Incorporating historical information in biosimilar trials

Johanna Mielke, Heinz Schmidli, Byron Jones

Acknowledgement & disclaimer

This project was supported by the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract number 999754557. The project is part of the IDEAS European training network (http://www.ideas-itn.eu/) from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 633567.

The opinions expressed in this presentation and on the following slides are solely those of the authors and not necessarily those of Novartis, Sandoz, Bayer or the Swiss Government.

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

IDEAS
Improving drug development

Historical data and biosimilars

- In biosimilar development:
  - the originator product has already been on the market for several years when the biosimilar development begins
  - the originator was already studied very often, both prior to market authorization and in post-marketing studies
- Idea: incorporate this historical information into the Phase III studies that are used for the approval of the biosimilar with a Bayesian approach
  - Summarize historical data in a prior distribution
  - Combine historical data with data in new study using Bayes' theorem to obtain posterior distribution
- Challenge: Type I error rate inflation is expected

Considered setting

- Parallel groups design
- Binary endpoint (responder, non-responder)
- Goal is to confirm equivalence in response rates of the test (T) and reference (R) product:

$$H_0 : |p_R - p_T| \geq \Delta \text{ vs. } H_1 : |p_R - p_T| < \Delta$$

- For Bayesian approach:
  - Informative prior for the reference product
  - Non-informative prior for the test product since no information is available prior to the study
  - Combine prior with the observed data with Bayes' theorem
- Benchmark: two-one-sided-test (TOST) frequentist approach that is the standard approach which considers the response rates in the new study only

Test decision & settings

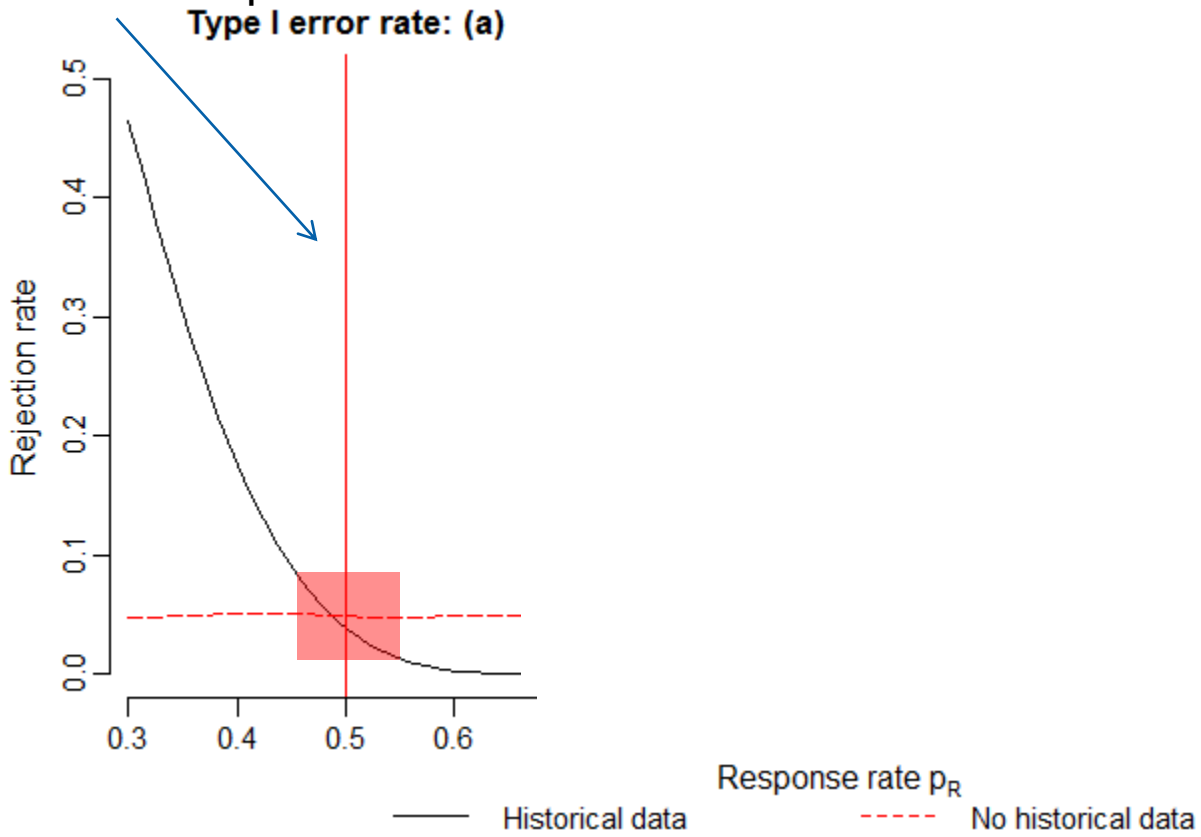
- Bayesian success criterion: Let X_R, X_T be random variables that follow the posterior distribution of T and R. Then, claim equivalence if:

$$B = P(|X_R - X_T| < \Delta) > c$$

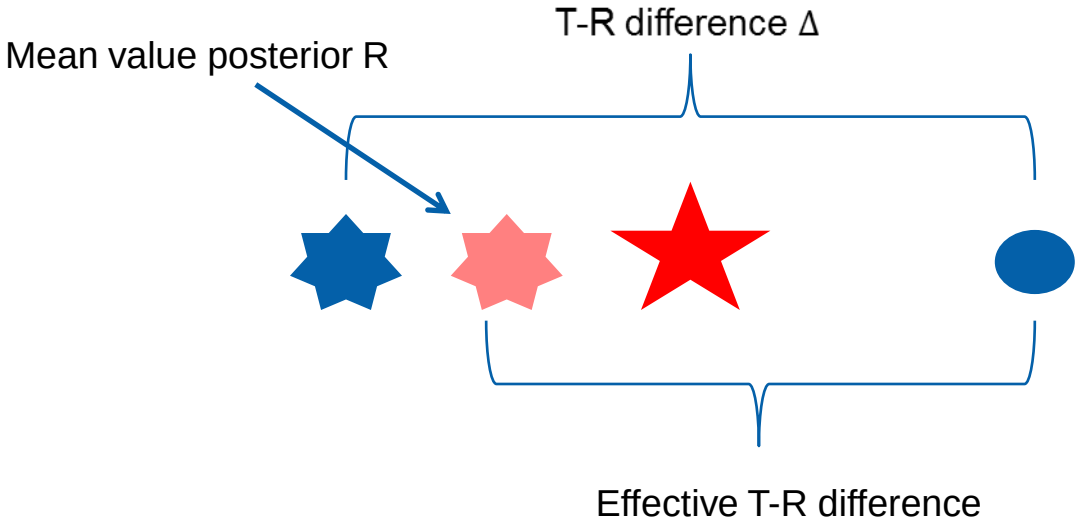
- We evaluate the operating characteristics for several true response rates for R and consider three different constellations of response rates for T (dependent on R):
  - $p_T = p_R + \Delta$ (Type I error rate: Situation (a))
  - $p_T = p_R - \Delta$ (Type I error rate: Situation (b))
  - $p_R = p_T$ (power)

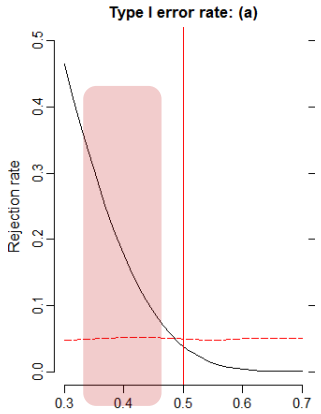
Operating characteristics

Mean value of the prior



Why do we observe this profile?





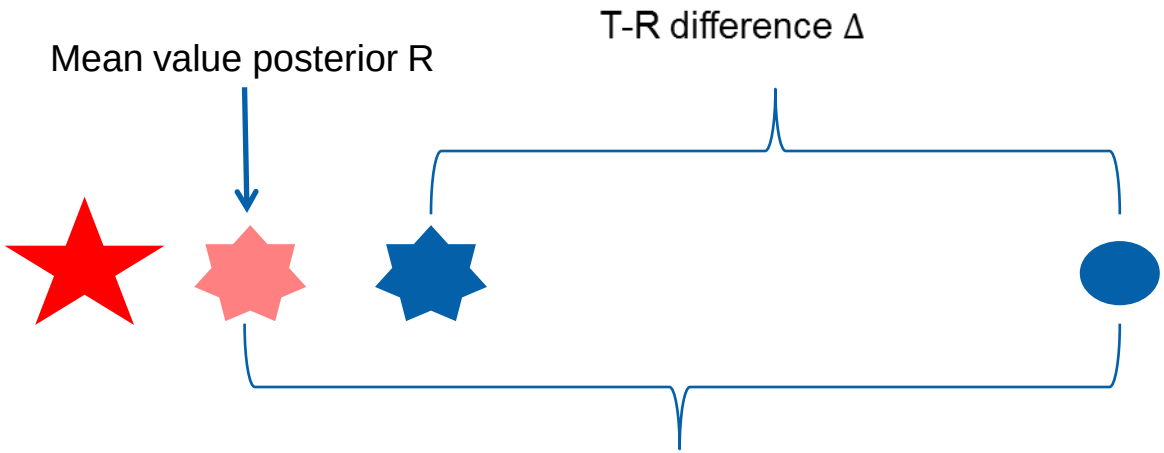
Response rate

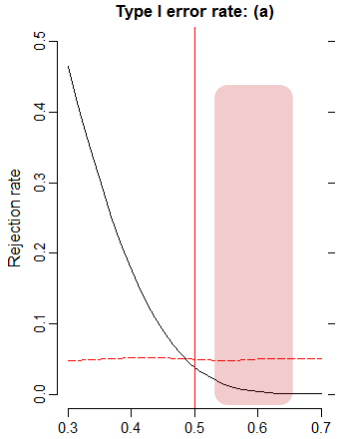
Mean value prior

Response rate T

Response rate R

Why do we observe this profile?





