Bayesian methods in the development and assessment of new therapies
Workshop of the IBS-DR working group “Bayes Methods”
Göttingen, Dec 6-7 2018

Bayesian dynamic borrowing of external information: What can be gained in terms of frequentist power?

Annette Kopp-Schneider, Silvia Calderazzo and
Manuel Wiesenfarth

Division of Biostatistics, German Cancer Research Center (DKFZ)
Heidelberg, Germany

GERMAN
CANCER RESEARCH CENTER
IN THE HELMHOLTZ ASSOCIATION

Motivation

- Adult trial in subjects with previously treated advanced or recurrent solid tumors harboring DNA repair deficiencies:

  Endpoint: response to treatment (dichotomous)

  Two arms: Targeted therapy vs. Physician's choice

- DNA repair deficiencies also occur in children

  → investigate targeted therapy in a single-arm pediatric trial

<u>Question:</u> Should this single pediatric arm be designed as stand-alone arm or can power gain be expected when borrowing information from the adult targeted therapy arm?

Planning the pediatric arm with stand-alone evaluation: Bayesian approach (1)

- Number of responders in children, $R_{ped} \sim \text{Bin}(n_{ped}, p)$
- Test H_0: $p = p_0$ vs. H_1: $p > p_0$, $p_0 = 0.2$
- Type I error rate $\alpha = 0.05$
- $n_{ped} = 40$

- Bayesian approach: Use beta-binomial model

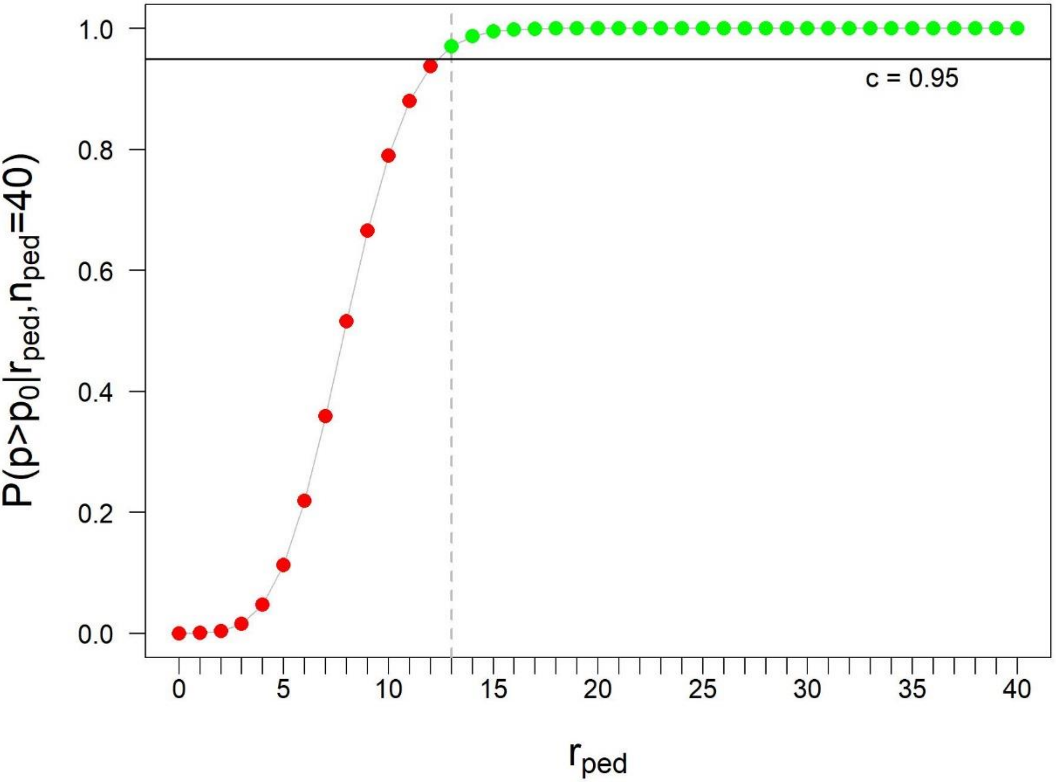
$$R_{ped} \mid p \sim \text{Bin}(n_{ped}, p), \pi(p) = \text{Beta}(0.5, 0.5)$$

- Evaluate efficacy based on Bayesian posterior probability:

$$P\left(p > p_0 | r_{ped}\right) \geq c, \text{ e.g., } c = 0.95.$$

Planning the pediatric arm with stand-alone evaluation: Bayesian approach (2)

Posterior probability $P(p > p_0 | r_{ped})$ as a function of r_{ped}



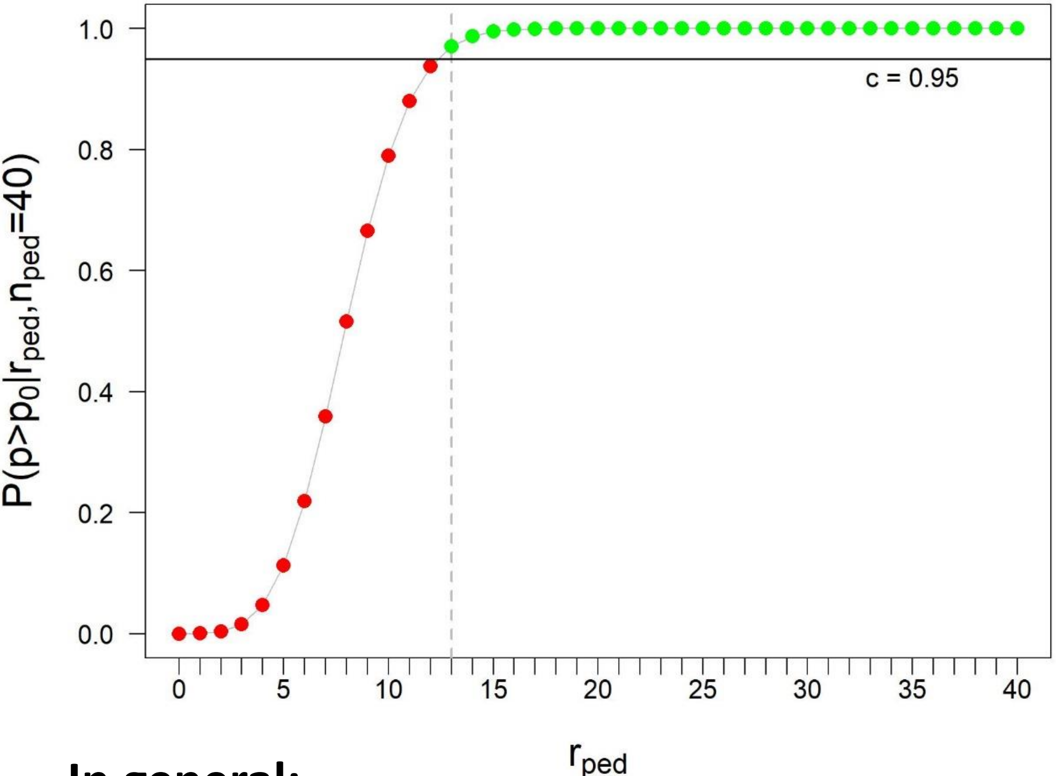
For $\boldsymbol{n_{ped} = 40}$:

$$P(p > p_0 | r_{ped}) \geq 0.95 \Leftrightarrow$$

$$r_{ped} \geq 13$$

Planning the pediatric arm with stand-alone evaluation: Bayesian approach (3)

