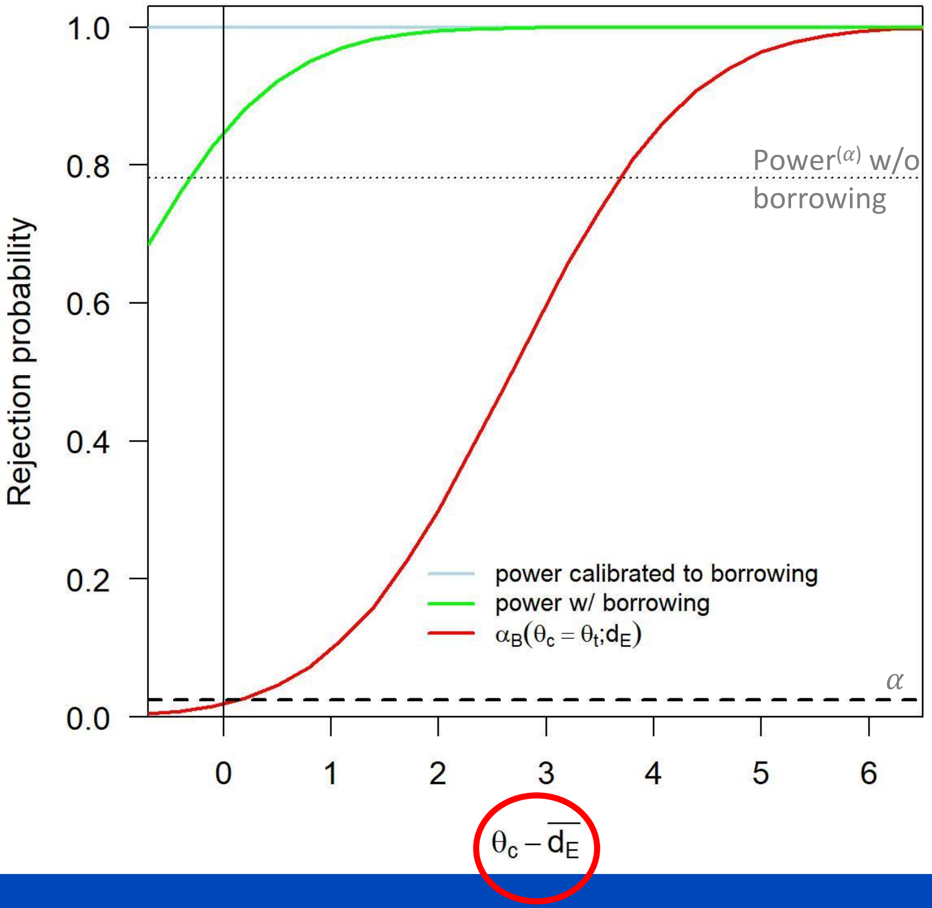
α

dkfz.

One-sided two-arm normal test with borrowing to control arm (aka hybrid control arm trial)

- Current treatment data $D_t \sim \mathcal{N}\left(\theta_t, {}^{1}/_{\sqrt{n}}\right), n = 15$
- Current control data $D_c \sim \mathcal{N}\left(\theta_c, {}^{1}/_{\sqrt{n}}\right), n = 15$
- External control data $D_{E,j} \sim \mathcal{N}\left(\theta_E, {}^{1}/_{\sqrt{n_E}}\right), n_E = 10$
- Two-arm one-sided test situation: $H_0: \theta_t \leq \theta_c$ vs. $H_1: \theta_t > \theta_c$
- Type I error rate obtained for $\theta_t - \theta_c = 0$
- Power evaluated at $\theta_t - \theta_c = 1$