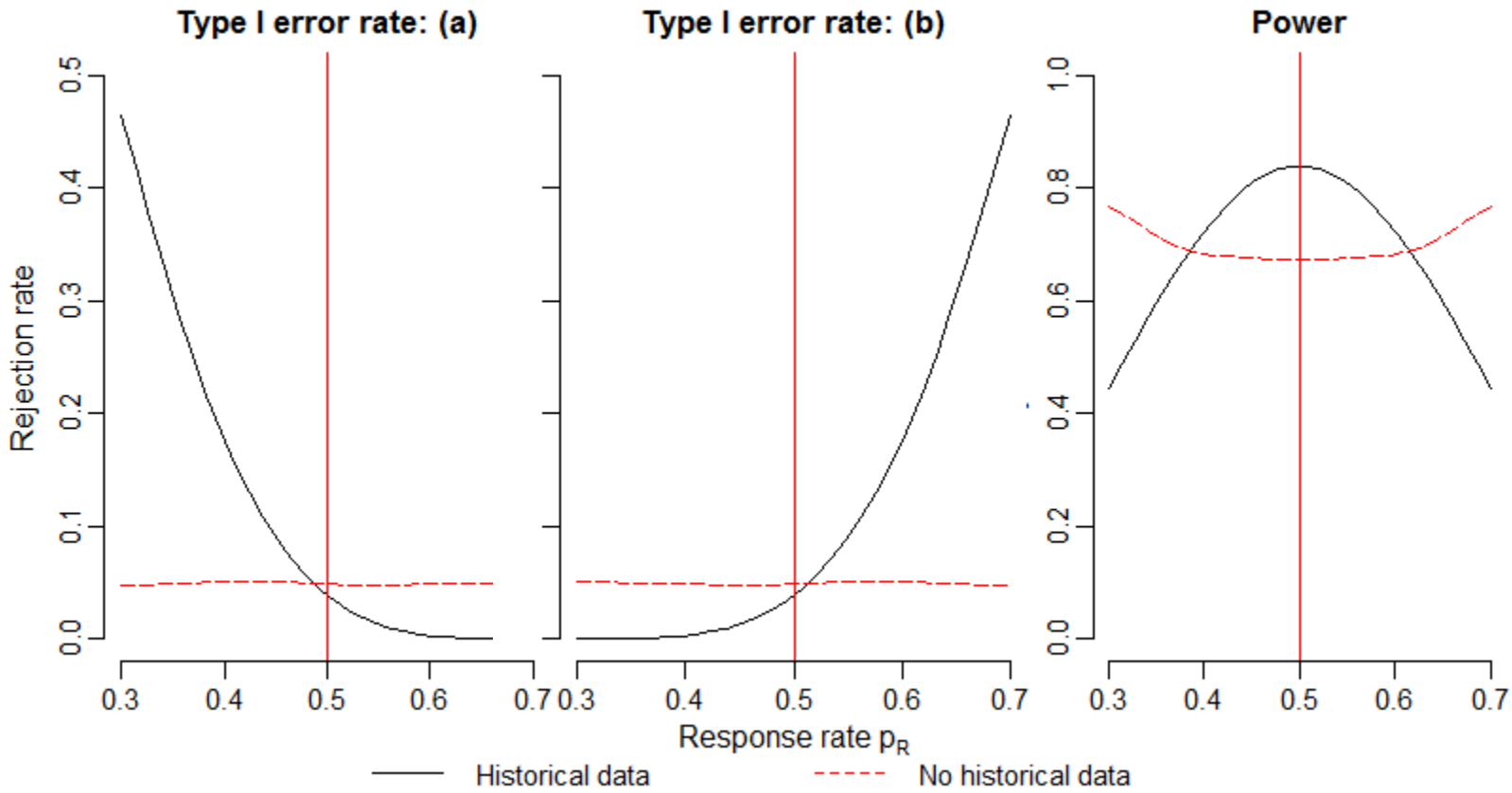
Response rate

Mean value prior

Response rate T

Response rate R

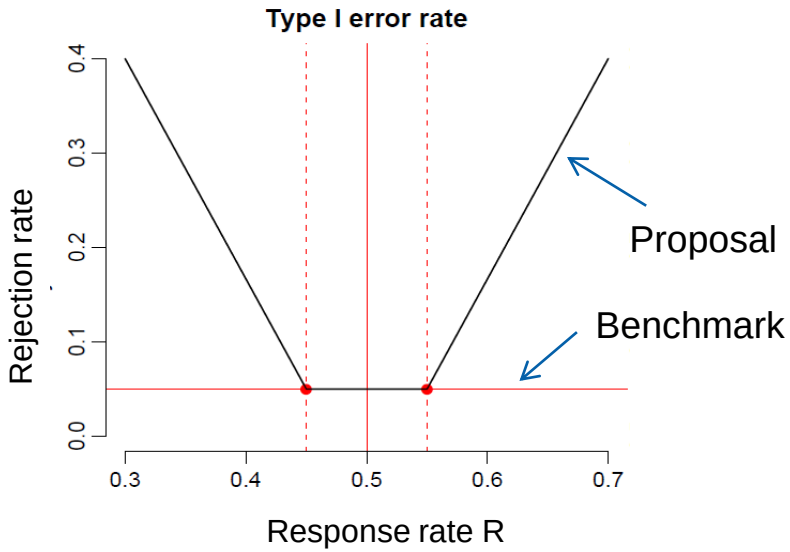
Operating characteristics

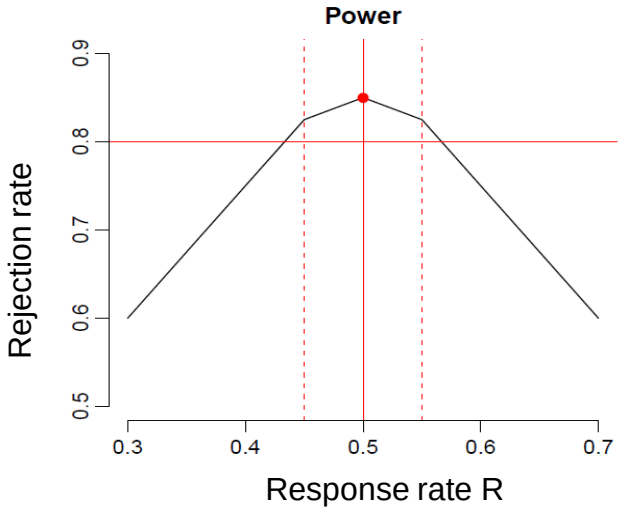


No gain in power is possible if Type I error rate is strictly controlled!

Partial Type I error rate control

- We accept that strict Type I error rate control is incompatible with a gain in power
- For biosimilars, we expect that it is possible to conduct a “similar” study
- We define an interval C in which we aim to control the Type I error rate
- Note: standard approaches do not give a relevant gain in power even if only partial Type I error rate control is required





The hybrid frequentist-Bayesian approach

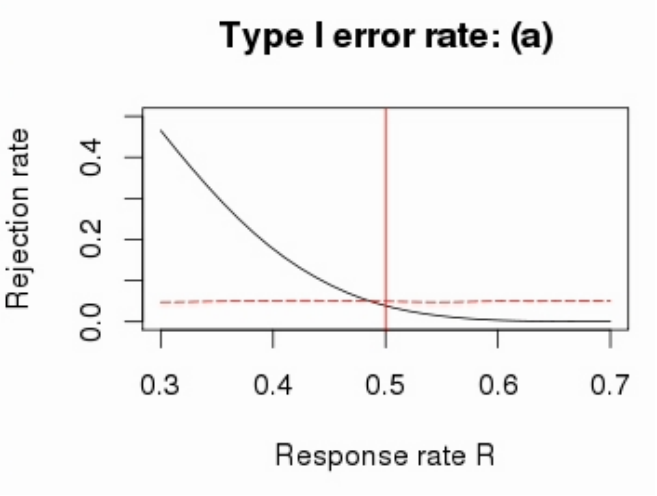
Overview of proposed method

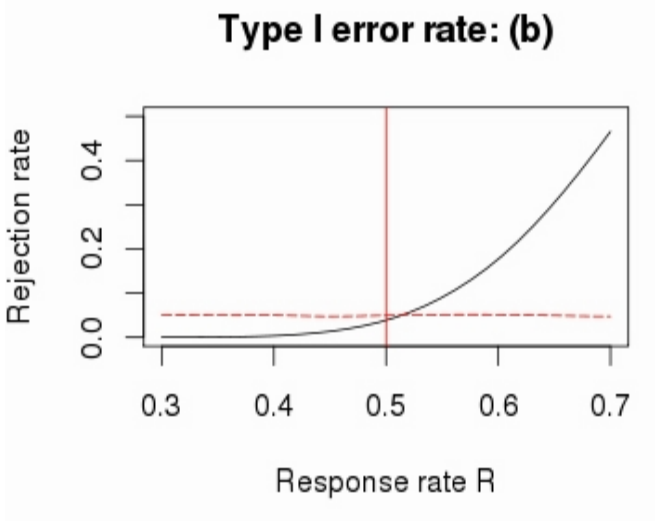
- Main goal: gain in power while controlling the Type I error rate in interval C
- Main concepts:
  - Switching rule I: if response rate of R in the new study and in the historical data are *very** different, do not use historical data
  - Switching rule II: if the response rates for T and R are *very** similar, use *lower** critical value
  - *Response rate-dependent critical values**