Posterior probability $P(p > p_0 | r_{ped})$ as a function of r_{ped}



For $\boldsymbol{n_{ped} = 40}$:

$$P(p > p_0 | r_{ped}) \geq 0.95 \Leftrightarrow$$

$$r_{ped} \geq 13$$

In general:

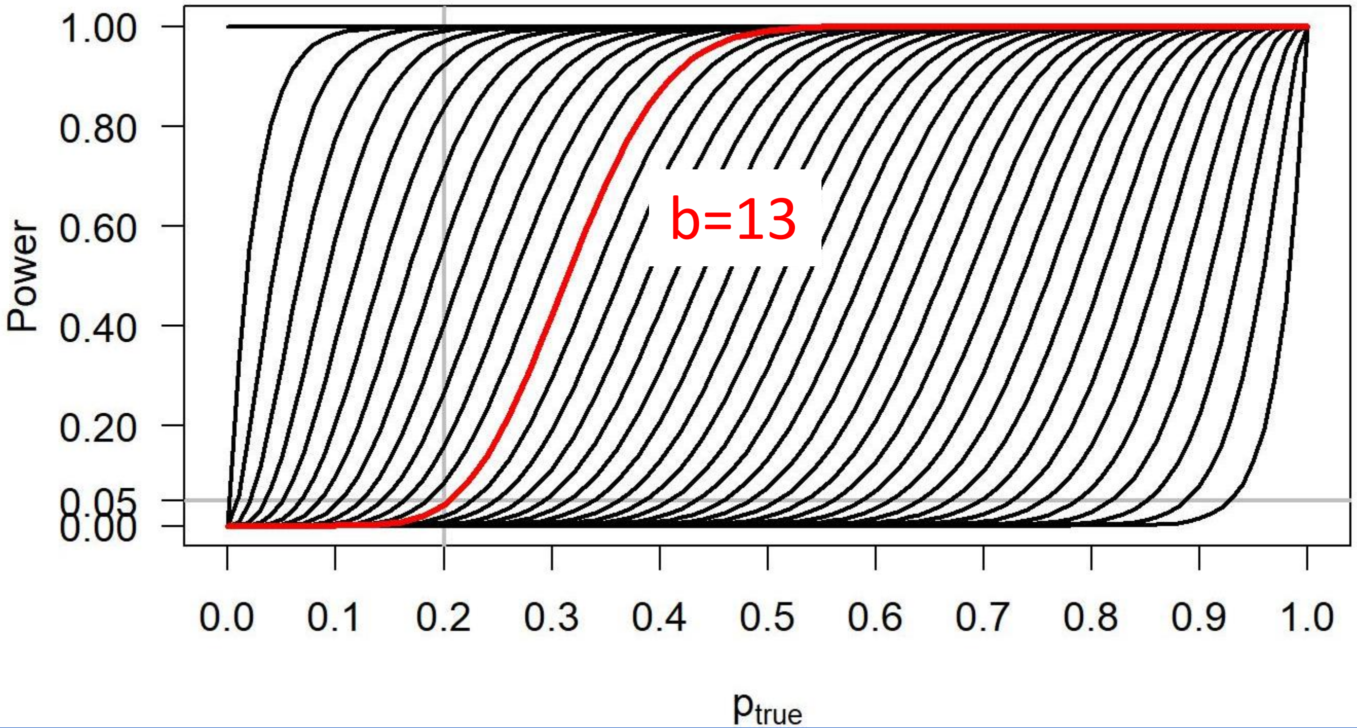
For every $c \in [0, P(p > p_0 | r_{ped} = n_{ped})]$ there exists a unique $b \in \{0,1, \dots, n_{ped}\}$ with $P(p > p_0 | r_{ped}) \geq c \Leftrightarrow r_{ped} \geq b$ (Kopp-Schneider et al., 2018)

Planning the pediatric arm with stand-alone evaluation: Frequentist approach

- Test H_0: $p = p_0$ vs. H_1: $p > p_0$
- Type I error rate α, e.g., $\alpha = 0.05$
- Uniformly most powerful (UMP) level α test is given by:

  reject $H_0 \Leftrightarrow r_{ped} \geq b_{\text{UMP}}(\alpha)$
- Here: $b_{\text{UMP}}(0.05) = 13$

Power: $P\left(R_{ped} \geq \ b | p_{true}\right)$



Planning the pediatric arm with stand-alone evaluation: Power function (1)

$$\text{Power} = f(p_{true})$$

$$= P(R_{ped} \geq b | p_{true})$$

$$= \sum_{r_{ped}=0}^{n} P\left(R_{ped} = r_{ped} | p_{true}\right) 1_{\{r_{ped} \geq b\}}$$

Planning the pediatric arm with stand-alone evaluation: Power function (2)

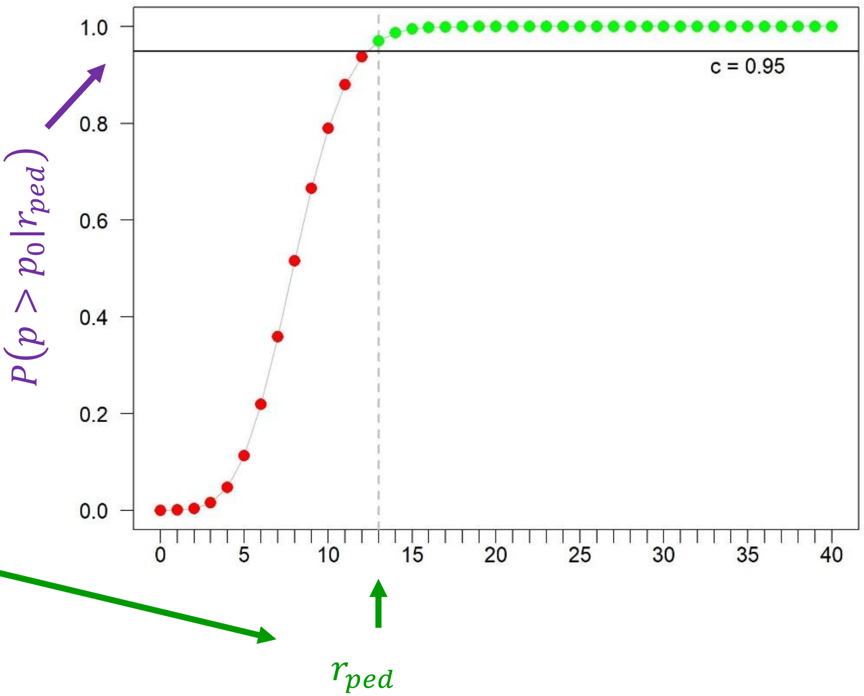
Power $= f(p_{true})$

$= P\left(R_{ped} \geq b | p_{true}\right)$

$$= \sum_{r_{ped}=0}^{n} P\left(R_{ped} = r_{ped} | p_{true}\right) 1_{\{r_{ped} \geq b\}}$$

$$= \sum_{r_{ped}=0}^{n} P\left(R_{ped} = r_{ped} | p_{true}\right) 1_{\{P(p > p_0 | r_{ped}) \geq c\}}$$

