Response-Adaptive Randomization and Adaptive Combination Test for Clinical Trials with Limited Number of Patients: Practical Guide

S. Krasnozhon[1], D. Schindler[2], R.-D. Hilgers[2], N. Heussen[2], W.F. Rosenberger[3] and F. König[1]

[1]Section for Medical Statistics/CeMSIIS, Medical University of Vienna, Wien, Austria,
[2]Department of Medical Statistics, RWTH Aachen University, Aachen, Germany,
[3]George Mason University, Fairfax, Virginia, USA

June 24 - 26, 2015, Cologne, Germany

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement number FP HEALTH 2013–602552.

Outline

Aims

- investigate **Response-Adaptive (RA) Randomization Procedures** for **Small Population Two-Arm Clinical Trials**
  - Urn Models
  - Sequential Estimation Designs
- discuss extensions to **Three-Arm 'Gold Standard' Non-Inferiority Trials**
- scrutinise **Adaptive Designs (AD)** using adaptive combination tests and investigate influence of
  - the number and timing of interim analyses (IA)
  - adaptation of allocation ratios
  - sample size reassessment

Outline

Statistical Model

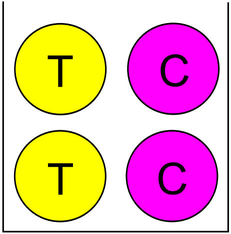
- Consider **two treatment groups** - treatment (T) and control (C).
- Y_n is a **response** of patient n (binary or continuous).
- Consider the **hypothesis**

$$H_0 : \theta_C = \theta_T \quad \text{versus} \quad H_1 : \theta_C \neq \theta_T$$

at **level** α (e.g., $\alpha = 0.05$).

The Klein (KLEIN) urn design (Galbete et al. [2014])

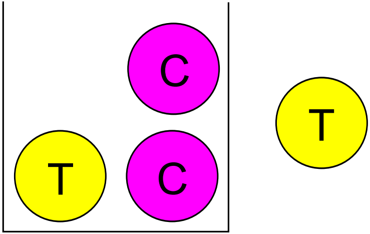
- an **urn** with $2w$ balls of type 'T' and type 'C';



$$P(T_{n+1} = T | previous\ Responses,\ Allocations) = \frac{w + F_{n,C} - F_{n,T}}{2w}$$

The Klein (KLEIN) urn design (Galbete et al. [2014])

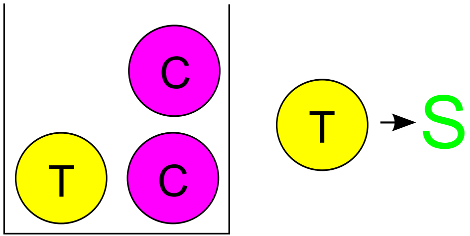
- an **urn** with $2w$ balls of type 'T' and type 'C';



$$P(T_{n+1} = T | previous\ Responses,\ Allocations) = \frac{w + F_{n,C} - F_{n,T}}{2w}$$

The Klein (KLEIN) urn design (Galbete et al. [2014])

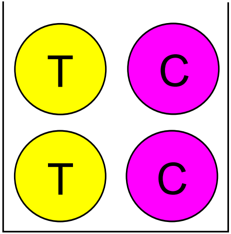
- an **urn** with $2w$ balls of type 'T' and type 'C';



$$P(T_{n+1} = T | previous\ Responses,\ Allocations) = \frac{w + F_{n,C} - F_{n,T}}{2w}$$

The Klein (KLEIN) urn design (Galbete et al. [2014])

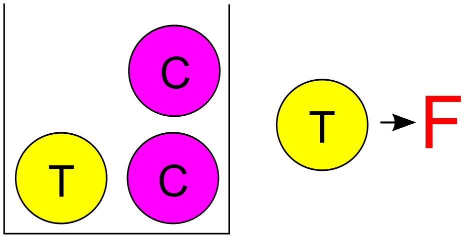
- an **urn** with $2w$ balls of type 'T' and type 'C';



$$P(T_{n+1} = T | previous\ Responses,\ Allocations) = \frac{w + F_{n,C} - F_{n,T}}{2w}$$

The Klein (KLEIN) urn design (Galbete et al. [2014])

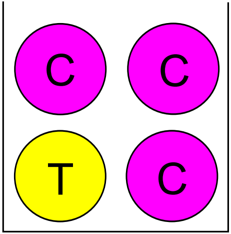
- an **urn** with $2w$ balls of type 'T' and type 'C';



$$P(T_{n+1} = T | previous\ Responses,\ Allocations) = \frac{w + F_{n,C} - F_{n,T}}{2w}$$

The Klein (KLEIN) urn design (Galbete et al. [2014])

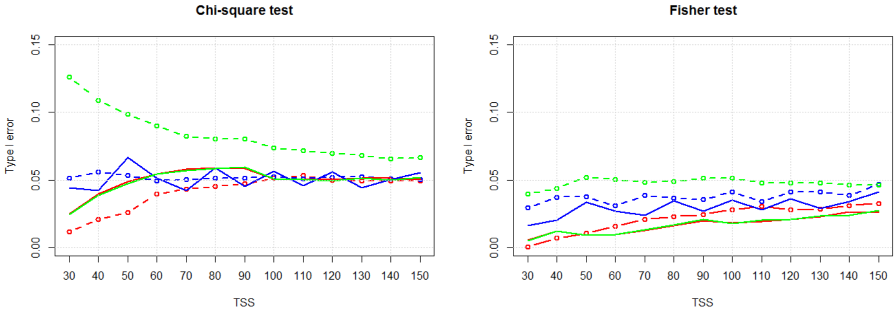
- an **urn** with $2w$ balls of type 'T' and type 'C';



$$P(T_{n+1} = T | previous\ Responses,\ Allocations) = \frac{w + F_{n,C} - F_{n,T}}{2w}$$

Simulation Results (KLEIN)

- $w = 1$.



$p_T = p_C = 0.1$

$p_T = p_C = 0.5$

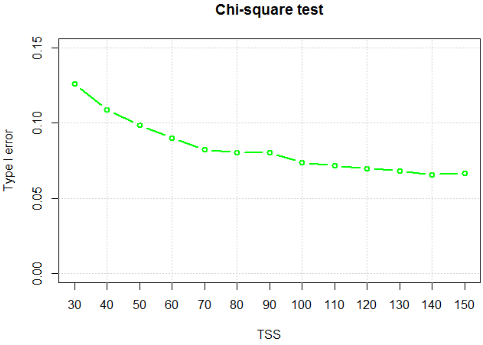
$p_T = p_C = 0.9$

0 - - 0 KLEIN design

——— Equal Allocation design

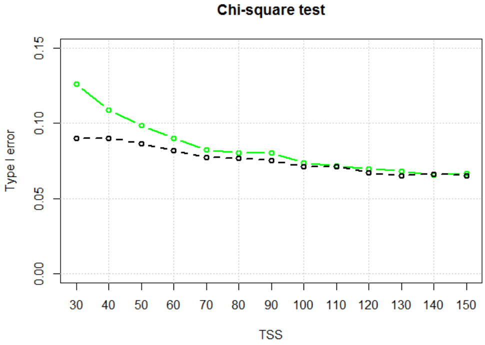
Simulation Results (KLEIN)

- $p_T = p_C = 0.9$;
- $w = \{1\}$.