**: tuning parameters, can be chosen either automatically or be specified by the user*

Response rate-dependent critical values

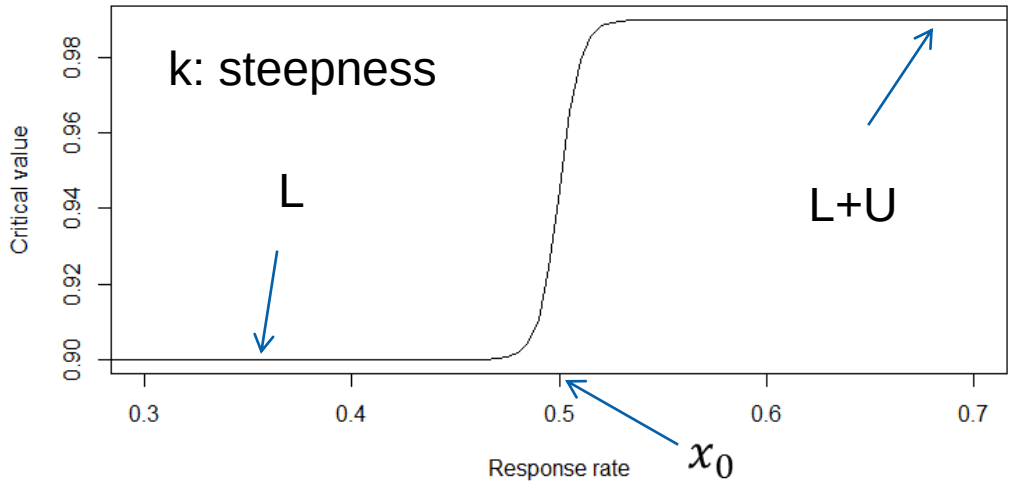
- The Type I error rate highly depends on the response rate in the new study
  ➔ Set the critical value high in regions in which the test is too liberal, and low in regions in which the test is too conservative
- The location of these regions depends on the ordering of the response rate of T and R in the new study (Situation (a) vs. Situation (b))
  ➔ Use different critical values for Situations (a) and (b)
- True response rate is not known
  ➔ Use estimated response rate





Response rate-dependent critical values

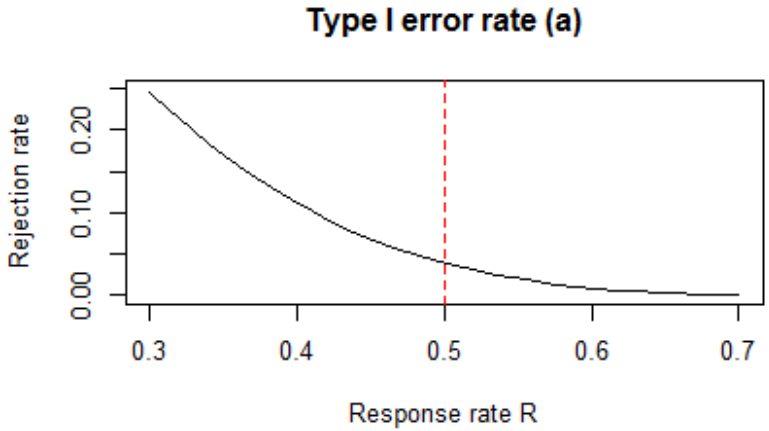
- Response rate-dependent critical values are chosen such that the Type I error rate is controlled in the interval C while the power is maximised under equality of response rates of T, R and the historical data
- Maximising a function without any assumptions on the functional form is difficult, we assume a logistic function

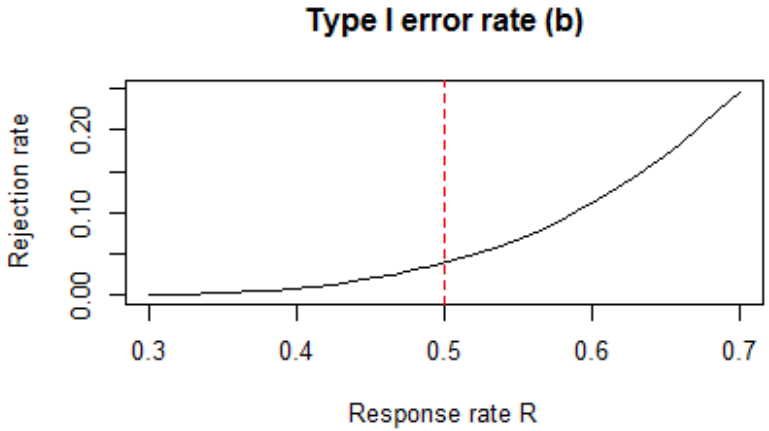


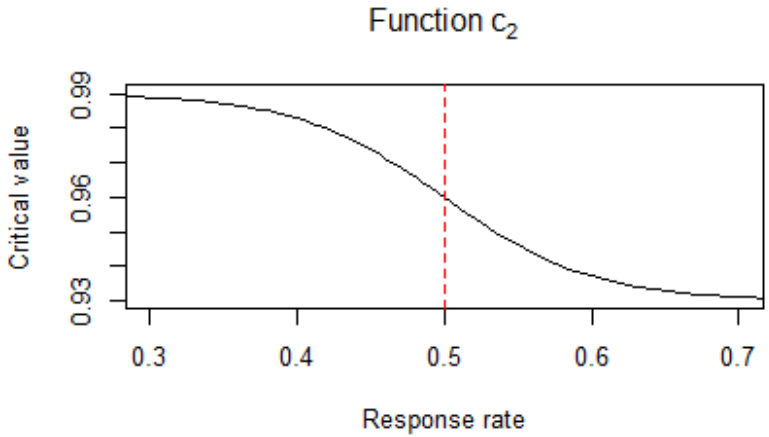
$$c_1(x) = L + \frac{U}{1 + \exp(-k(x - x_0))}$$

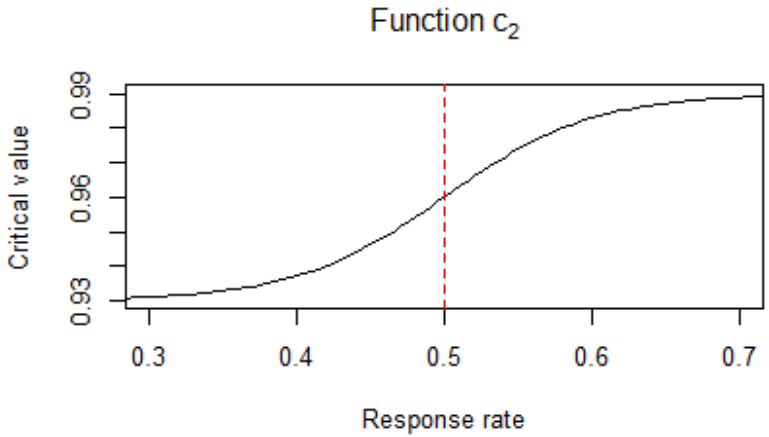
$$c_2(x) = L + \frac{U}{1 + \exp(k(x - x_0))}$$