Hybrid control arm trial: Fixed Power Prior

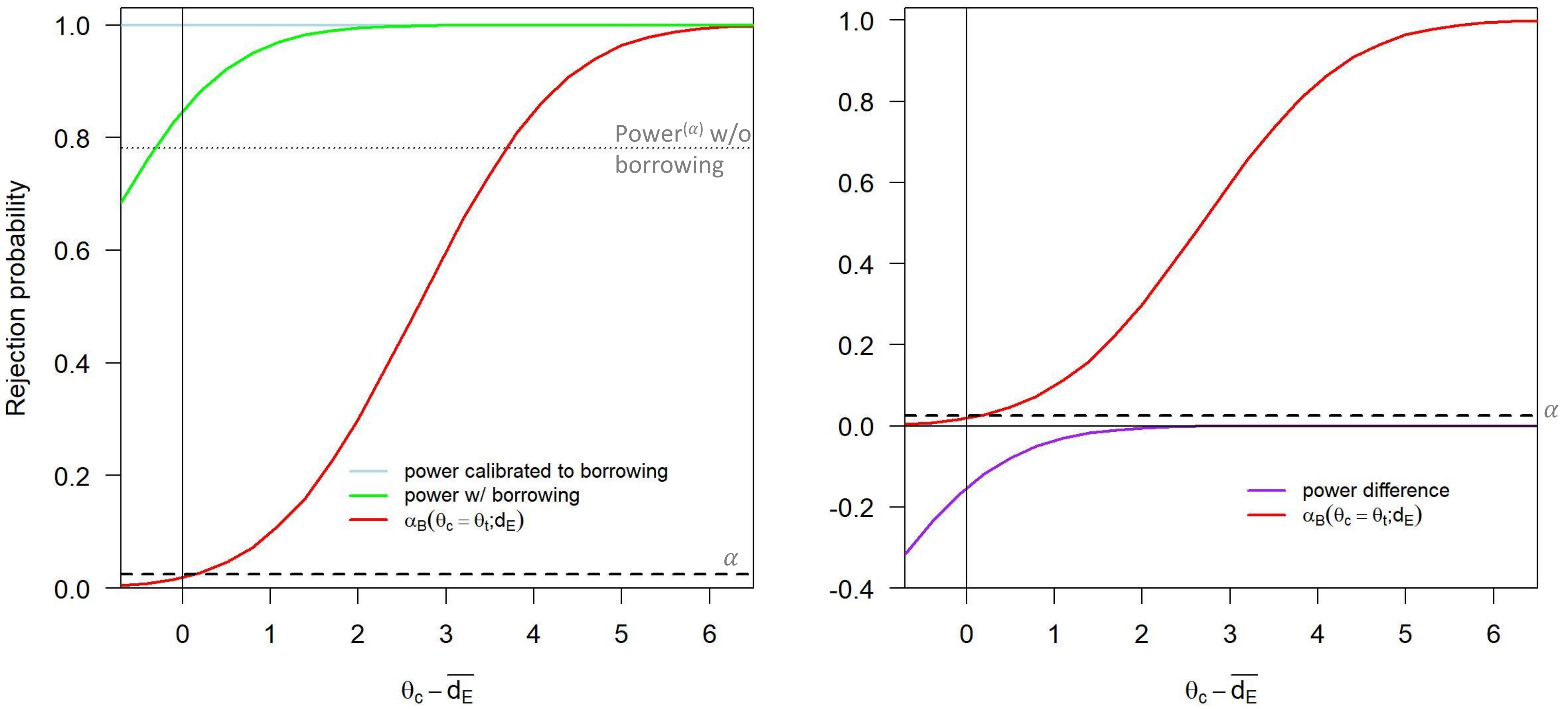


Power (at $\theta_t - \theta_c = 1$) of test calibrated to borrowing:

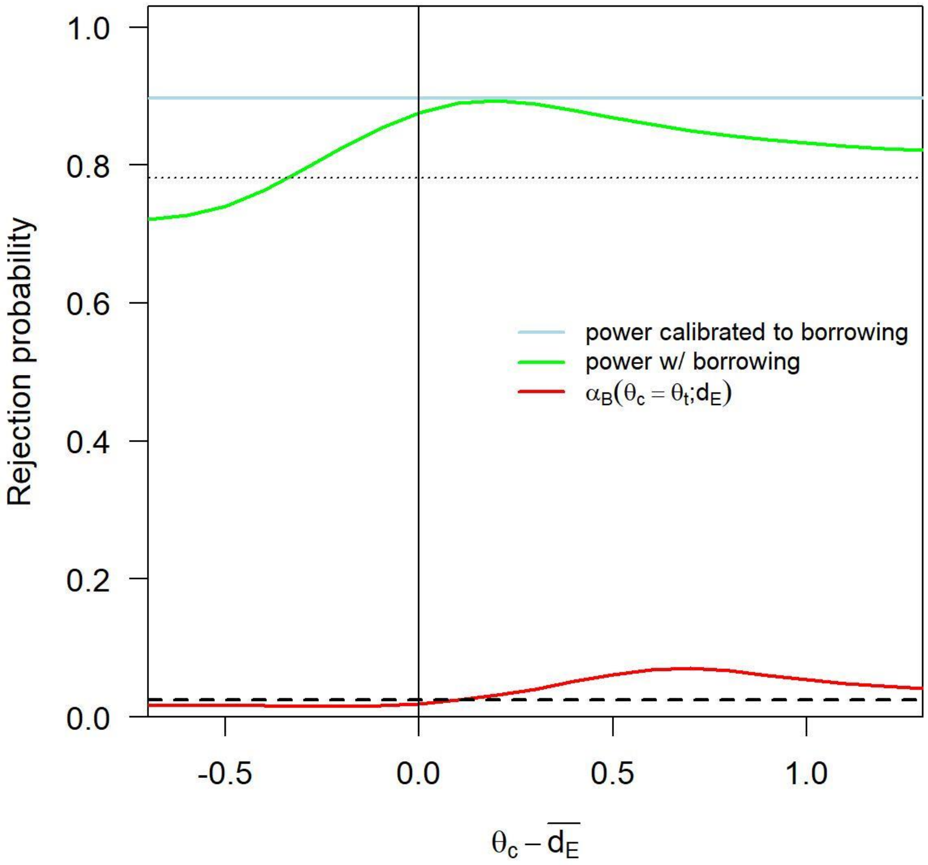
- $\alpha_B(d_E)$ varies with θ_c $(= \theta_t)$
- For increasing $\theta_c - \overline{d_E}$: $\alpha_B(d_E)$ goes to 1
- Since θ_c is unknown:
  need to calibrate to $\max_{\theta_c} \alpha_B(\theta_c = \theta_t; d_E)$