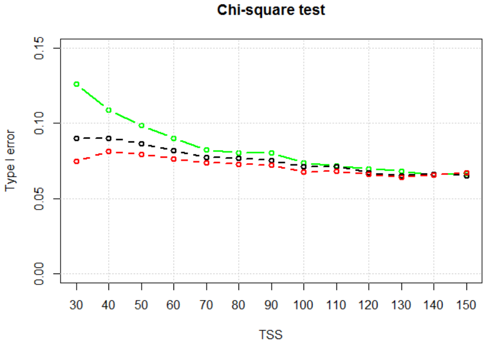
Simulation Results (KLEIN)

- $p_T = p_C = 0.9$;
- $w = \{1, 2\}$.



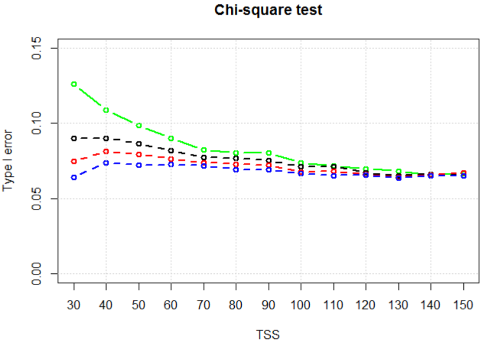
Simulation Results (KLEIN)

- $p_T = p_C = 0.9$;
- $w = \{1, 2, 3\}$.



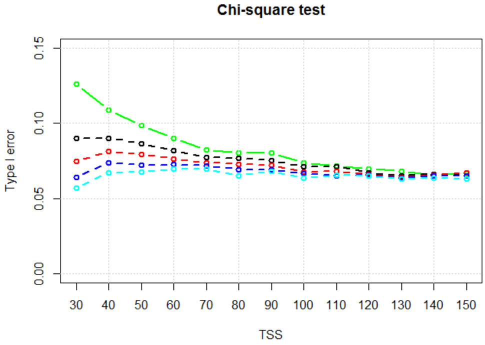
Simulation Results (KLEIN)

- $p_T = p_C = 0.9$;
- $w = \{1, 2, 3, 4\}$.



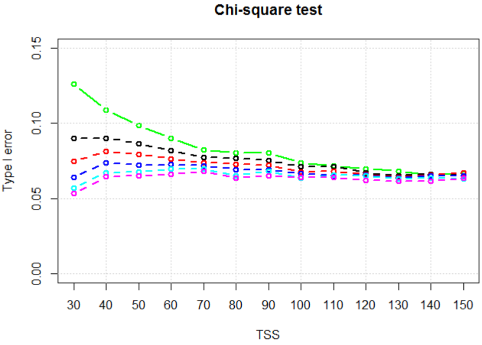
Simulation Results (KLEIN)

- $p_T = p_C = 0.9$;
- $w = \{1, 2, 3, 4, 5\}$.



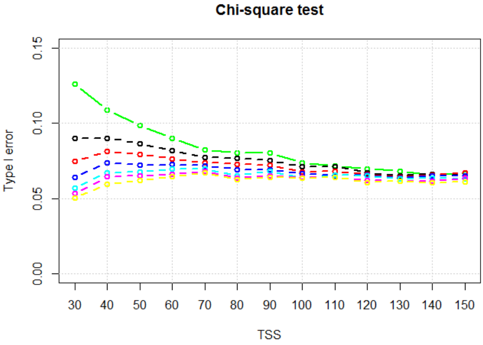
Simulation Results (KLEIN)

- $p_T = p_C = 0.9$;
- $w = \{1, 2, 3, 4, 5, 6\}$.



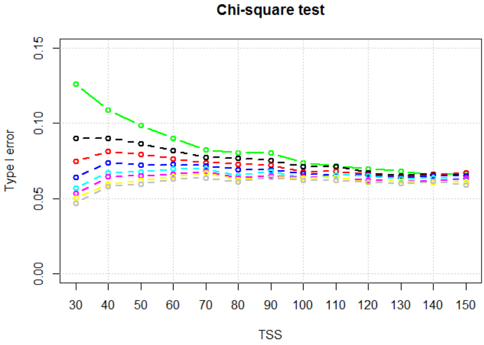
Simulation Results (KLEIN)

- $p_T = p_C = 0.9$;
- $w = \{1, 2, 3, 4, 5, 6, 7\}$.



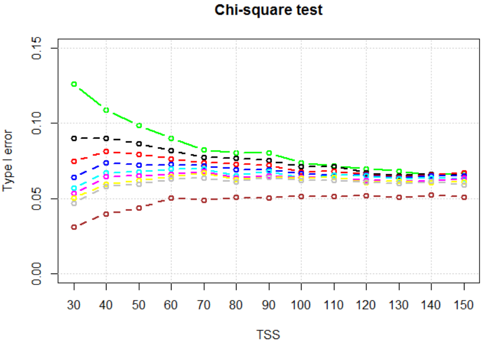
Simulation Results (KLEIN)

- $p_T = p_C = 0.9$;
- $w = \{1, 2, 3, 4, 5, 6, 7, 8\}$.



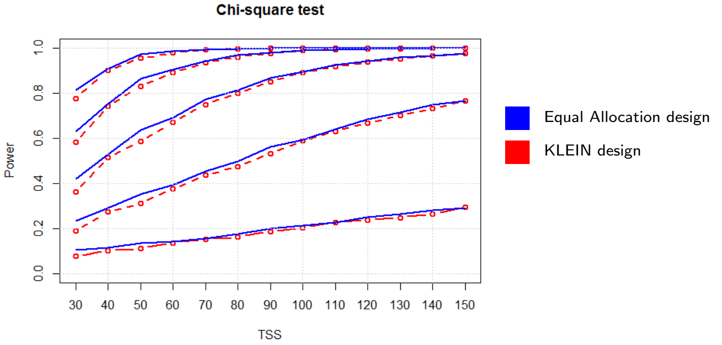
Simulation Results (KLEIN)

- $p_T = p_C = 0.9$;
- $w = \{1, 2, 3, 4, 5, 6, 7, 8, 100\}$.



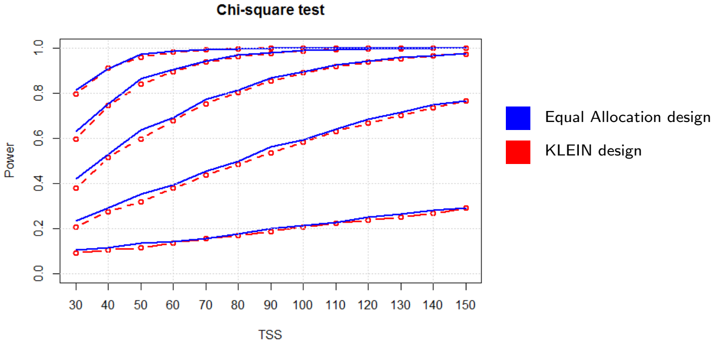
Simulation Results (KLEIN)

- $w = 1$;
- $p_C = 0.2$ and $p_T = \{0.3, 0.4, 0.5, 0.6, 0.7\}$.