Example: Response rate-dependent critical values

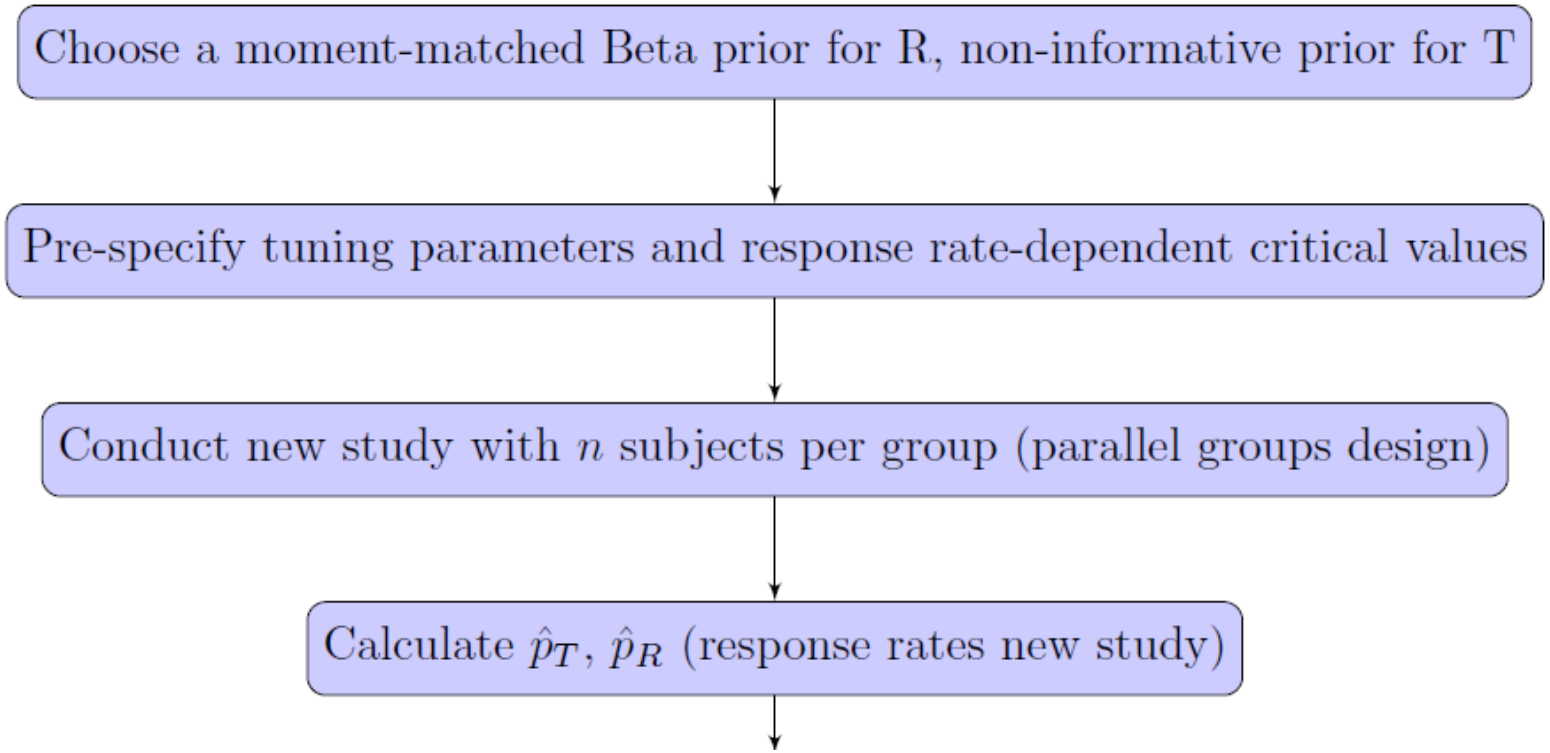




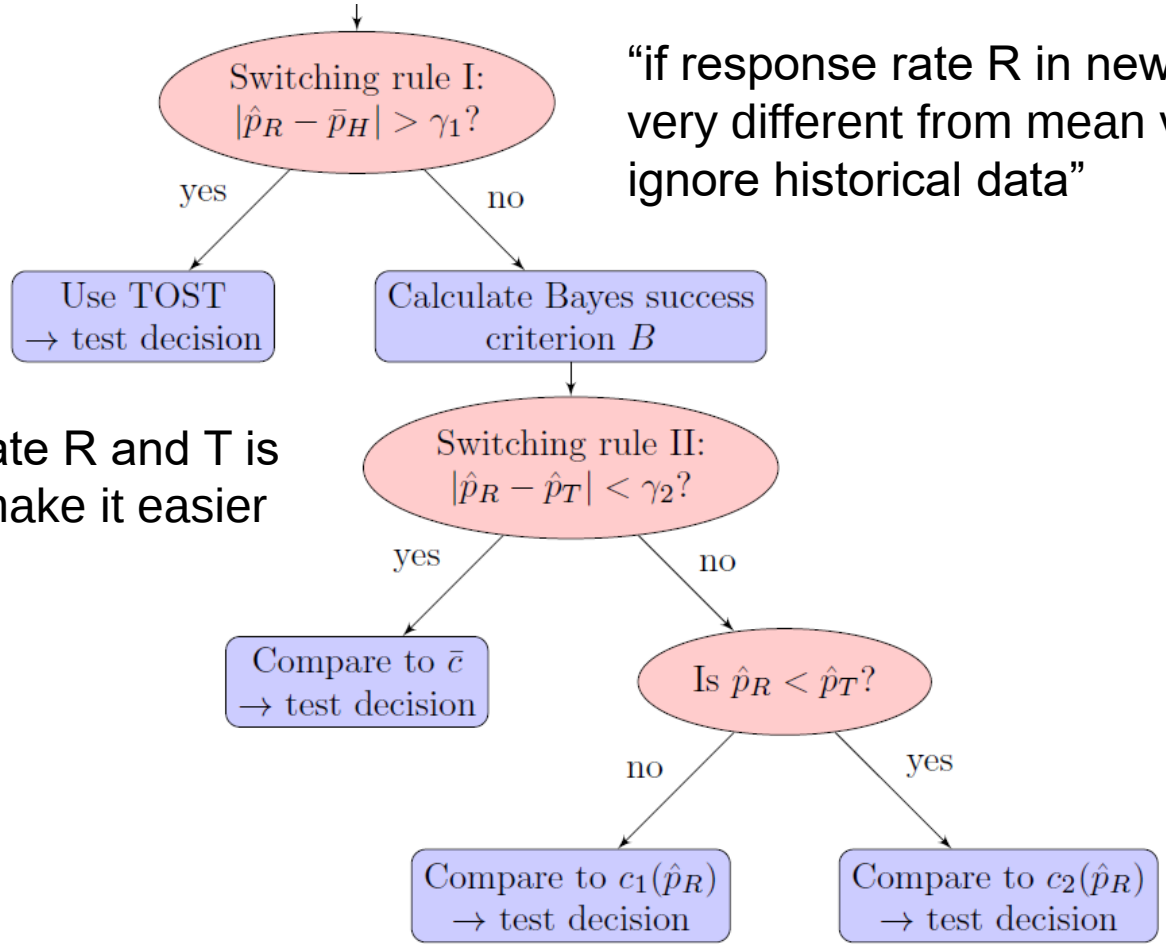




Proposed approach



Proposed approach



"if response rate R in new study is very different from mean value prior, ignore historical data"

"if response rate R and T is very similar, make it easier to reject"

Technical challenge: choice of tuning parameters

- In total 7 parameters have to be chosen:
  - 4 parameters of the response rate-dependent critical values
  - 3 tuning parameters
- In the binary case, it is possible to calculate exact rejection rates for a specific setting

$$r = \sum_{r_R=0}^{n} \sum_{r_T=0}^{n} P(X = r_T) P(Y = r_R) d_{r_T, r_R}$$

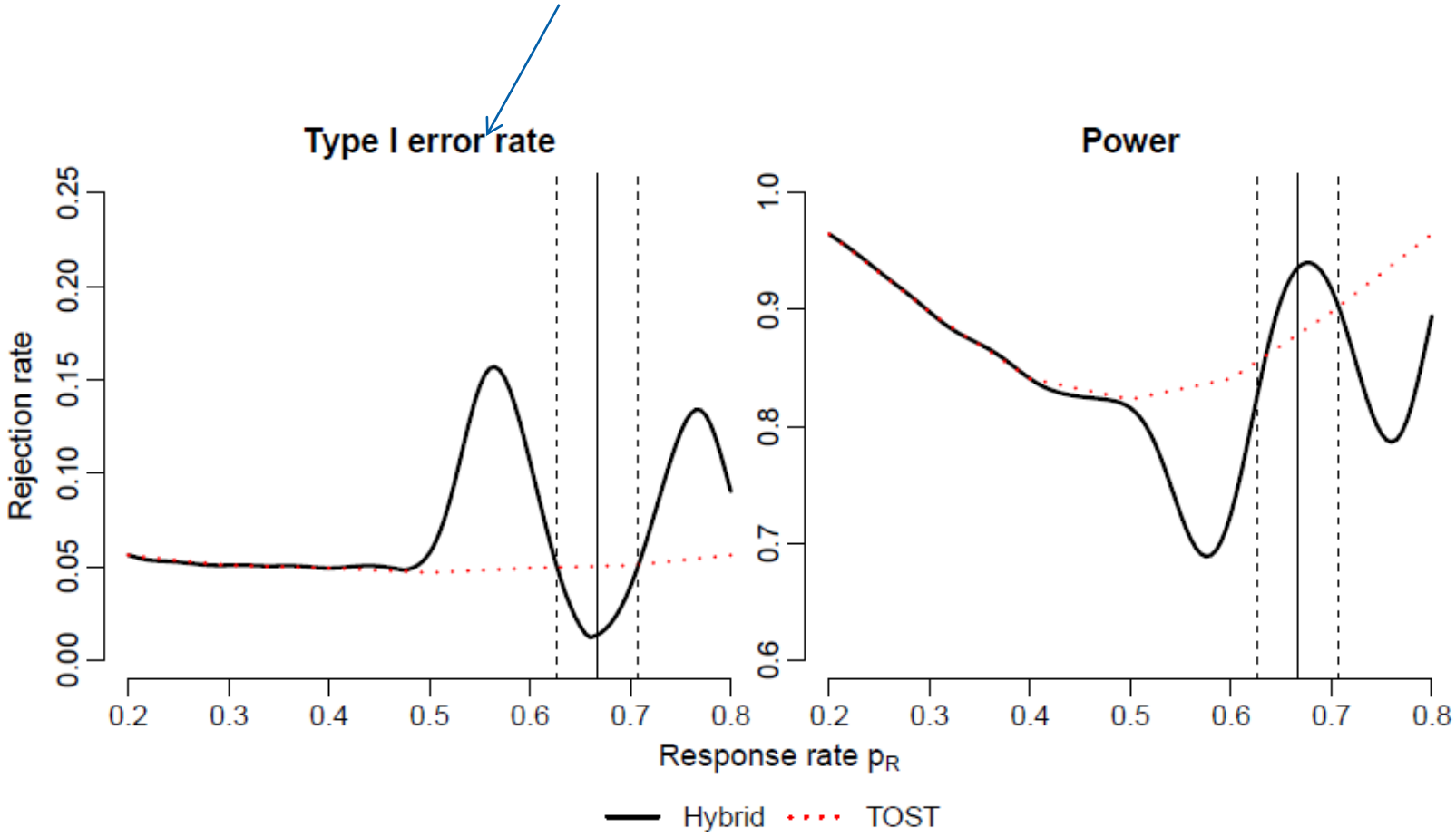
Test decision for the observed pair of responders