(c selected appropriately)

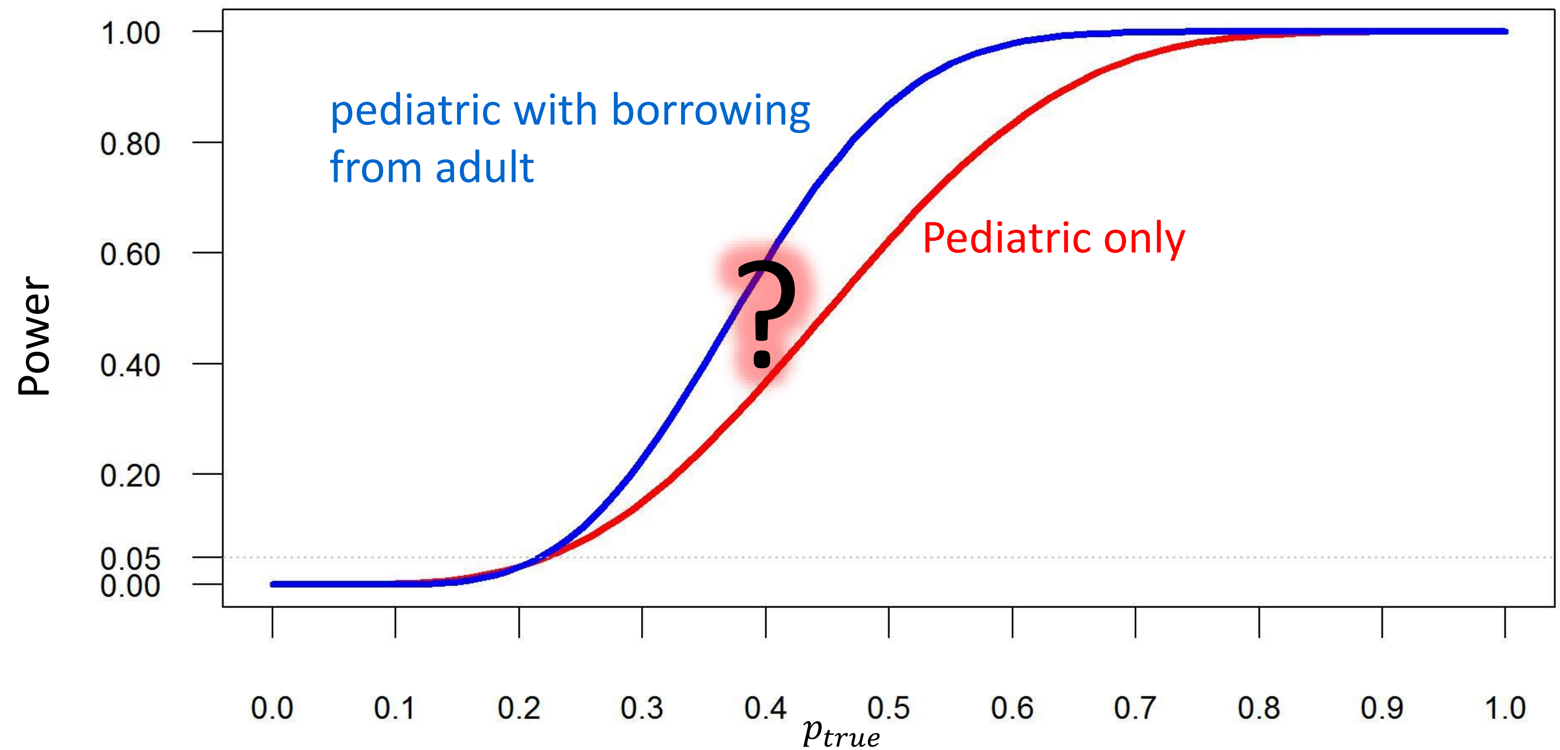


Borrowing from adult information for the pediatric arm

Use information from adults to inform the prior for the pediatric trial.

Hope

If treatment is successful in adults, then power is increased for pediatric trial:



Adaptive power parameter (1)

Power prior approach with power parameter $\delta \in [0, 1]$:

$$\pi(p|r_{adu}, \delta) \ \propto L(p; r_{adu})^{\delta} \pi(p)$$

Adapt $\delta = \delta(r_{ped}, r_{adu})$ such that information is only borrowed for similar adult and pediatric data:

$\rightarrow \delta(r_{ped}, r_{adu})$ large when adult and children data are similar

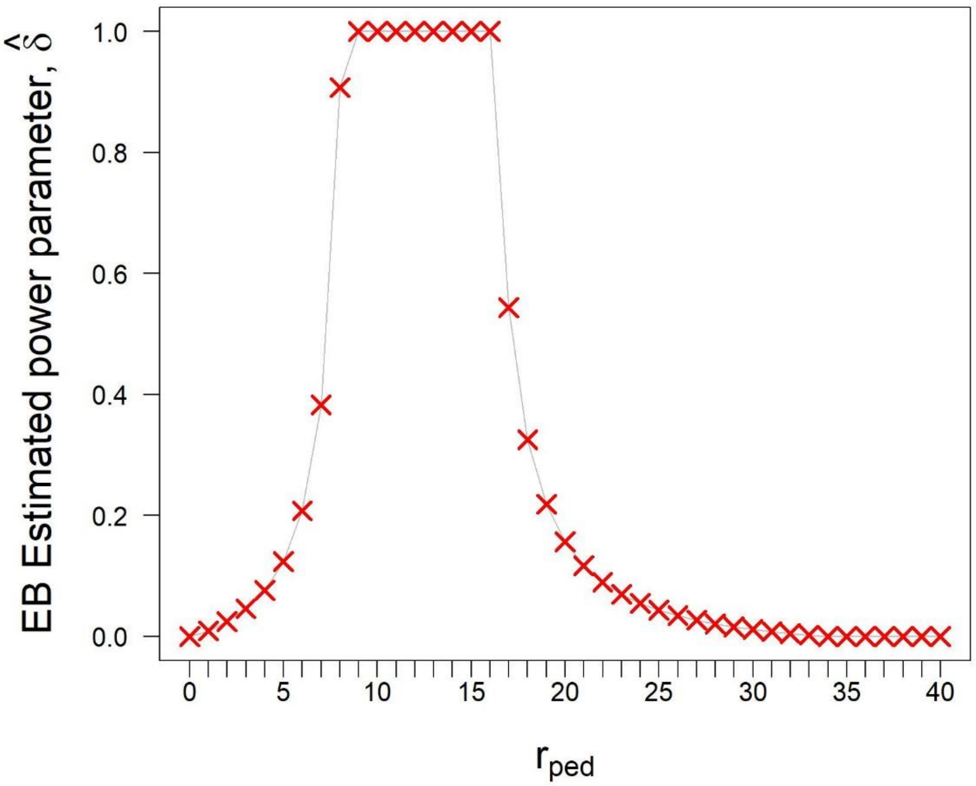
$\rightarrow \delta(r_{ped}, r_{adu})$ small in case of prior-data conflict.