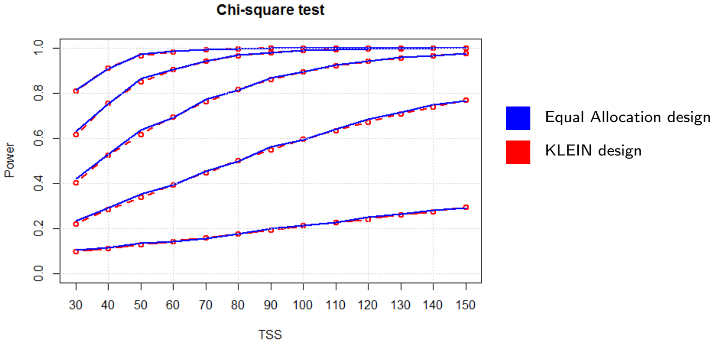
Simulation Results (KLEIN)

- $w = 6$;
- $p_C = 0.2$ and $p_T = \{0.3, 0.4, 0.5, 0.6, 0.7\}$.



Simulation Results (KLEIN)

- $w = 100$;
- $p_C = 0.2$ and $p_T = \{0.3, 0.4, 0.5, 0.6, 0.7\}$.



Sequential Estimation Design (BIN) - the Doubly adaptive Biased Coin Design (DBCD) (Eisele [1994])

- **minimize** the expected **number of failures**
  - fix the **allocation ratio**, ρ

$$\rho = \frac{\sqrt{p_T}}{\sqrt{p_T} + \sqrt{p_C}};$$

- estimate ρ after each patient to determine the **allocation probability** for the patient $n+1$ using **DBCD**.

Sequential Estimation Design (BIN) - the Doubly adaptive Biased Coin Design (DBCD) (Eisele [1994])

$$g_\alpha(x,y) = \begin{cases} 1, & \text{if } x = 0 \\ 0, & \text{if } x = 1 \\ \frac{y(y/x)^\alpha}{y(y/x)^\alpha + (1-y)((1-y)/(1-x))^\alpha}, & \text{if } x \in (0,1) \end{cases}$$

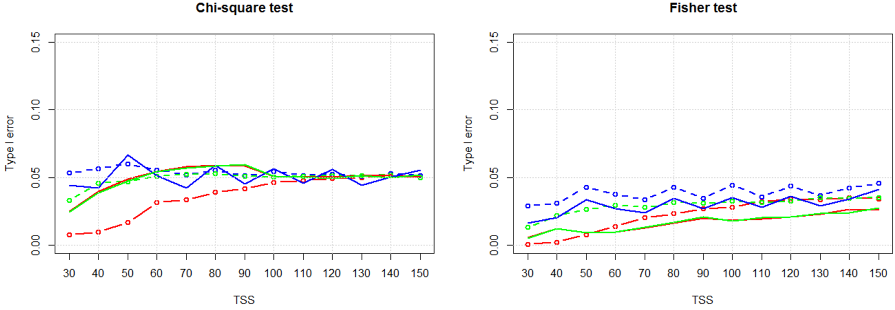
$$P(T_{n+1} = T | previous\ Responses,\ Allocations) = g_\alpha\left(\frac{N_{n,T}}{n}, \hat{\rho}_n\right)$$

with $\alpha \geq 0$ ($\alpha = 2$ Hu and Rosenberger [2003]) and $N_{n,i}$ is the **number of patients** assigned to treatment i, $i = T, C$, up to the patient n.

- **Warning**: at least one success in every group needs to be observed before starting RA allocation!
  - $N_{6,T} = 3$ and $\hat{\rho}_6 = 0$;
  - $P(T_7 = T | previous\ Responses,\ Allocations) = 0 \Rightarrow T_7 = C$;
  - cycle.

Simulation Results (DBCD BIN)

- $\alpha = 2$ (Hu and Rosenberger [2003]);
- **burn-in period** $n_B = 12$.



$p_T = p_C = 0.1$

$p_T = p_C = 0.5$

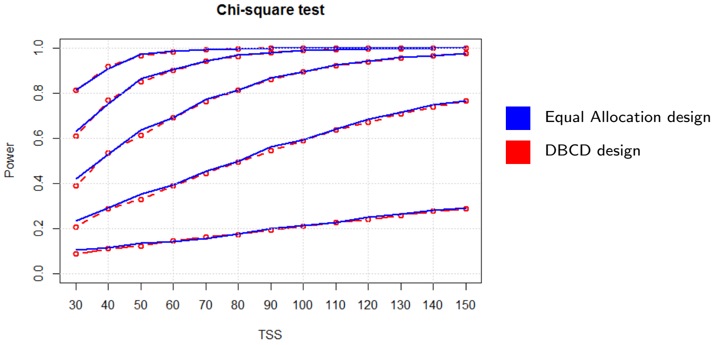
$p_T = p_C = 0.9$

0 - - 0 DBCD design

——— Equal Allocation design

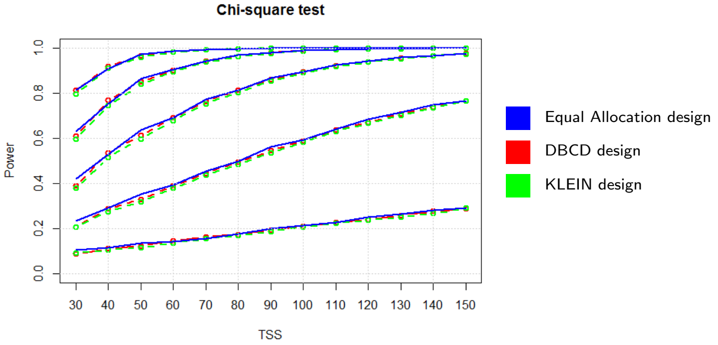
Simulation Results (DBCD BIN)

- $\alpha = 2$ (Hu and Rosenberger [2003]);
- $p_C = 0.2$ and $p_T = \{0.3, 0.4, 0.5, 0.6, 0.7\}$.



Simulation Results (DBCD vs. KLEIN BIN)

- $\alpha = 2$ (Hu and Rosenberger [2003]);
- $p_C = 0.2$ and $p_T = \{0.3, 0.4, 0.5, 0.6, 0.7\}$.