Probabilities to observe a specific number of responders

- Problems:
  - computational very expensive (n=150: 22801 settings)
  - local optima, flat curve
- We propose an algorithm, but recommend manual fine tuning

Example

Maximum of Situation (a) and (b)!



Case study

Planning of a hypothetical study

- Phase III study for a proposed biosimilar with the active substance adalimumab (Humira)
- Indication: Psoriasis
- Endpoint: PASI 90 (Psoriasis Area and Severity Index) responder rate at week 16
- Equivalence margin: $\Delta = 0.15$
- Sample size new study: $n = 175$

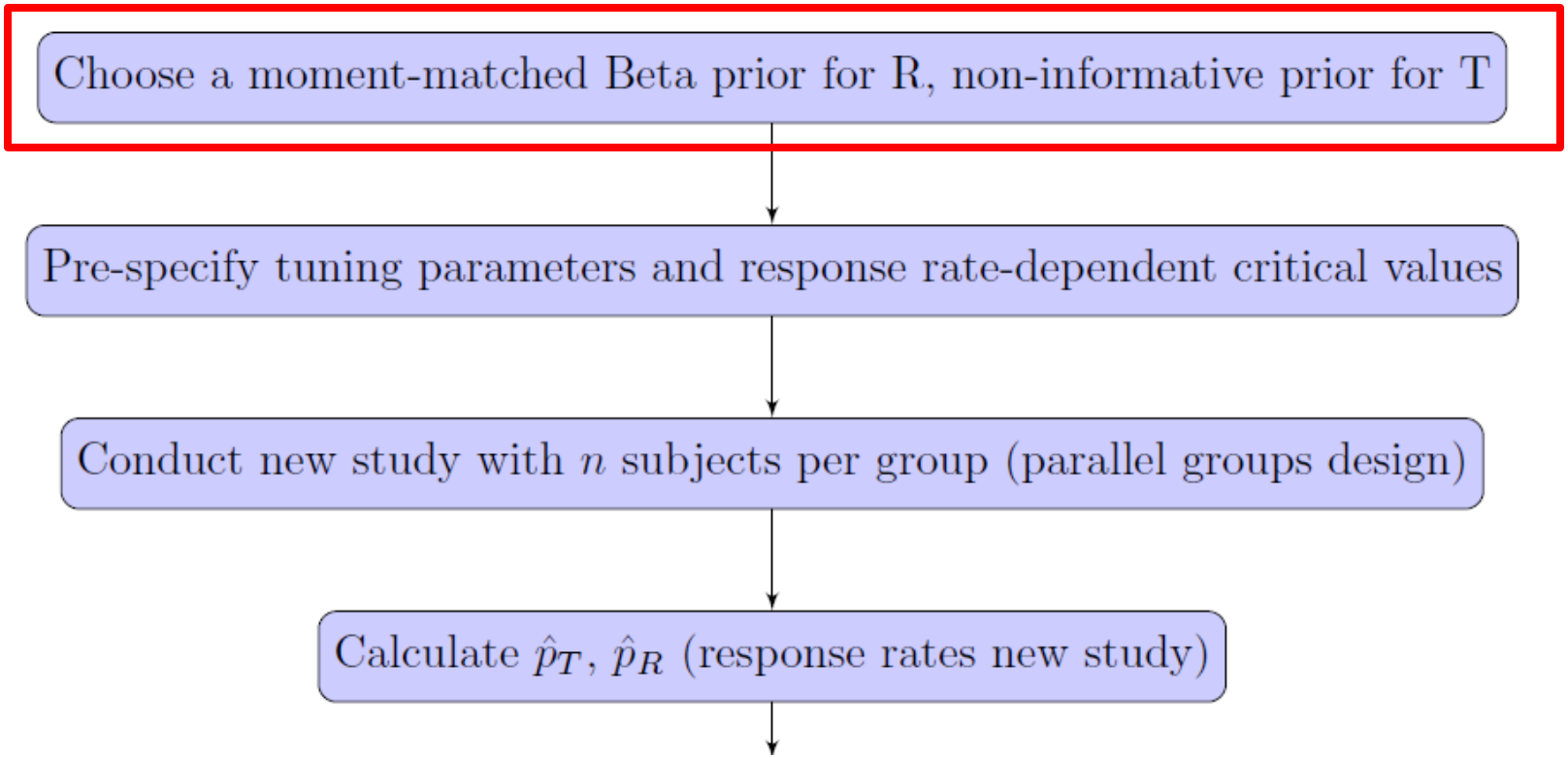
Historical information

Study	Publication	Indication	Responders/Sample size (%)
1	Menter *et al.* (2008)	moderate to severe psoriasis	366/814 (45.0)*
2	Saurat *et al.* (2008)	moderate to severe plaque psoriasis	55/108 (51.3)*
3	Thaçi *et al.* (2010)	moderate to severe psoriasis	183/364 (50.0)
4	Blauvelt *et al.* (2017)	moderate to severe psoriasis	166/334 (49.7)
5	Reich *et al.* (2017)	moderate to severe psoriasis	116/248 (46.8)
Total		886/1858 (47.7)	

- Choice of interval C:

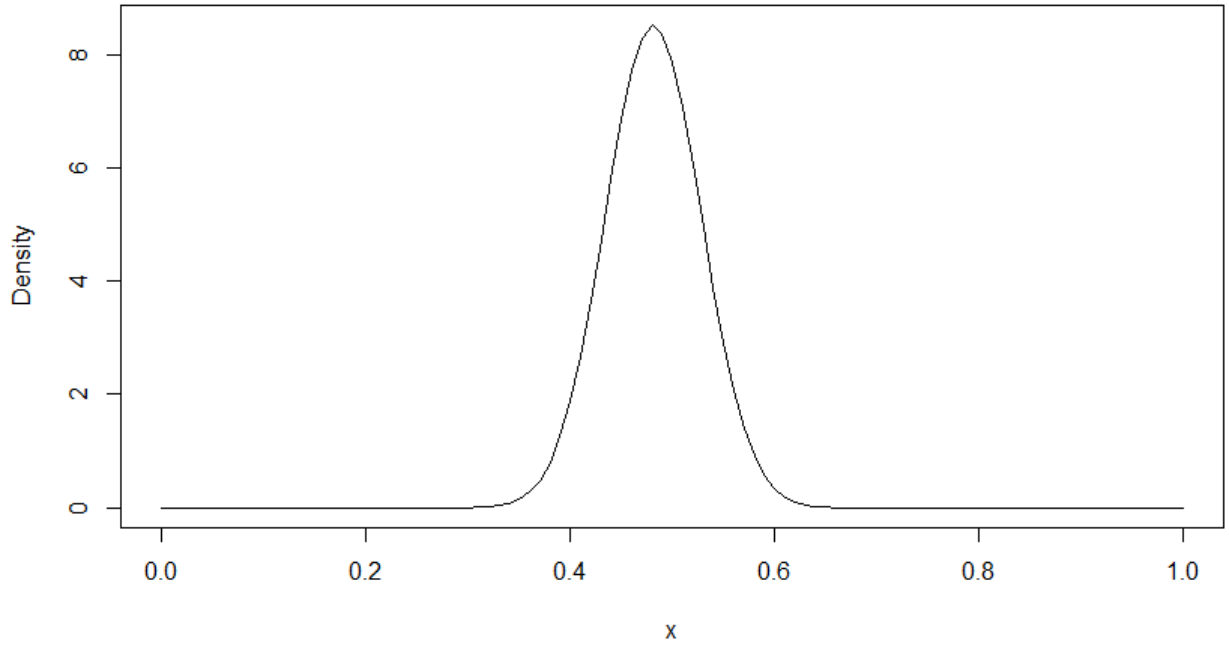
$$C = [\bar{p}_H - 0.05, \bar{p}_H + 0.05] = [0.4313, 0.5313]$$