Adaptive power parameter (2)

Result from adult trial: e.g., $r_{adu} = 12$ among $n_{adu} = 40$ ($\hat{p}_{adu} = 0.3$)

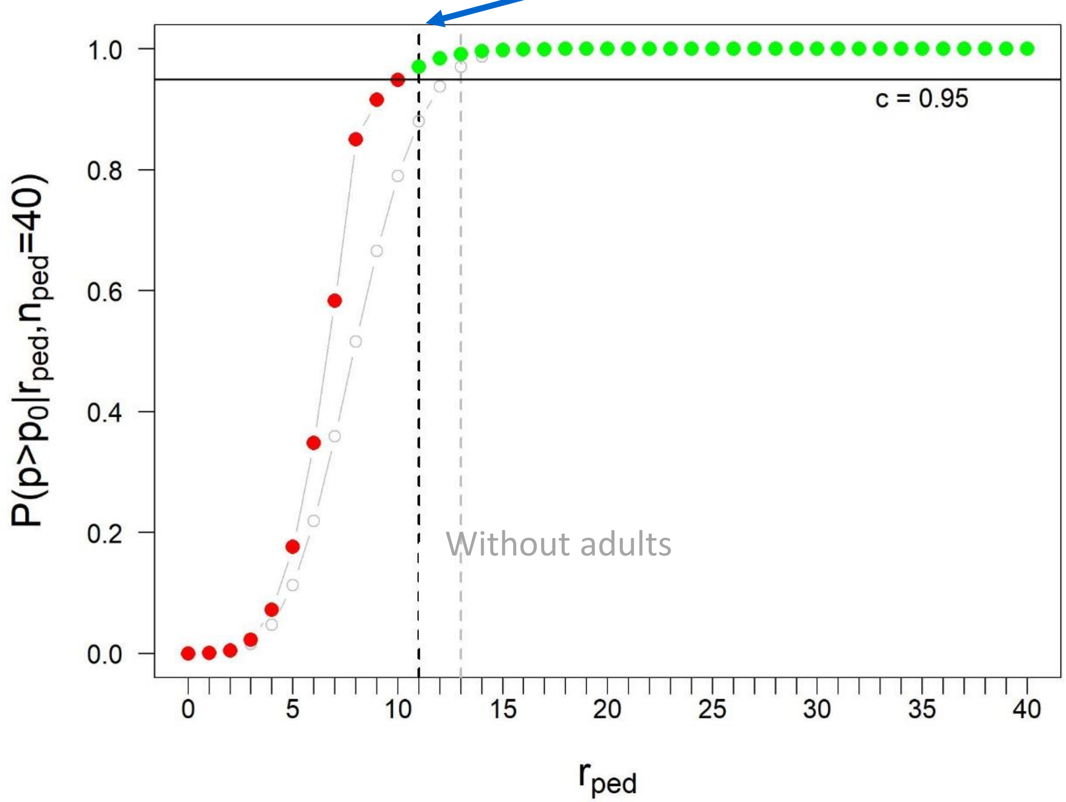
Use an Empirical Bayes approach where $\hat{\delta}(r_{ped}; r_{adu} = 12)$ maximizes the marginal likelihood of δ (Gravestock, Held et al. 2017):

$\hat{\delta}(r_{ped}; r_{adu} = 12)$:



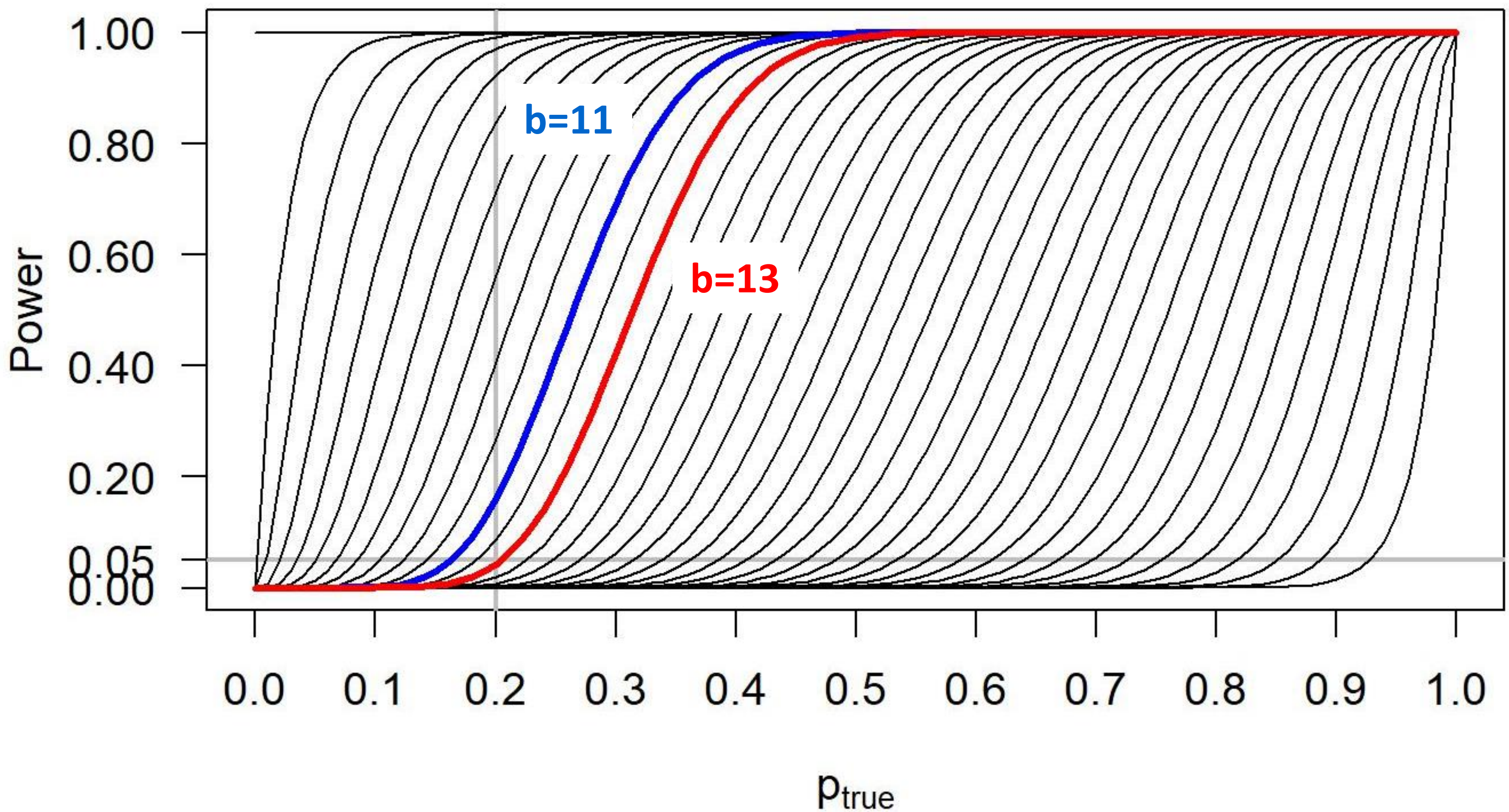
Adaptive power parameter (3)