→ power calibrated to borrowing $= 1$

Hybrid control arm trial: Fixed Power Prior



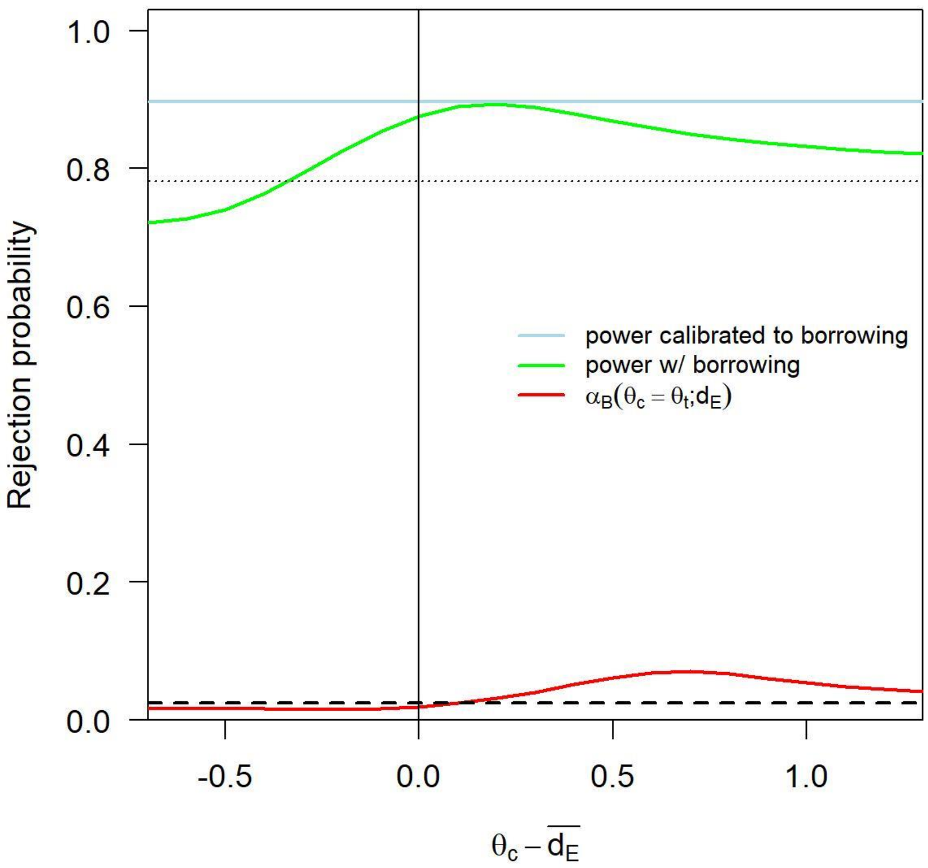
Hybrid control arm trial: Empirical Bayes Power Prior