Application of the proposed method

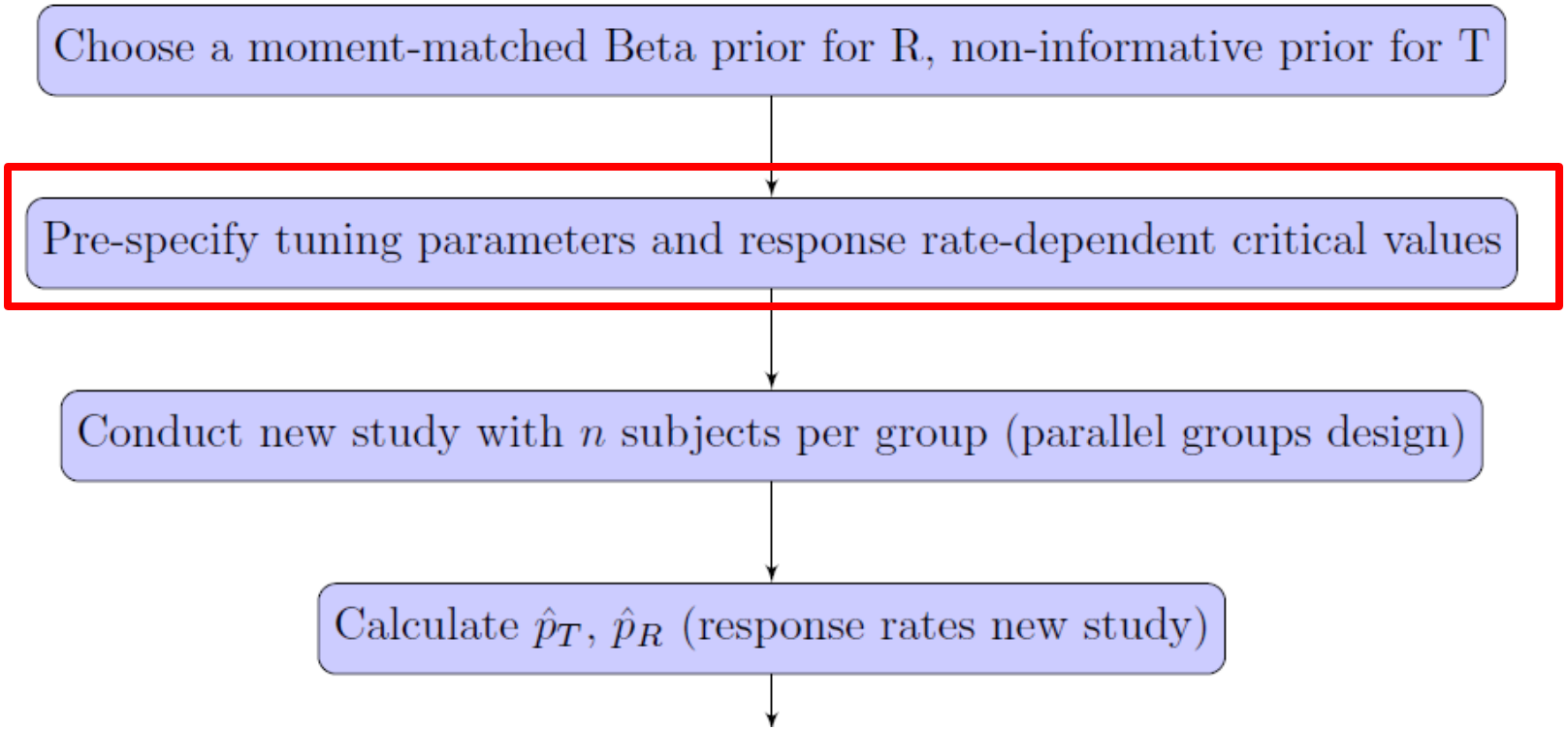


Step 1: Derivation of MAP prior

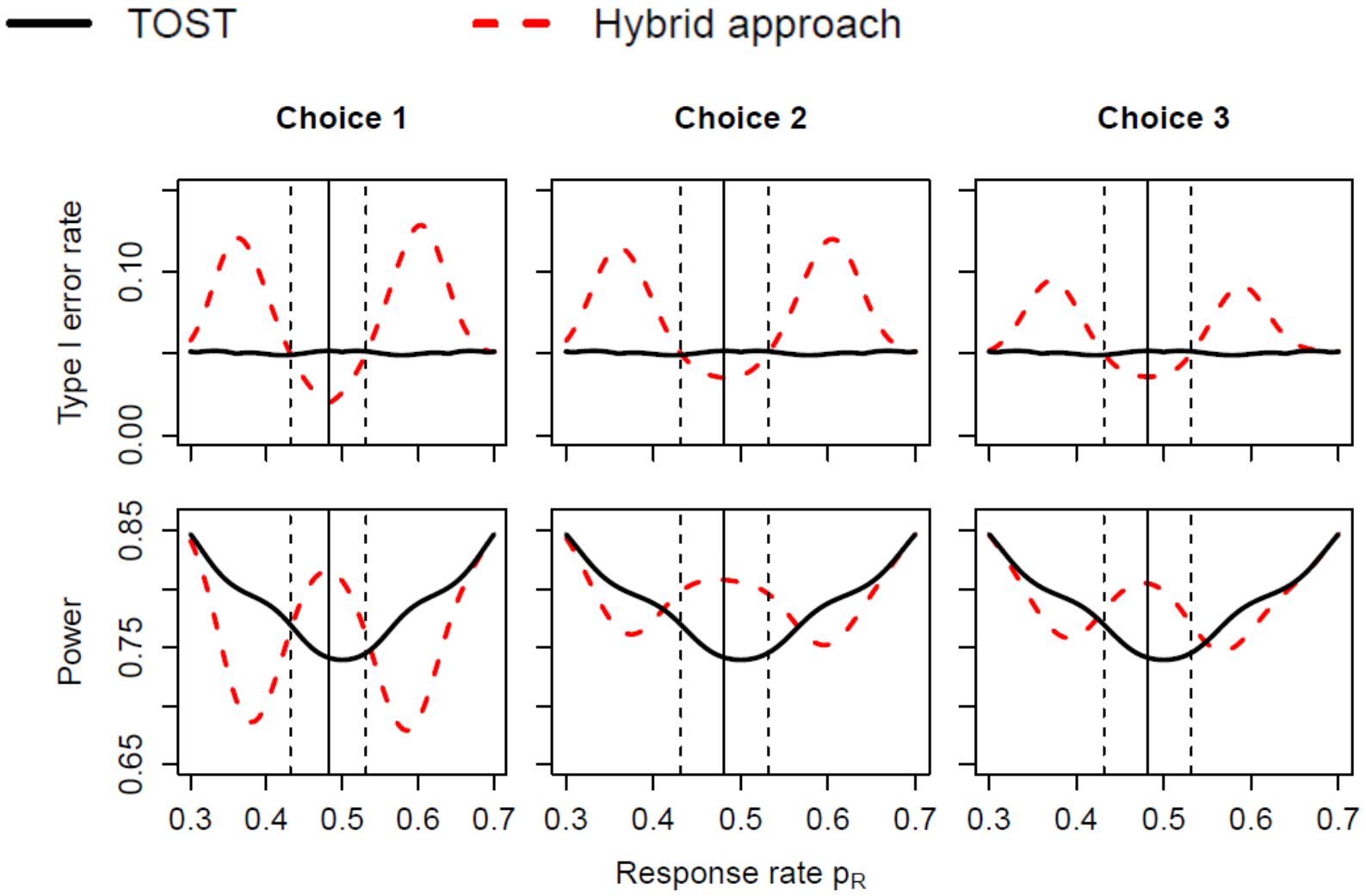
- R-package RBesT (Weber, 2017)
- Use of default assumption for hyper-parameters
- Best fit: Beta distribution with $a = 55.0844, b = 59.3647$