$P\left(p > p_0 | r_{ped}, r_{adu}, \hat{\delta}(r_{ped}; r_{adu})\right) > c = 0.95$ corresponds to $r_{ped} \geq b = 11$



Adaptive power parameter (4)

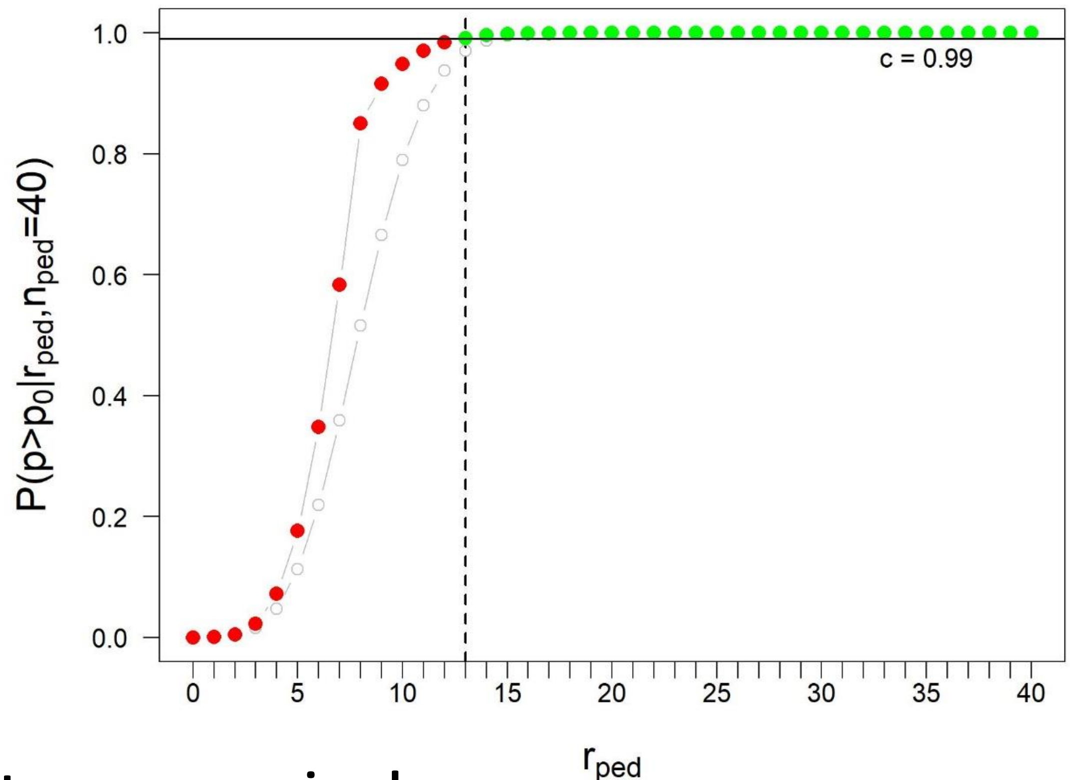
$P\left(p > p_0 | r_{ped}, r_{adu}, \hat{\delta}(r_{ped}; r_{adu})\right) > c = 0.95$ corresponds to $r_{ped} \geq b = 11$



→ **power gain but type I error inflation**

Adaptive power parameter (5)

- For this situation: $P\left(p > p_0 | r_{ped}, r_{adu}, \hat{\delta}(r_{ped}, r_{adu})\right)$ is monotonically increasing in r_{ped}
- $P(p > p_0 | r_{ped}, r_{adu}, \hat{\delta}) > c' = 0.99$ corresponds to $x_{ped} \geq b = 13$



→ **type I error controlled but no power gained**

Robust mixture prior (1)

