Selecting an appropriate randomization procedure for a small population group trial on the basis of a linked optimization criterion

David Schindler, Ralf-Dieter Hilgers

Department of Medical Statistics
RWTH Aachen University

24. September 2015

FP7 HEALTH 2013 - 602552

Motivation

- No scientific evaluation of randomization procedures in the presence of several types of bias found in literature.
- Several demands on randomization procedures have been well studied independently of each other, but not simultaneously.
- Urgent need for a score that unifies several issues for measuring the different demands on the randomization process.

⇒ Propose a new framework for the selection of an appropriate randomization procedure based on desirability functions.

FP7 HEALTH 2013 - 602552

Motivation

Examplary application of the new framework on

- Selection bias:
  - Assessed issue: correct guesses.
- Chronological bias
  - Assessd issue: type-I-error, power.
- Balancing behavior
  - Assessed issue: power loss due to differences in group sizes.

$\Rightarrow$ Propose a new framework for the selection of an appropriate randomization procedure based on desirability functions.

FP7 HEALTH 2013 - 602552

Terminology

- Two-armed clinical trial with parallel group design with continuous endpoint and total sample size N.
- Experimental treatment E and control treatment C.
- Let $\boldsymbol{T} = (T_1, \ldots, T_N)' \in \{E, C\}^N$ be a randomization sequence and T_i be the ith element of $\boldsymbol{T}$.
- Let $N_s(i, \boldsymbol{T})$ be the number of patients assigned to $s \in \{E, C\}$ after i allocations.

FP7 HEALTH 2013 - 602552

Selection bias

Assuming a balanced trial it is opportune for the experimenter to guess the ith allocation according to the convergence strategy:

(Blackwell and Hodges Jr., 1957)

$$g_{CS}(i, \boldsymbol{T}) = \begin{cases} E, & \text{if } N_E(i-1, \boldsymbol{T}) < N_C(i-1, \boldsymbol{T}) \\ \text{random guess}, & \text{if } N_E(i-1, \boldsymbol{T}) = N_C(i-1, \boldsymbol{T}) \\ C, & \text{if } N_E(i-1, \boldsymbol{T}) > N_C(i-1, \boldsymbol{T}) \end{cases}.$$

Expected proportion of Correct Guesses (CG) of $\boldsymbol{T}$ is defined as:

$$CG(\boldsymbol{T}) = \frac{\mathbb{E}\left(\sum_{i=1}^{N} \mathbb{1}_{\{T_i = g_{CS}(i, \boldsymbol{T})\}}\right)}{N}$$

FP7 HEALTH 2013 - 602552

Chronological bias

Model for chronological bias: (Tamm and Hilgers, 2014; Rosenkranz, 2011)

$$\boldsymbol{Y} = \begin{pmatrix} 1 & \tilde{T}_1 & 1 \\ 1 & \tilde{T}_2 & 2 \\ \vdots & \vdots & \vdots \\ 1 & \tilde{T}_N & N \end{pmatrix} \begin{pmatrix} \mu \\ \xi \\ \vartheta \end{pmatrix} + \boldsymbol{\epsilon},$$

with $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I}_{N\times N})$ and $\tilde{T}_i := t(T_i) = \begin{cases} 1, & \text{if } T_i = E \\ -1, & \text{if } T_i = C \end{cases}$

- The trial is evaluated with a model including the effects μ and ξ, although the time effect $\vartheta \neq 0$ is present (misspecification).

  $\Rightarrow$ The type-I-error α and the power $(1-\beta)$ when testing $\xi = 0$ using a t-test is biased, due to not adjusting for ϑ.

Balancing behavior

Due to differences in group sizes $N_E(N, \boldsymbol{T}) - N_C(N, \boldsymbol{T})$ arising at the end of a clinical trial a **loss in the power** when conducting Student's t-test occurs.

Example:
Assuming a total sample size of $N = 50$, an effect size of $\Delta = 0.81$, and a type-I-error probability of $\alpha = 0.05$ it follows:

$N_E(N, \boldsymbol{T})$	25	24	23	20	15
$N_C(N, \boldsymbol{T})$	25	26	27	30	35
$1 - \beta_0(\boldsymbol{T})$	0.800	0.799	0.797	0.784	0.728

FP7 HEALTH 2013 - 602552

Right-sided Derringer-Suich desirability function

Definition: (Derringer and Suich, 1980)

$$d_i(\boldsymbol{T}) := d(c_i(\boldsymbol{T})) = \begin{cases} 1, & \text{if } c_i(\boldsymbol{T}) \leq TV_i \\ \frac{USL_i - c_i(\boldsymbol{T})}{USL_i - TV_i}, & \text{if } TV_i < c_i(\boldsymbol{T}) < USL_i \\ 0, & \text{if } c_i(\boldsymbol{T}) \geq USL_i \end{cases}$$

- $c_i(\boldsymbol{T})$: value of the i-th issue for $\boldsymbol{T}$.
- TV_i: Target Value of the i-th issue.
- USL_i: Upper Specification Limit of the i-th issue.