Sequential Estimation Design (CONT) - the Doubly adaptive Biased Coin Design (DBCD) (Eisele [1994])

- fix the **allocation ratio**, ρ (Zhang and Rosenberger [2006])

$$\rho = \begin{cases} \frac{\sigma_T\sqrt{\mu_C}}{\sigma_T\sqrt{\mu_C}+\sigma_C\sqrt{\mu_T}}, & \text{if } s = 1 \\ \frac{1}{2}, & \text{otherwise} \end{cases}$$

where

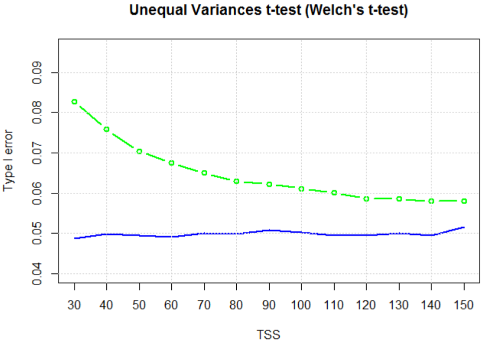
$$s = \begin{cases} 1, & \text{if } (\mu_T < \mu_C \cap r > 1) \vee (\mu_T > \mu_C \cap r < 1) \\ 0, & \text{otherwise} \end{cases}$$

and $r = \sigma_T\sqrt{\mu_C}/\sigma_C\sqrt{\mu_T}$;

- estimate ρ after each patient to determine the **allocation probability** for the patient $n+1$ using **DBCD**.

Simulation Results (DBCD CONT)

- $\alpha = 2$ (Hu and Rosenberger [2003]);
- **burn-in period** $n_B = \{4\}$.



Equal Allocation design

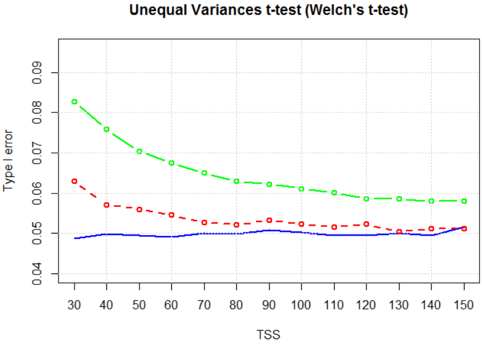
0 - - 0 DBCD design

$\mu_T = \mu_C = 13$

$\sigma_T = \sigma_C = 2.5$

Simulation Results (DBCD CONT)

- $\alpha = 2$ (Hu and Rosenberger [2003]);
- **burn-in period** $n_B = \{4, 6\}$.



Equal Allocation design

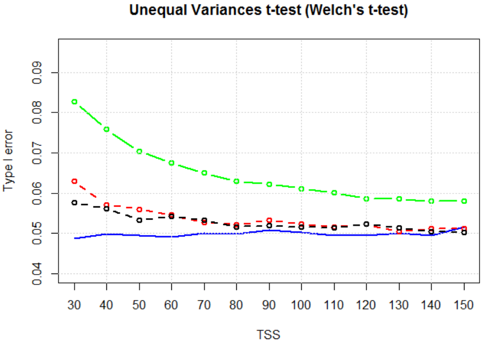
0 - - 0 DBCD design

$\mu_T = \mu_C = 13$

$\sigma_T = \sigma_C = 2.5$

Simulation Results (DBCD CONT)

- $\alpha = 2$ (Hu and Rosenberger [2003]);
- **burn-in period** $n_B = \{4, 6, 8\}$.



Equal Allocation design

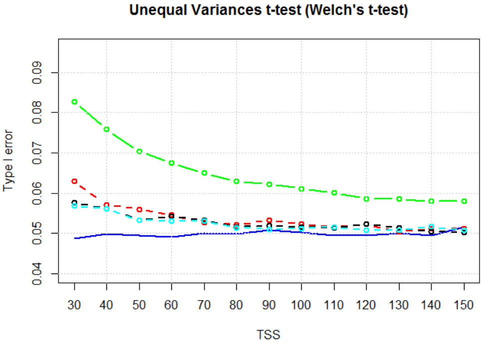
0 - - 0 DBCD design

$\mu_T = \mu_C = 13$

$\sigma_T = \sigma_C = 2.5$

Simulation Results (DBCD CONT)

- $\alpha = 2$ (Hu and Rosenberger [2003]);
- **burn-in period** $n_B = \{4, 6, 8, 10\}$.



Equal Allocation design

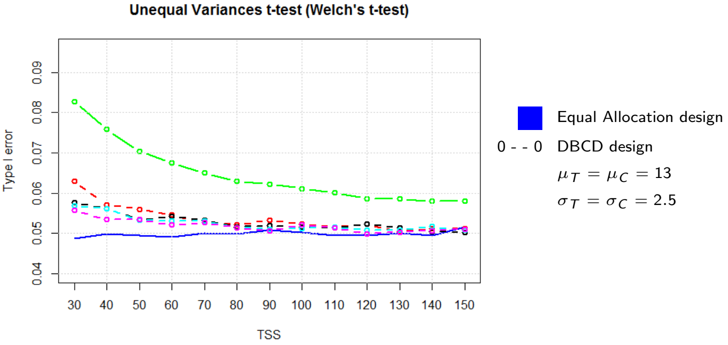
0 - - 0 DBCD design

$\mu_T = \mu_C = 13$

$\sigma_T = \sigma_C = 2.5$

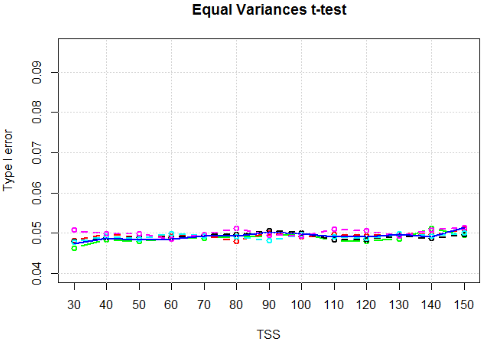
Simulation Results (DBCD CONT)

- $\alpha = 2$ (Hu and Rosenberger [2003]);
- **burn-in period** $n_B = \{4, 6, 8, 10, 12\}$.



Simulation Results (DBCD CONT)

- $\alpha = 2$ (Hu and Rosenberger [2003]);
- **burn-in period** $n_B = \{4, 6, 8, 10, 12\}$.



Equal Allocation design

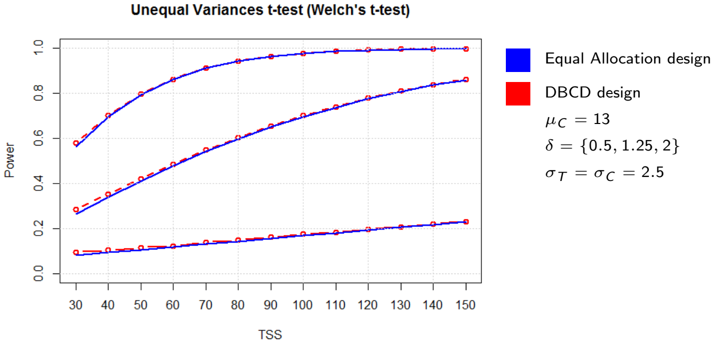
0 - - 0 DBCD design

$\mu_T = \mu_C = 13$

$\sigma_T = \sigma_C = 2.5$

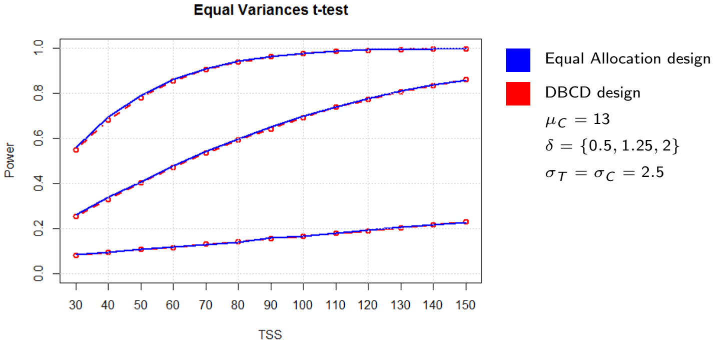
Simulation Results (DBCD CONT)

- $\alpha = 2$ (Hu and Rosenberger [2003]);
- **burn-in period** $n_B = \{12\}$.