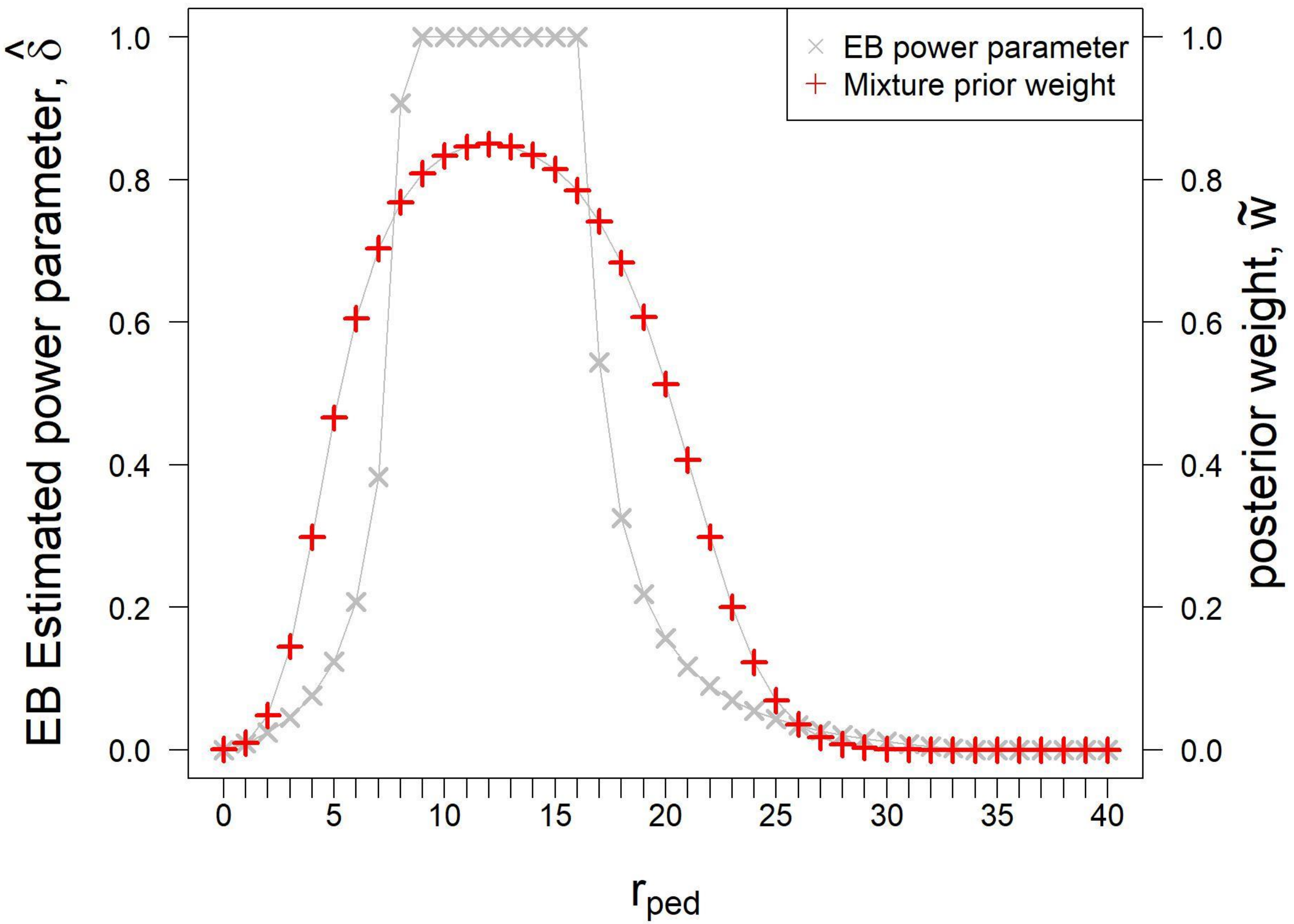
- Another way of discounting prior information is given by the use of robust mixture prior as convex combination of an uninformative prior and a prior that incorporates external information (e.g., Schmidli et al. (2014))

$$\pi(p) \;= w\ \text{Beta}(0.5{+}r_{adu},\, 0.5{+}n_{adu}{-}r_{adu}) + (1-w)\ \text{Beta}(0.5,\, 0.5)$$

- Here: $w = 0.5$

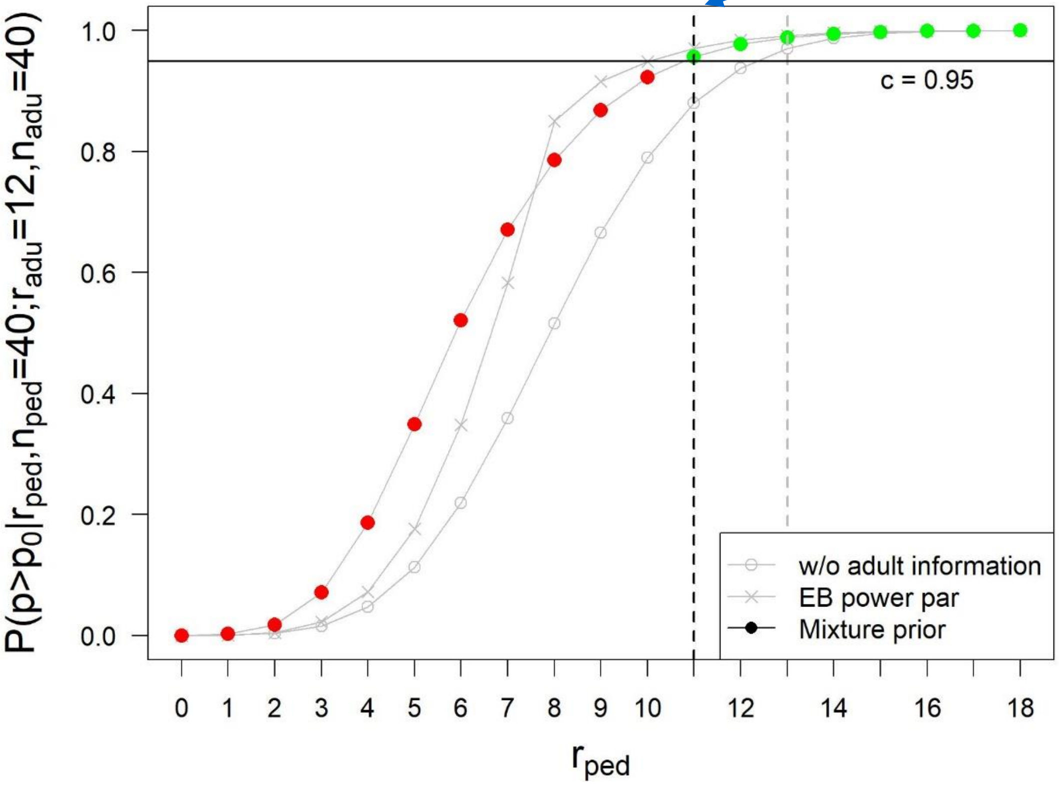
- Posterior is convex combination of Beta distributions with weight $\widetilde{w}$

Robust mixture prior (2)



Robust mixture prior (3)

$P(p > p_0 | r_{ped}, r_{adu}, \tilde{w}) > c = 0.95$ corresponds to $r_{ped} \geq b = 11$



→ **type I error inflation**

→ select $c' = 0.98 \rightarrow b = 13 \rightarrow$ **type I error controlled but no power gained.**

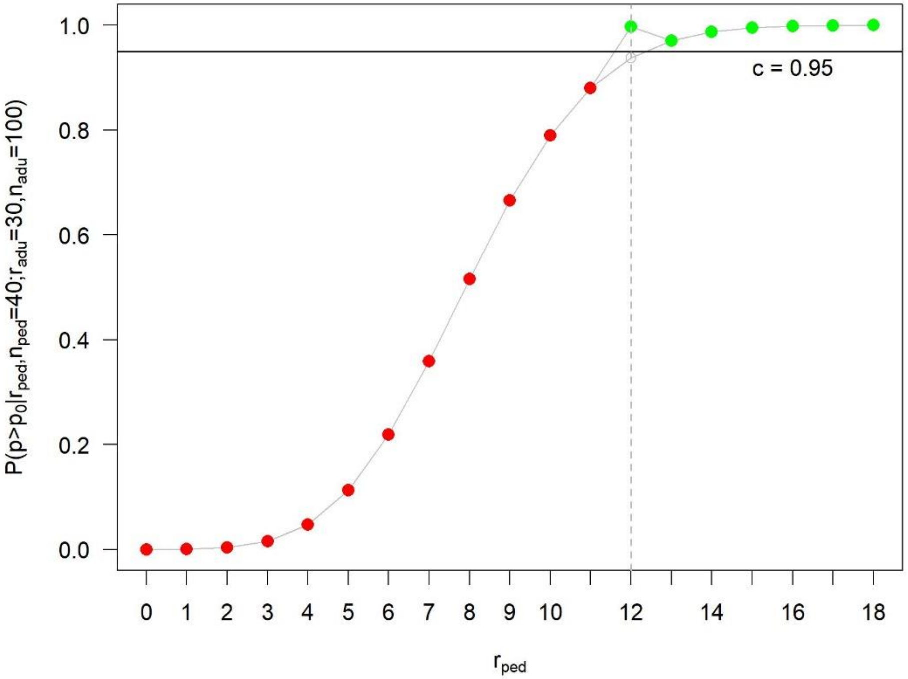
"Extreme borrowing" (1)