Application of the proposed method



Step 2: Choice of tuning parameters

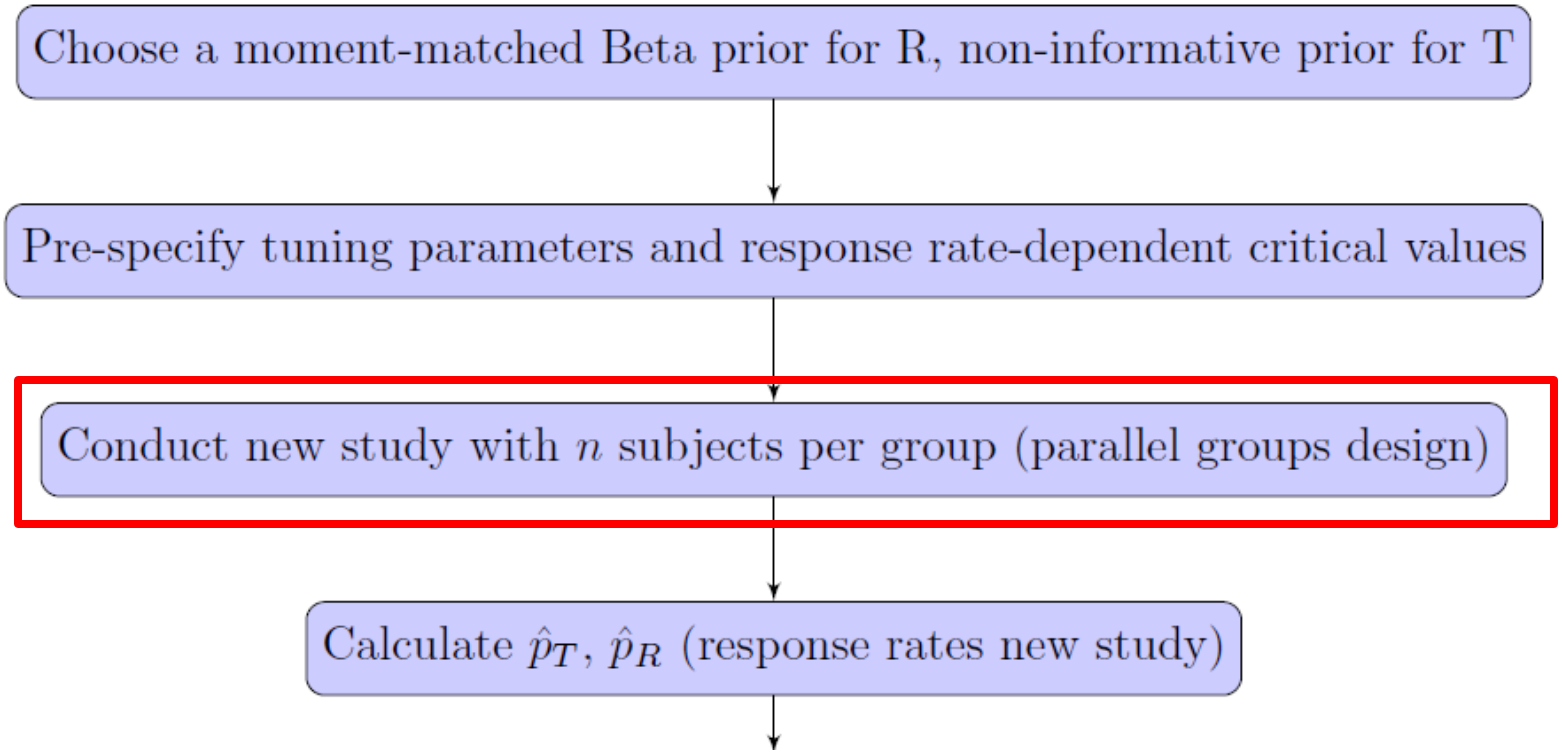


$\gamma_1 = 0.1133$
$\gamma_2 = 0.0378$
$c = 0.9$

$\gamma_1 = 0.1133$
$\gamma_2 = 0.0567$
$c = 0.9$

$\gamma_1 = 0.0944$
$\gamma_2 = 0.0567$
$c = 0.9$

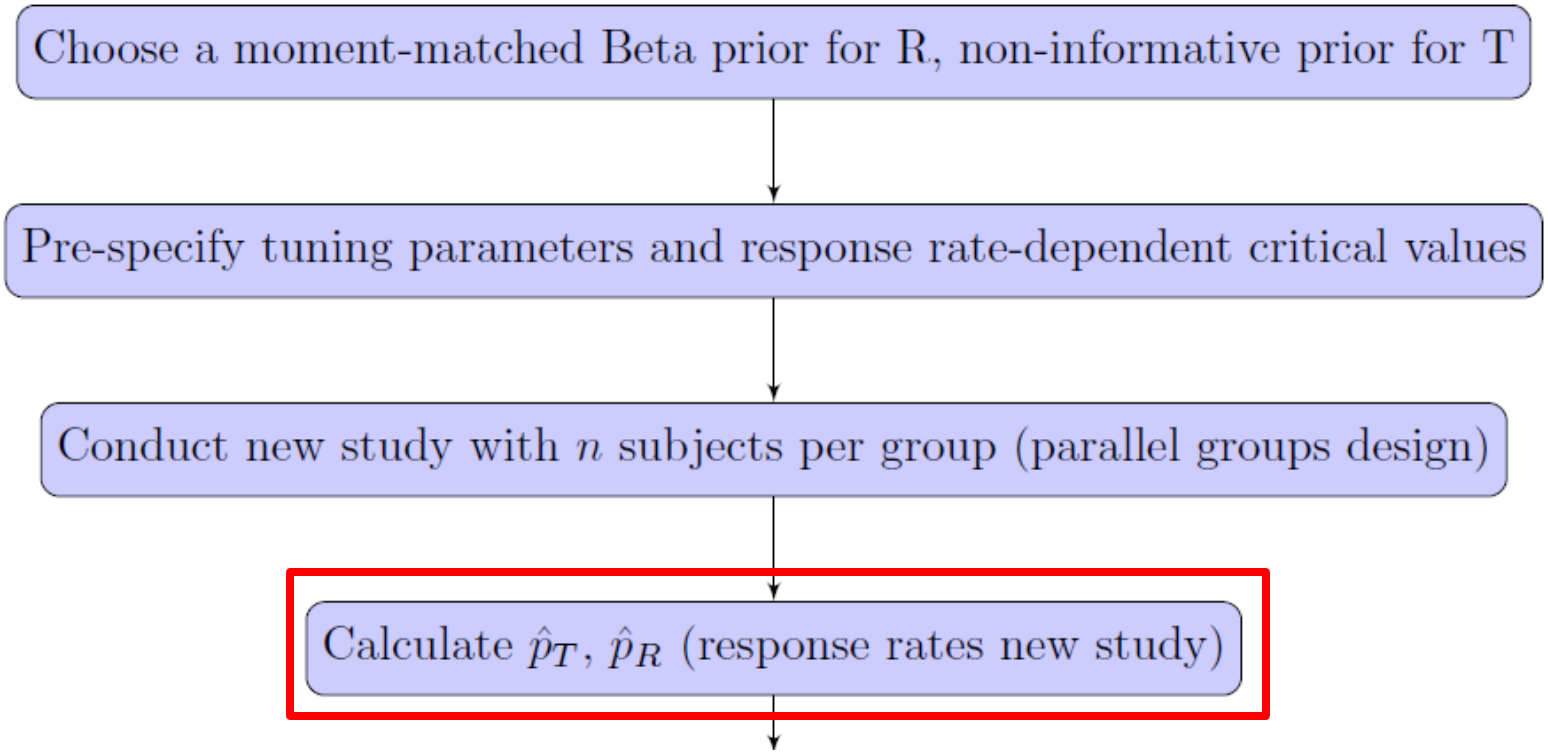
Application of the proposed method



Step 3: Conduct new study

- We consider the Phase III in psoriasis which was undertaken for the approval of Amgevita (Amgen)
  - Indication: stable moderate to severe plaque psoriasis
  - One of the endpoints: PASI 90 at week 16
  - 172 evaluated subjects in the test (biosimilar) group
  - 173 evaluated subjects in the reference (originator) group

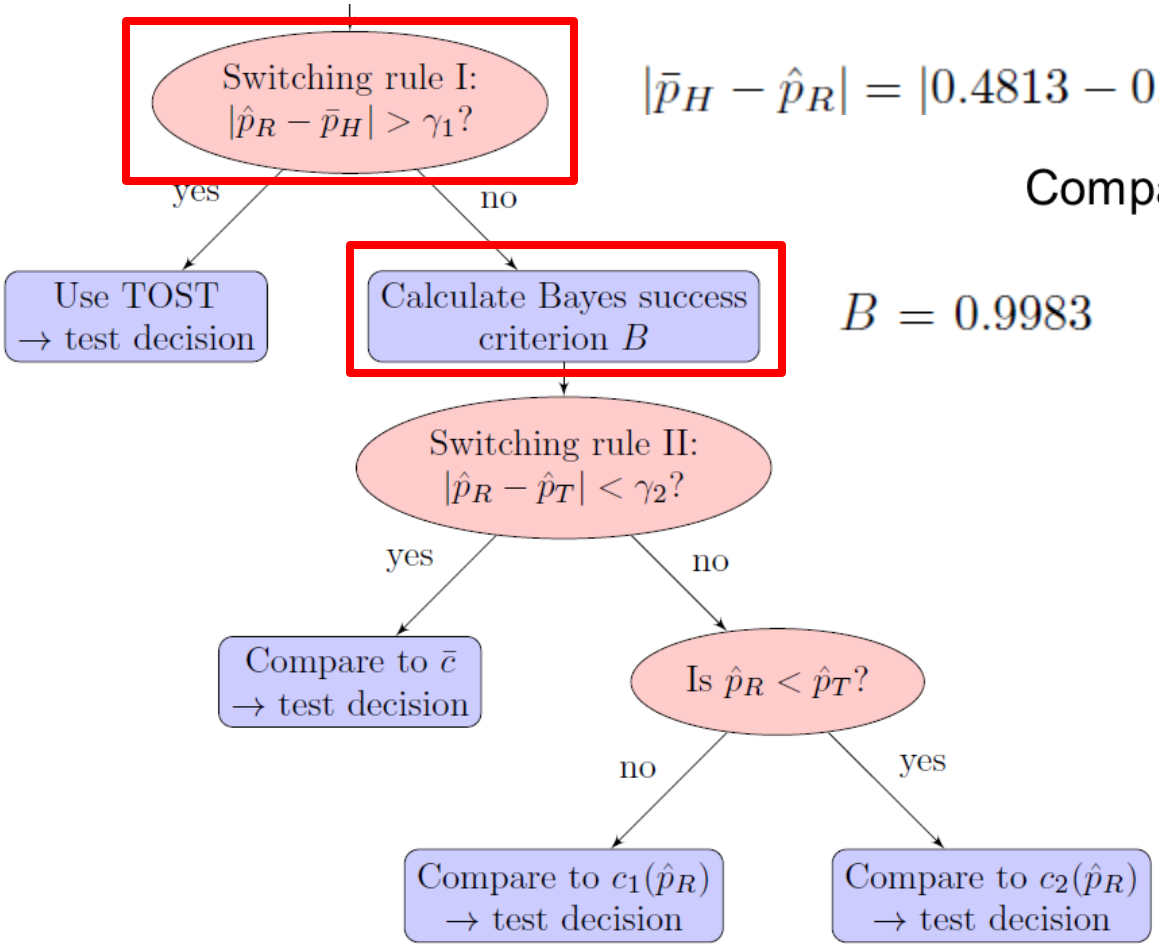
Application of the proposed method



Step 4: Study results

- 81 of 172 subjects responded on test (0.471)
- 82 of 173 subjects responded on reference (0.474)