Simulation Results (DBCD CONT)

- $\alpha = 2$ (Hu and Rosenberger [2003]);
- **burn-in period** $n_B = \{12\}$.



$\mu_C = 13$

$\delta = \{0.5, 1.25, 2\}$

$\sigma_T = \sigma_C = 2.5$

Outline

Statistical Model

- Consider **three treatment groups** - treatment (T), active control (C) and placebo (P).
- Y_n is a **response** of patient n (binary).
- Consider the **hypotheses**

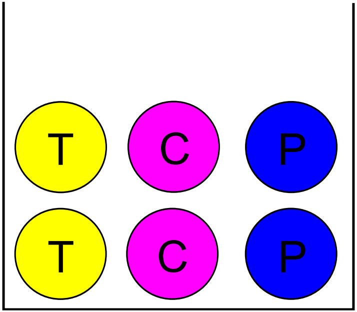
$$H_{0,TP}: \ \theta_T \leq \theta_P \ \text{ vs. } \ H_{1,TP}: \ \theta_T > \theta_P$$

$$H_{0,TC}: \ \theta_T \leq \theta_C - \delta \ \text{ vs. } \ H_{1,TC}: \ \theta_T > \theta_C - \delta$$

  where δ is **non-inferiority margin**.
- α-**adjustment**, e.g., hierarchical order.
- Statistical **test procedure** is defined according to Farrington and Manning [1990].

The Generalized Pólya's Urn Design (GPUD) (Wei [1979])

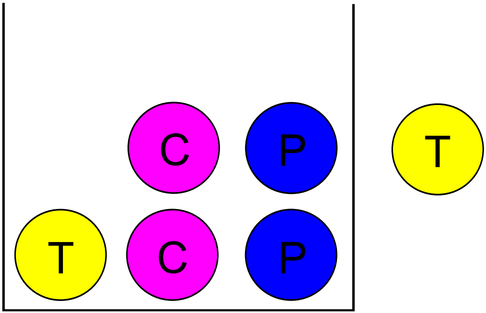
- an **urn** with $3w$ balls of type 'T', type 'C' and type 'P';



$$P(T_{n+1} = i | previous\ Responses,\ Allocations) = \frac{w + \beta S_{n,i} + \alpha \sum_{j \neq i} F_{n,j}}{3w + 2\alpha n}$$

The Generalized Pólya's Urn Design (GPUD) (Wei [1979])

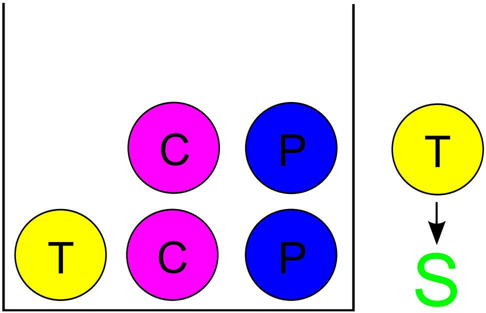
- an **urn** with $3w$ balls of type 'T', type 'C' and type 'P';



$$P(T_{n+1} = i | previous\ Responses,\ Allocations) = \frac{w + \beta S_{n,i} + \alpha \sum_{j \neq i} F_{n,j}}{3w + 2\alpha n}$$

The Generalized Pólya's Urn Design (GPUD) (Wei [1979])

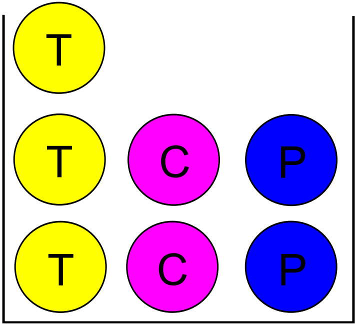
- an **urn** with $3w$ balls of type 'T', type 'C' and type 'P';



$$P(T_{n+1} = i | previous\ Responses,\ Allocations) = \frac{w + \beta S_{n,i} + \alpha \sum_{j \neq i} F_{n,j}}{3w + 2\alpha n}$$

The Generalized Pólya's Urn Design (GPUD) (Wei [1979])

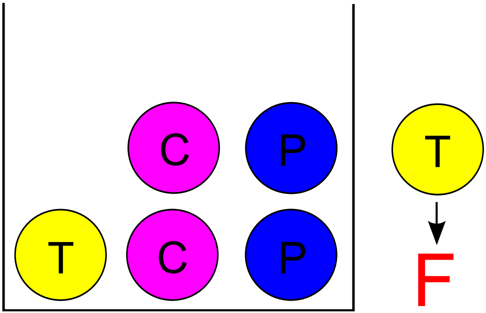
- an **urn** with $3w$ balls of type 'T', type 'C' and type 'P';



$$P(T_{n+1} = i | previous\ Responses,\ Allocations) = \frac{w + \beta S_{n,i} + \alpha \sum_{j \neq i} F_{n,j}}{3w + 2\alpha n}$$

The Generalized Pólya's Urn Design (GPUD) (Wei [1979])

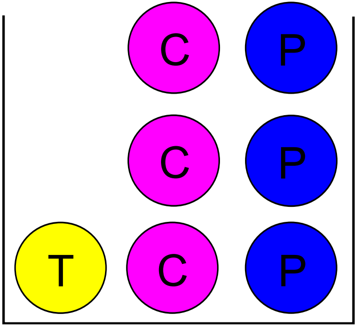
- an **urn** with $3w$ balls of type 'T', type 'C' and type 'P';



$$P(T_{n+1} = i | previous\ Responses,\ Allocations) = \frac{w + \beta S_{n,i} + \alpha \sum_{j \neq i} F_{n,j}}{3w + 2\alpha n}$$

The Generalized Pólya's Urn Design (GPUD) (Wei [1979])

- an **urn** with $3w$ balls of type 'T', type 'C' and type 'P';



$$P(T_{n+1} = i | previous\ Responses,\ Allocations) = \frac{w + \beta S_{n,i} + \alpha \sum_{j \neq i} F_{n,j}}{3w + 2\alpha n}$$

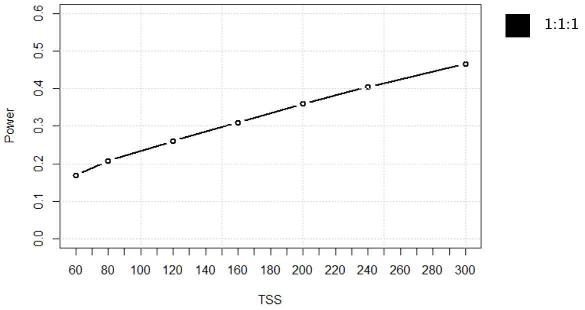
Sequential Estimation Design - the Doubly adaptive Biased Coin Design (DBCD) (Hu and Zhang [2004])