- Artificial method for illustration of not monotonically increasing $P(p > p_0|r_{ped}, r_{adu})$: borrow adult information $\Leftrightarrow \hat{p}_{adu} = \hat{p}_{ped}$
- Assume $n_{adu} = 100$, $r_{adu} = 30 \Rightarrow \hat{p}_{adu} = 0.3$
- Here: borrow all adult information if $\hat{p}_{ped} = 0.3 \Rightarrow r_{ped} = 12$

"Extreme borrowing" (2)

- Artificial method for illustration of not monotonically increasing $P(p > p_0 | r_{ped}, r_{adu})$: borrow adult information $\Leftrightarrow \hat{p}_{adu} = \hat{p}_{ped}$
- Assume $n_{adu} = 100, r_{adu} = 30 \Rightarrow \hat{p}_{adu} = 0.3$
- Here: borrow all adult information if $\hat{p}_{ped} = 0.3 \Rightarrow r_{ped}$ = 12

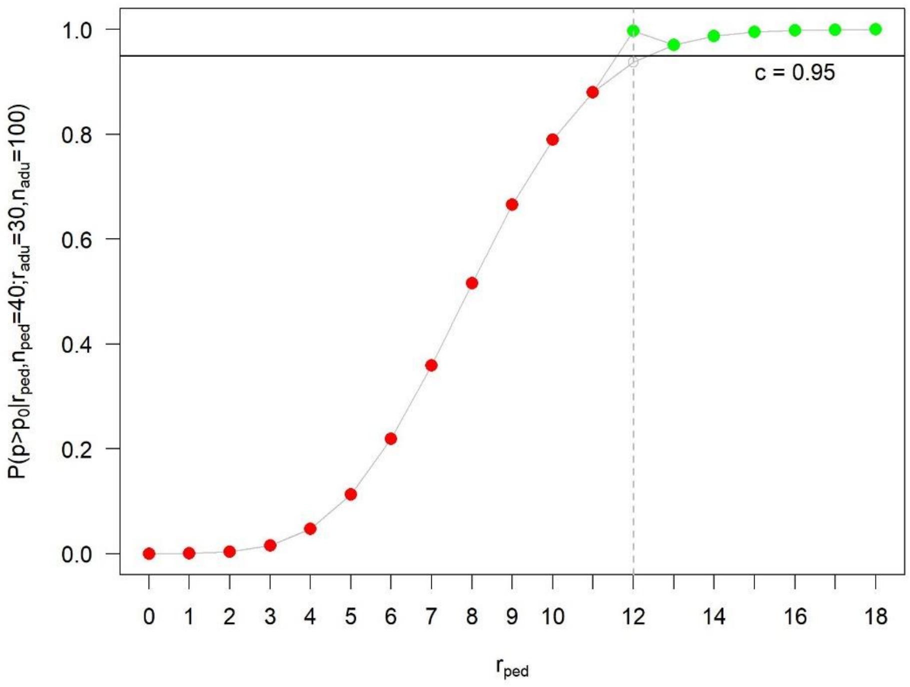
For $c = 0.95 \Rightarrow b = 12$
$\Rightarrow$ type I error rate $= 0.088$



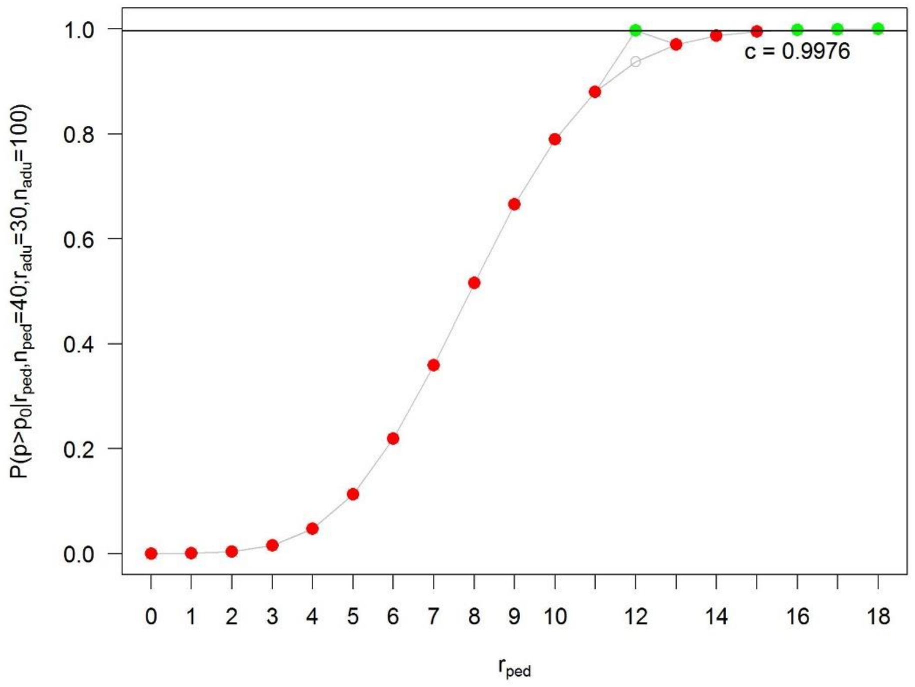
"Extreme borrowing" (3)

- Artificial method for illustration of not monotonically increasing $P(p > p_0 | r_{ped}, r_{adu})$: borrow adult information $\Leftrightarrow \hat{p}_{adu} = \hat{p}_{ped}$
- Assume $n_{adu} = 100$, $r_{adu} = 30 \Rightarrow \hat{p}_{adu} = 0.3$
- Here: borrow all adult information if $\hat{p}_{ped} = 0.3 \Rightarrow r_{ped} = 12$