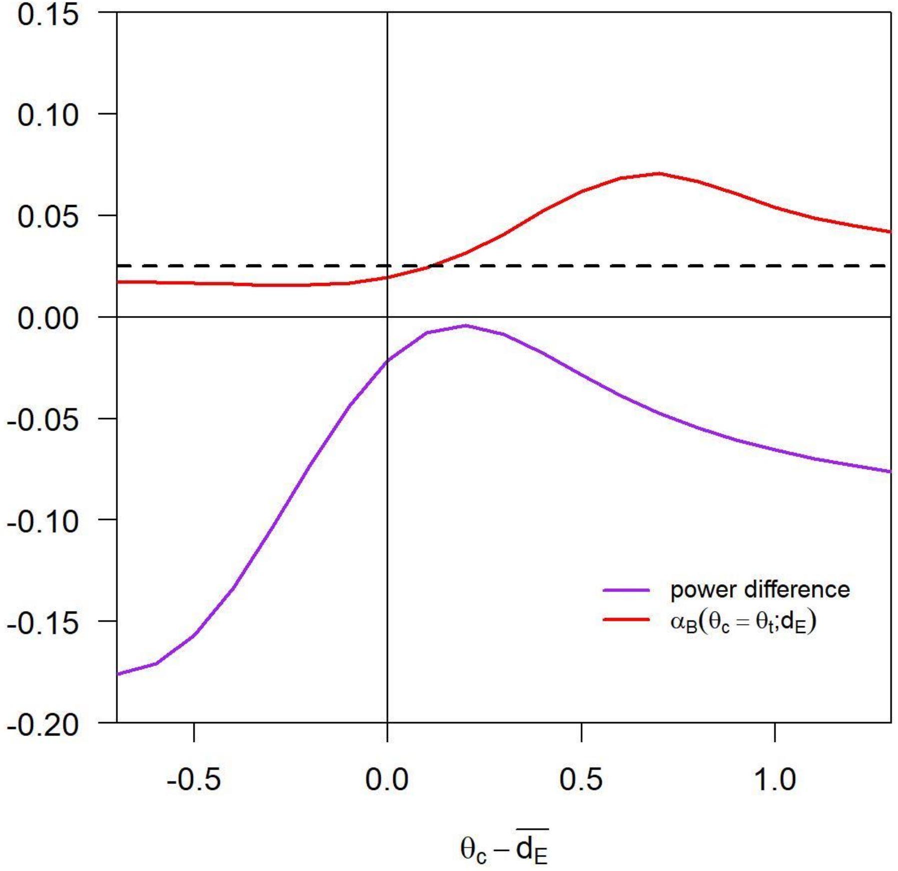


Power$^{(\alpha)}$ w/o borrowing

α

Hybrid control arm trial: Empirical Bayes Power Prior





Summary and Conclusions

- For fair comparison of OCs w/ and w/o borrowing:
  calibrate test w/o borrowing to α_B → evaluate **Power$_B(\theta)$** - **Power$^{(\alpha_B)}(\theta)$**.

- Focus here on OCs for one-sided tests for normal outcomes, external data considered fixed

- Borrowing of external information to one-arm trial:
  - fixed Power Prior: α_B goes to 1 when „$\overline{d_E}$ far in H_1"
  - EB Power Prior: T1E inflation is bounded, „extreme borrowing" can lead to power loss
  - Robust Mixture Prior: dynamic borrowing approach but α_B goes to 1 when „$\overline{d_E}$ far in H_1".

- Hybrid control trial:
  - fixed Power Prior: $\max_{\theta_c} \alpha_B(\theta_c = \theta_t; d_E) = 1$ for every $\overline{d_E}$.
  - EB Power Prior: T1E inflation less pronounced, but Power$_B$ - Power$^{(\alpha_B)}$ always < 0

- Not shown here: for random external data there seems to always be a (greater) power loss (proven for fixed PP borrowing in one-arm one-sided normal test).

- If prior information is reliable and consistent with new information, OCs can be improved
  → Incorporation of prior information can give a rationale for T1E inflation and power gain.

References

Gravestock, I. and Held, L. (2017) Adaptive power priors with empirical Bayes for clinical trials. *Pharmaceutical Statistics*, *16*, 349–360

Kopp-Schneider A, Calderazzo S, Wiesenfarth M. (2020) Power gains by using external information in clinical trials are typically not possible when requiring strict type I error control. *Biometrical Journal* 62(2):361-374.

Kopp-Schneider A, Wiesenfarth M, Held L, Calderazzo S, (2023) Simulating and reporting frequentist operating characteristics of clinical trials that borrow external information: Towards a fair comparison in case of one-arm and hybrid control two-arm trials. *Pharmaceutical Statistics*.

Neuenschwander B, Capkun-Niggli G, Branson M, and Spiegelhalter DJ (2010) Summarizing historical information on controls in clinical trials. *Clinical Trials* 7, 5 (2010).

Viele, K., Berry, S., Neuenschwander, B., Amzal, B., Chen, F., Enas, N.…Thompson, L. (2014). Use of historical control data for assessing treatment effects in clinical trials. *Pharmaceutical Statistics*, *13*(1), 41–54.

dkfz.