- fix the **allocation ratio**, ρ $(i = T, C, P)$;
- estimate ρ_j $(j = T, C, P)$ after each patient to determine the **allocation probability** for the patient $n+1$ using

$$P(T_{n+1} = j | previous\ Responses,\ Allocations) = \frac{\hat{\rho}_{n,j}(\frac{\hat{\rho}_{n,j}}{N_{n,j}/n})^{\alpha}}{\sum_{i=1}^{3} \hat{\rho}_{n,i}(\frac{\hat{\rho}_{n,i}}{N_{n,i}/n})^{\alpha}}.$$

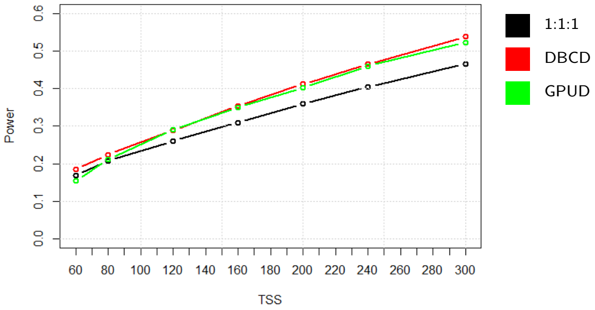
Simulation Results (Three-Arm Trials BIN)

- $\delta = 0.1$;
- $p_T = p_C = 0.7$ and $p_P = 0.1$;
- reject H_{TP} and H_{TC}.



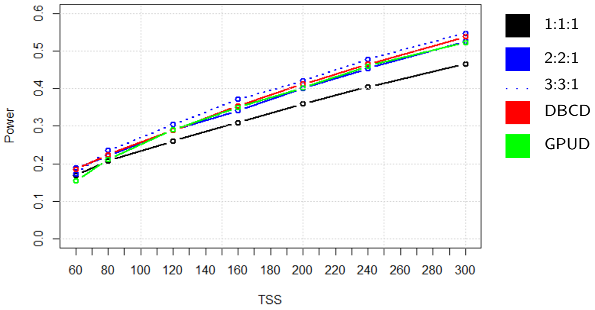
Simulation Results (Three-Arm Trials BIN)

- $\delta = 0.1$;
- $p_T = p_C = 0.7$ and $p_P = 0.1$;
- reject H_{TP} and H_{TC}.



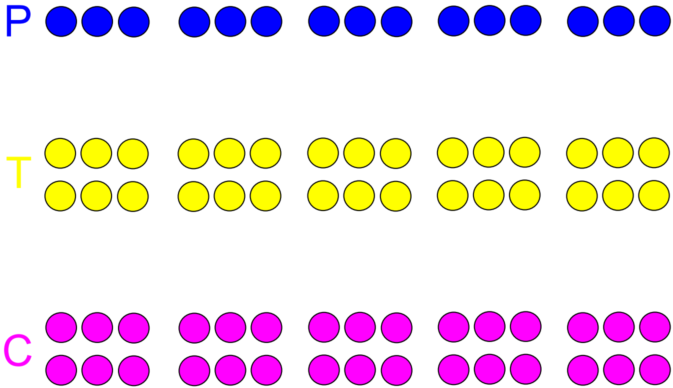
Simulation Results (Three-Arm Trials BIN)

- $\delta = 0.1$;
- $p_T = p_C = 0.7$ and $p_P = 0.1$;
- reject H_{TP} and H_{TC}.

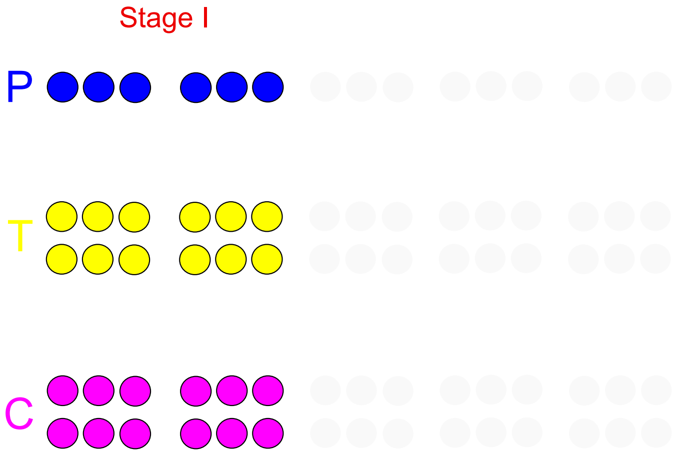


Outline

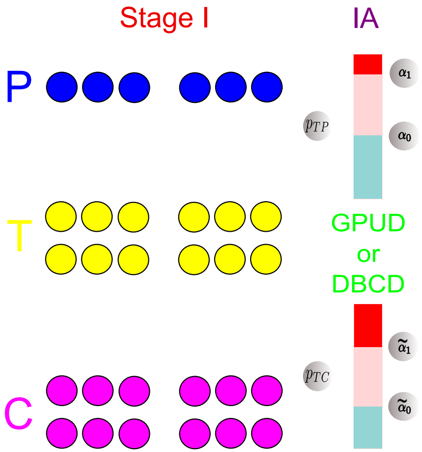
Adaptive Design based on Adaptive Combination Test



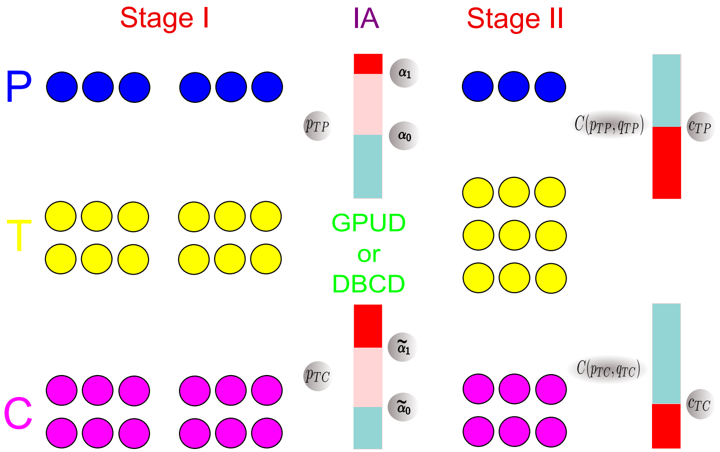
Adaptive Design based on Adaptive Combination Test