For $c = 0.95 \Rightarrow b = 12$
$\Rightarrow$ type I error rate $= 0.088$



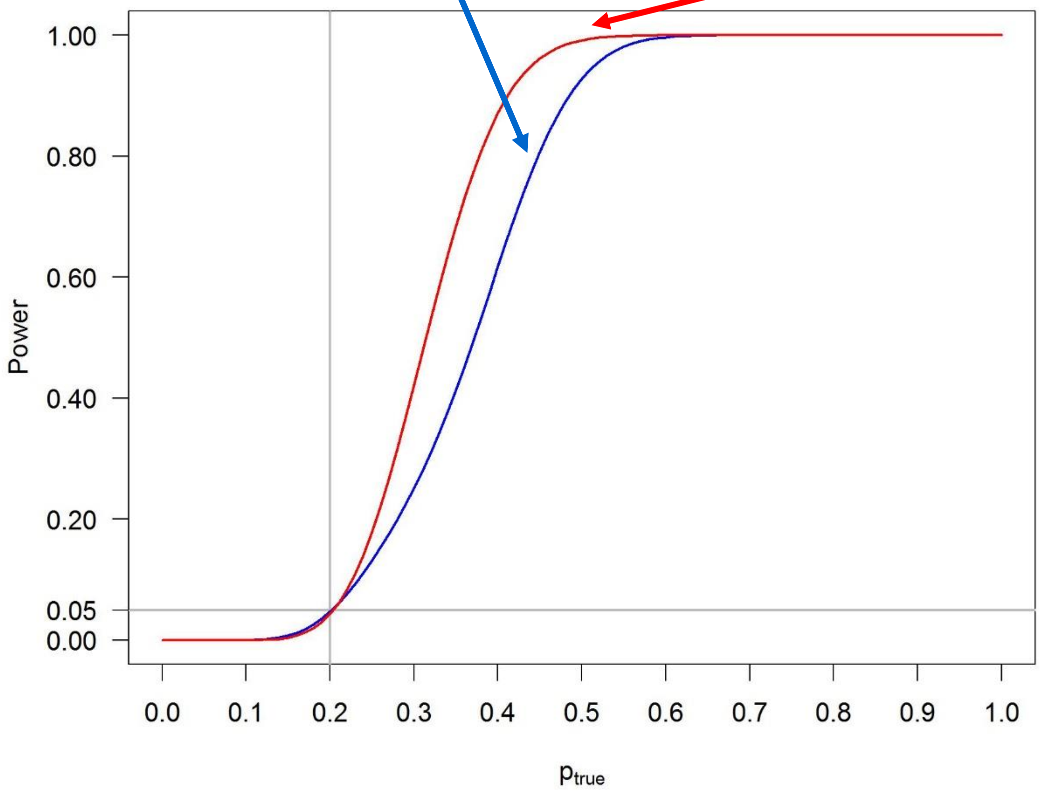
For $c = 0.9976$
$\Rightarrow$ reject H_0 if $b = 12$ or $b \geq 16$
$\Rightarrow$ type I error rate $= 0.047$



“Extreme borrowing” (4)

Reject H_0 if $b \in \{12\} \cup \{16, 17, \dots, 40\}$

Compare to: Reject H_0 if $b \in \{13, 17, \dots, 40\}$



→ **type I error controlled but power decreased**

Borrowing from adult information in general (1)

- If $P\left(p > p_0 | r_{ped}, r_{adu}\right)$ is monotonically increasing in r_{ped}, then there exists c' with

$$P\left(p > p_0 | r_{ped}, r_{adu}\right) \geq c' \Leftrightarrow r_{ped} \geq b_{\mathrm{UMP}}(\alpha) \; (*)$$

and $b_{\mathrm{UMP}}(\alpha)$ is the level α UMP test boundary.

Borrowing from adult information in general (2)

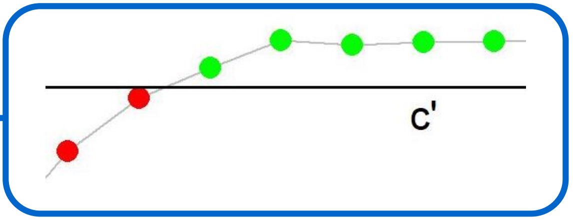
- If $P(p > p_0|r_{ped}, r_{adu})$ is monotonically increasing in r_{ped}, then there exists c' with

$$P(p > p_0|r_{ped}, r_{adu}) \geq c' \Leftrightarrow r_{ped} \geq b_{\text{UMP}}(\alpha) \ (*)$$

and $b_{\text{UMP}}(\alpha)$ is the level α UMP test boundary.

- If $P(p > p_0|r_{ped}, r_{adu})$ is not monotonically increasing in r_{ped}, then there are 3 options:

1. a threshold c' with $(*)$ can still be identified.

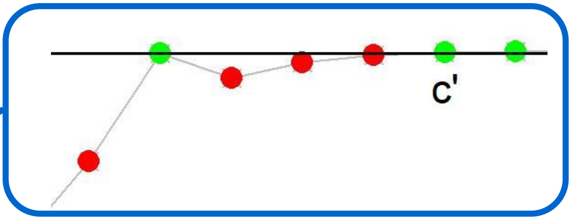


2. if no c' with $(*)$ can be identified, then either

   a. the test does not control type I error

   or

   b. the test controls type I error but is not UMP.



$\longrightarrow$ The trial may be considered a success for r_{ped} responses and a failure for one more pediatric response $(r_{ped} + 1)$

Summary

View decision rule as test function $\varphi(r_{ped}) = 1_{\{P(p>p_0|r_{ped},r_{adu})\geq c\}}$

$\rightarrow$ There is nothing better than the UMP test!

- This holds for all situations in which UMP tests exist:
  exponential family distribution
  one-sided tests, two-sided tests (equivalence situation)
  one-sided comparison of two normal variables …
- This should also hold in situations in which UMP unbiased tests exist
  since decision rule should be unbiased:
  two-sided comparisons
  comparison of two proportions …
- True for any (adaptive) borrowing mechanism (power prior, mixture prior, …)

- Proven by Psioda and Ibrahim (2018) for one-sample one-sided normal test with borrowing using a fixed power prior.

Conclusion

- If strong frequentist type I error control is desired in a situation where a UMP test exists, external information is effectively discarded.

- However, if prior information is reliable and consistent with the new information, the final operating characteristics of the trial can be improved: increased power or lower type I error, depending on where prior information is placed (but at expense of the other characteristic).

  → Incorporation of prior information can give a rationale for type I error inflation with benefit of a power gain.

References

- Gravestock I, Held L; COMBACTE-Net consortium (2017). Adaptive power priors with empirical Bayes for clinical trials. *Pharmaceutical Statistics* 16(5): 349-360.
- Kopp-Schneider A, Wiesenfarth M, Witt R, Edelmann D, Witt O, Abel U. (2018) Monitoring futility and efficacy in phase II trials with Bayesian posterior distributions – A calibration approach. *Biometrical Journal* online.
- Psioda MA, Ibrahim JG (2018) Bayesian clinical trial design using historical data that inform the treatment effect. *Biostatistics* online.
- Schmidli H, Gsteiger S, Roychoudhury S, O'Hagan A, Spiegelhalter D, Neuenschwander B (2014). Robust meta-analytic-predictive priors in clinical trials with historical control information. *Biometrics* 70(4):1023-32.