Application of proposed method

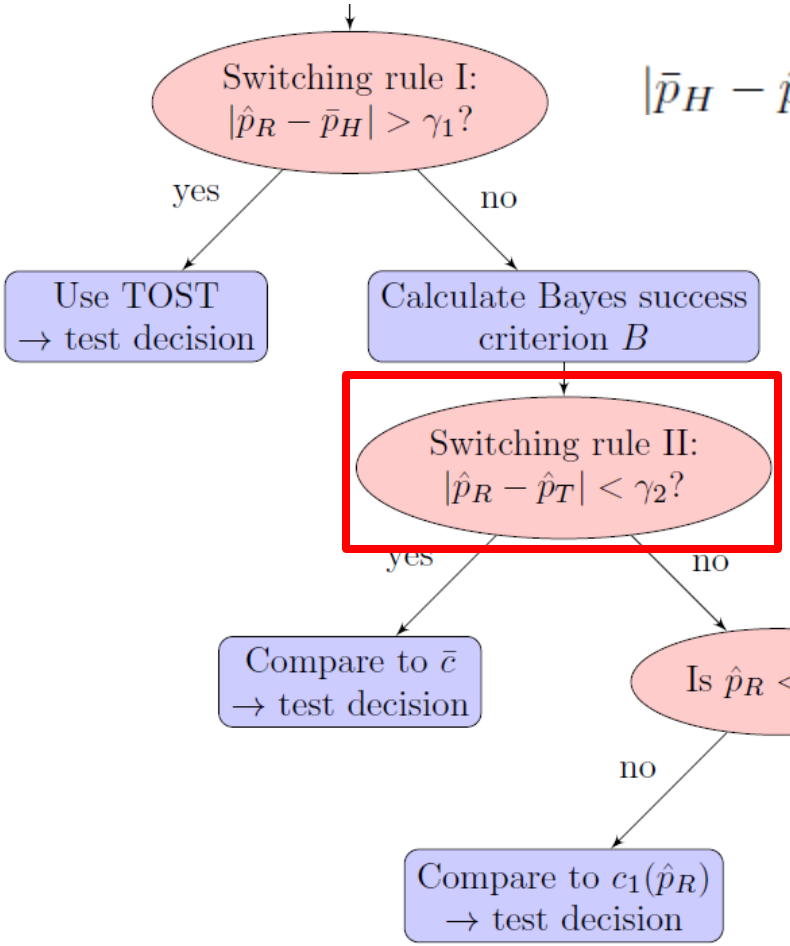


$$|\bar{p}_H - \hat{p}_R| = |0.4813 - 0.474| = 0.0073$$

Compare to: $\gamma_1 = 0.1133$

$$B = 0.9983$$

Application of proposed method



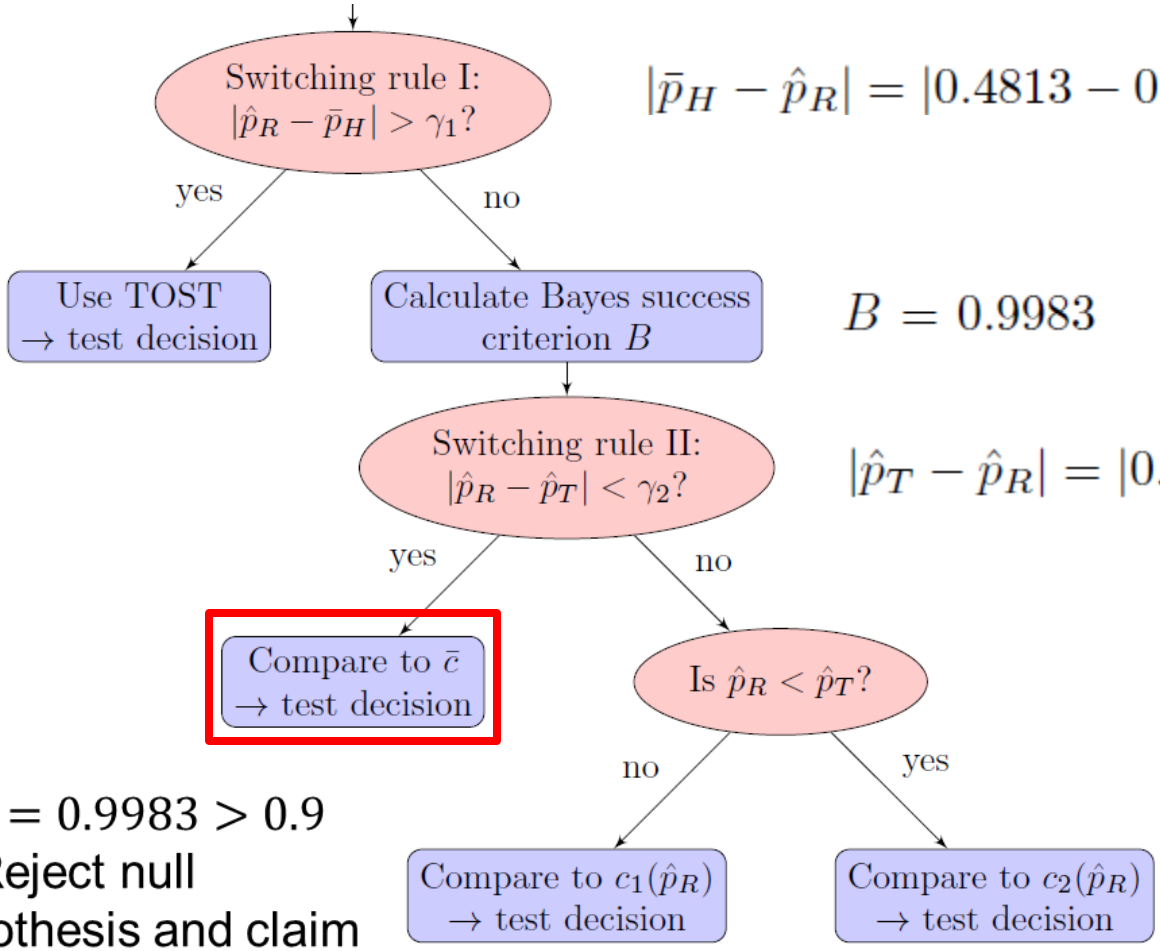
$$|\bar{p}_H - \hat{p}_R| = |0.4813 - 0.474| = 0.0073$$

$$B = 0.9983$$

$$|\hat{p}_T - \hat{p}_R| = |0.471 - 0.474| = 0.003$$

Compare to: $\gamma_2 = 0.0567$

Application of proposed method



$$|\bar{p}_H - \hat{p}_R| = |0.4813 - 0.474| = 0.0073$$

$$B = 0.9983$$

$$|\hat{p}_T - \hat{p}_R| = |0.471 - 0.474| = 0.003$$

$B = 0.9983 > 0.9$

➔ Reject null hypothesis and claim equivalence

Discussion

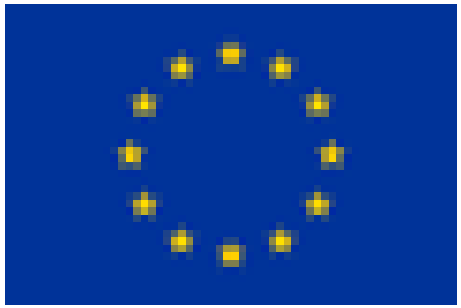
Discussion

- Proposed approach provides a gain in power in comparison to not using historical data while controlling the Type I error rate in the interval C
- Choice of the interval C dependent on knowledge and confidence in conducting a new study which is similar to the historical studies
- The choice of the response rate-dependent critical values and tuning parameters for the switching rules is computationally very expensive, but not difficult for the user to perform

Mielke, J., Schmidli, H. and Jones, B. (2018c): Incorporating historical information in biosimilar trials: challenges and a hybrid Bayesian-frequentist approach. *Biometrical Journal*, 60 (3), 564-582.

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Thank you



This project was supported by the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract number 999754557. The opinions expressed and arguments employed herein do not necessarily reflect the official views of the Swiss Government. The project is part of the IDEAS European training network (http://www.ideas-itn.eu/) from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 633567.