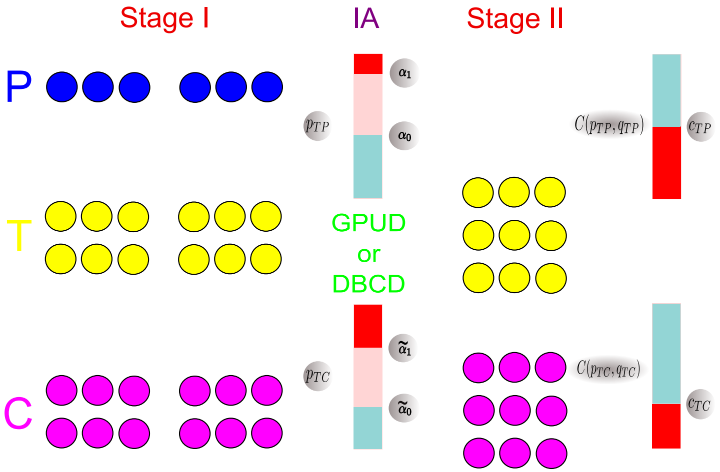
Adaptive Design based on Adaptive Combination Test



Adaptive Design based on Adaptive Combination Test

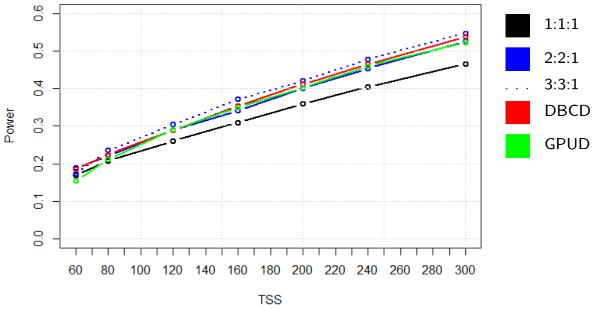


Adaptive Design based on Adaptive Combination Test



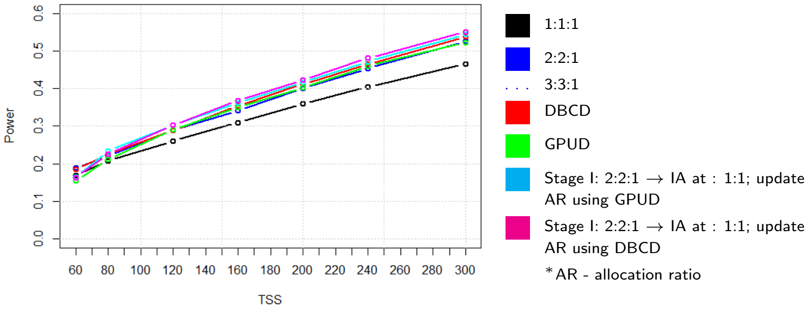
Simulation Results (Three-Arm Trials BIN)

- $\delta = 0.1$;
- $p_T = p_C = 0.7$ and $p_P = 0.1$;
- reject H_{TP} and H_{TC}.



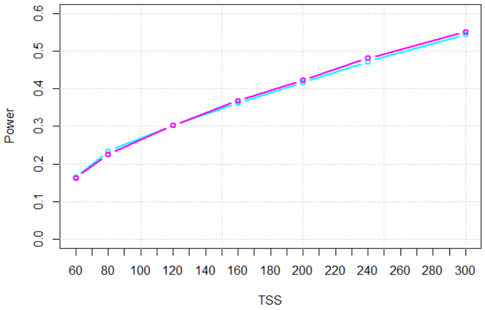
Simulation Results (Three-Arm Trials BIN)

- $\delta = 0.1$;
- $p_T = p_C = 0.7$ and $p_P = 0.1$;
- reject H_{TP} and H_{TC}.



Simulation Results (Three-Arm Trials BIN)

- Changing **timing of interim analysis**?
- $\delta = 0.1$;
- $p_T = p_C = 0.7$ and $p_P = 0.1$;
- reject H_{TP} and H_{TC}.



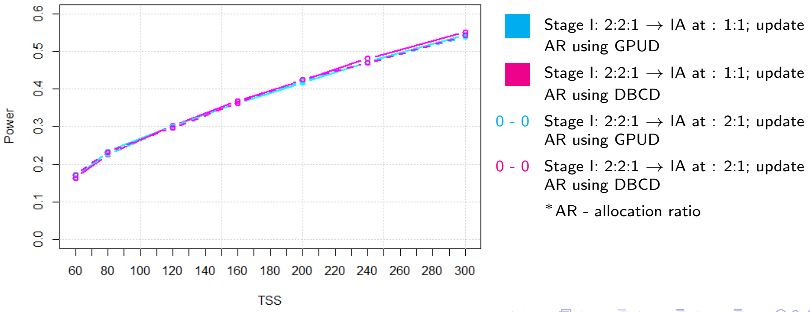
Stage I: 2:2:1 $\rightarrow$ IA at : 1:1; update AR using GPUD

Stage I: 2:2:1 $\rightarrow$ IA at : 1:1; update AR using DBCD

*AR - allocation ratio

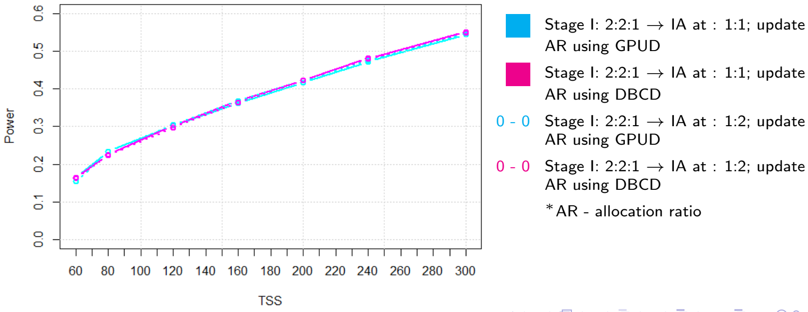
Simulation Results (Three-Arm Trials BIN)

- Changing **timing of interim analysis**?
- $\delta = 0.1$;
- $p_T = p_C = 0.7$ and $p_P = 0.1$;
- reject H_{TP} and H_{TC}.



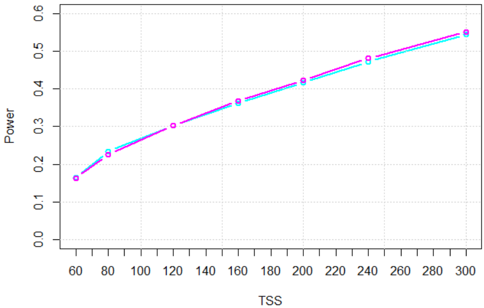
Simulation Results (Three-Arm Trials BIN)

- Changing **timing of interim analysis**?
- $\delta = 0.1$;
- $p_T = p_C = 0.7$ and $p_P = 0.1$;
- reject H_{TP} and H_{TC}.



Simulation Results (Three-Arm Trials BIN)

- **Early stopping** for efficacy? (OBF, Pocock)
- $\delta = 0.1$;
- $p_T = p_C = 0.7$ and $p_P = 0.1$;
- reject H_{TP} and H_{TC}.



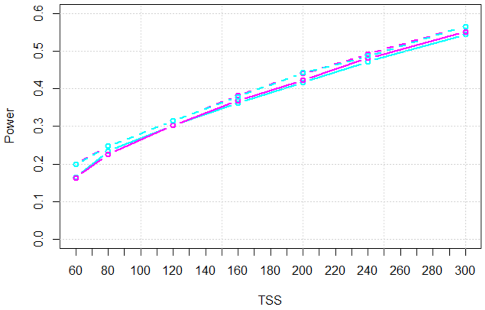
Stage I: 2:2:1 $\rightarrow$ IA at : 1:1; update AR using GPUD (NO EARLY STOP)

Stage I: 2:2:1 $\rightarrow$ IA at : 1:1; update AR using DBCD (NO EARLY STOP)

*AR - allocation ratio

Simulation Results (Three-Arm Trials BIN)

- **Early stopping** for efficacy? (OBF, Pocock)
- $\delta = 0.1$;
- $p_T = p_C = 0.7$ and $p_P = 0.1$;
- reject H_{TP} and H_{TC}.



Stage I: 2:2:1 $\rightarrow$ IA at : 1:1; update AR using GPUD (NO EARLY STOP)

Stage I: 2:2:1 $\rightarrow$ IA at : 1:1; update AR using DBCD (NO EARLY STOP)

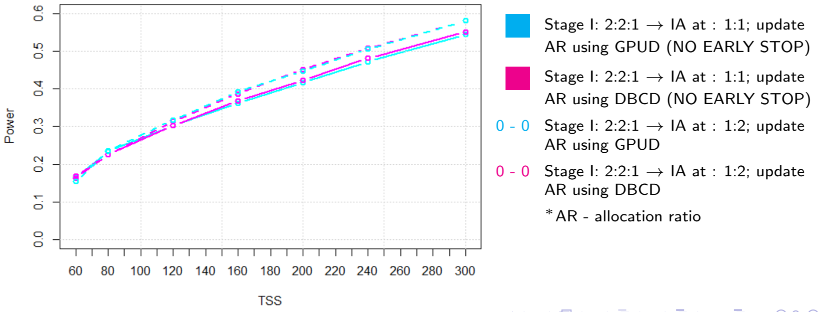
0 - 0 Stage I: 2:2:1 $\rightarrow$ IA at : 2:1; update AR using GPUD

0 - 0 Stage I: 2:2:1 $\rightarrow$ IA at : 2:1; update AR using DBCD

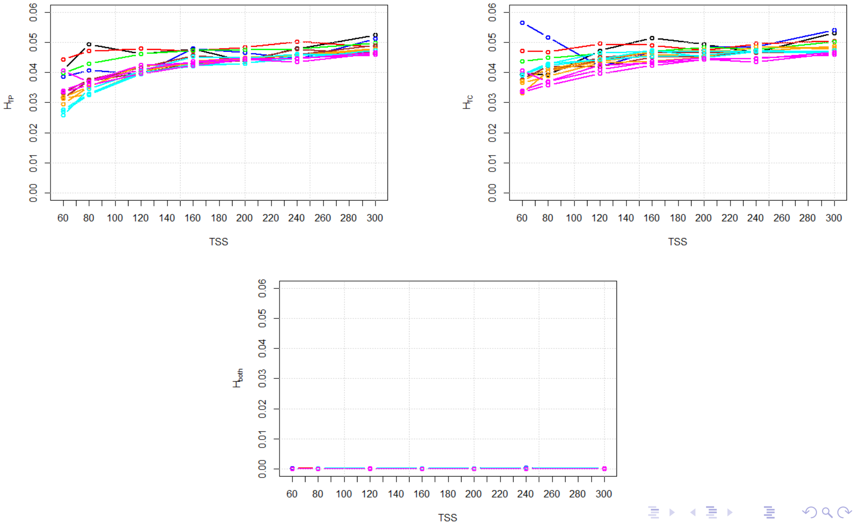
*AR - allocation ratio

Simulation Results (Three-Arm Trials BIN)

- **Early stopping** for efficacy? (OBF, Pocock)
- $\delta = 0.1$;
- $p_T = p_C = 0.7$ and $p_P = 0.1$;
- reject H_{TP} and H_{TC}.



Simulation Results (Three-Arm Trials BIN)



Outline

Conclusions

- minor changes in parameters may have a huge impact on performance (power, type I error, etc.);
- RA designs may not control the type I error rate;
- no "formal" proof of type I error control;
- extensive simulations are needed, but the question is, if simulations are sufficient to prove type I error control (Posch et al. [2011], Gutjahr et al. [2011]);
- by incorporating response-adaptive procedures into adaptive designs, we preserve type I error rate;
- in small populations, we should keep a number of IA to a minimum.

Future Work

- Incorporate appropriate test procedures, that reflect the design.
- Investigate impact of timing, early stopping, etc.
- What are the main reasons (advantages) to use RA procedures in sequential designs?
- When, if so, does the randomization procedure need to be changed?
- How to compare procedures and what criteria to use?

References I

A. Agresti and B. Caffo.
Simple and effective confidence intervals for proportions and differences of proportions result from adding two successes and two failures.
The American Statistician, 54(4):280–288, 2000.

P. Bauer and K. Kohne.
Evaluation of experiments with adaptive interim analyses.
Biometrics, pages 1029–1041, 1994.

J. R. Eisele.
The doubly adaptive biased coin design for sequential clinical trials.
Journal of Statistical Planning and Inference, 38(2):249–261, 1994.

C. P. Farrington and G. Manning.
Test statistics and sample size formulae for comparative binomial trials with null hypothesis of non-zero risk difference or non-unity relative risk.
Statistics in medicine, 9(12):1447–1454, 1990.

A. Galbete, J. A. Moler, and F. Plo.
A response-driven adaptive design based on the Klein urn.
Methodology and Computing in Applied Probability, 16(3):731–746, 2014.

G. Gutjahr, M. Posch, and W. Brannath.
Familywise error control in multi-armed response-adaptive two-stage designs.
Journal of biopharmaceutical statistics, 21(4):818–830, 2011.

References II

F. Hu and W. F. Rosenberger.
Optimality, variability, power: evaluating response-adaptive randomization procedures for treatment comparisons.
Journal of the American Statistical Association, 98(463):671–678, 2003.

F. Hu and L.-X. Zhang.
Asymptotic properties of doubly adaptive biased coin designs for multitreatment clinical trials.
Annals of Statistics, pages 268–301, 2004.

M. Posch, W. Maurer, and F. Bretz.
Type I error rate control in adaptive designs for confirmatory clinical trials with treatment selection at interim.
Pharmaceutical statistics, 10(2):96–104, 2011.

L. Wei.
The generalized Polya's urn design for sequential medical trials.
The Annals of statistics, pages 291–296, 1979.

L. Zhang and W. F. Rosenberger.
Response-adaptive randomization for clinical trials with continuous outcomes.
Biometrics, 62(2):562–569, 2006.

THANK YOU

BACK UP

- **statistical test** for non-inferiority:

$$Z_{TP} = \frac{\hat{p}_T - \hat{p}_P}{\sqrt{\frac{\hat{p}_T(1-\hat{p}_T)}{n_T} + \frac{\hat{p}_P(1-\hat{p}_P)}{n_P}}}$$

$$Z_{TC} = \frac{\hat{p}_T - \hat{p}_C + \delta}{\sqrt{\frac{\hat{p}_T(1-\hat{p}_T)}{n_T} + \frac{\hat{p}_C(1-\hat{p}_C)}{n_C}}}$$