FP7 HEALTH 2013 - 602552

Right-sided Derringer-Suich desirability function

Definition: (Derringer and Suich, 1980)

$$d_i(\boldsymbol{T}) := d(c_i(\boldsymbol{T})) = \begin{cases} 1, & \text{if } c_i(\boldsymbol{T}) \leq TV_i \\ \frac{USL_i - c_i(\boldsymbol{T})}{USL_i - TV_i}, & \text{if } TV_i < c_i(\boldsymbol{T}) < USL_i \\ 0, & \text{if } c_i(\boldsymbol{T}) \geq USL_i \end{cases}$$

- $c_i(\boldsymbol{T})$: value of the i-th issue for $\boldsymbol{T}$.
- TV_i: Target Value of the i-th issue.
- USL_i: Upper Specification Limit of the i-th issue.

$\Rightarrow$ Need a meaningful TV and USL dependent on the practical need.

FP7 HEALTH 2013 - 602552

Right-sided Derringer-Suich desirability function

Definition: (Derringer and Suich, 1980)

$$d_i(\boldsymbol{T}) := d(c_i(\boldsymbol{T})) = \begin{cases} 1, & \text{if } c_i(\boldsymbol{T}) \leq TV_i \\ \frac{USL_i - c_i(\boldsymbol{T})}{USL_i - TV_i}, & \text{if } TV_i < c_i(\boldsymbol{T}) < USL_i \\ 0, & \text{if } c_i(\boldsymbol{T}) \geq USL_i \end{cases}$$

Investigated standard setting:

i	$c_i(\boldsymbol{T})$	TV_i	USL_i
1	$CG(\boldsymbol{T})$	0.50	0.75
2	$\alpha_{TT}(\boldsymbol{T})$	0.05	0.10
3	$\beta_{TT}(\boldsymbol{T})$	0.20	0.40
4	$\beta_0(\boldsymbol{T})$	0.20	0.21

FP7 HEALTH 2013 - 602552

Properties of desirability scores

- Desirability scores are dimensionless and $\in [0, 1]$.
- Desirability scores are summarizeable with the geometric mean:

$$\bar{d}(\boldsymbol{T}) := \prod_{i=1}^{4} d_i(\boldsymbol{T})^{\omega_i} \text{ with } \sum_{i=1}^{4} \omega_i = 1.$$

- $\boldsymbol{T}$ with $\bar{d}(\boldsymbol{T}) = 0$ is called undesired randomization sequence.

FP7 HEALTH 2013 - 602552

Properties of desirability scores

- Desirability scores are dimensionless and $\in [0, 1]$.
- Desirability scores are summarizeable with the geometric mean:

$$\bar{d}(\boldsymbol{T}) := \prod_{i=1}^{4} d_i(\boldsymbol{T})^{\omega_i} \text{ with } \sum_{i=1}^{4} \omega_i = 1.$$

- $\boldsymbol{T}$ with $\bar{d}(\boldsymbol{T}) = 0$ is called undesired randomization sequence.

- Weights should be chosen dependent on the planned trial.
- Heuristical approach: Put one third of the weights on each demand.

$\Rightarrow \omega_1 = 1/3$, $\omega_2 = \omega_3 = 1/6$, and $\omega_4 = 1/3$

Correct guesses of PBR(4) for $N = 4$

- `randomizeR` was used for the evaluation (Schindler and Uschner, 2015).

j	$\boldsymbol{T}_j'$	$P(\boldsymbol{T}_j)$	$CG(\boldsymbol{T}_j)$	$d_1(\boldsymbol{T}_j)$
1	EECC	1/6	0.625	
2	ECEC	1/6	0.750	
3	CEEC	1/6	0.750	
4	ECCE	1/6	0.750	
5	CECE	1/6	0.750	
6	CCEE	1/6	0.625	
	average value:		0.708	

PBR(k) (Permuted Block Randomization with block length k) Within each block half of the patients are assigend to E and C.

FP7 HEALTH 2013 - 602552

Correct guesses of PBR(4) for $N = 4$

- `randomizeR` was used for the evaluation (Schindler and Uschner, 2015).

j	$\boldsymbol{T}_j'$	$P(\boldsymbol{T}_j)$	$CG(\boldsymbol{T}_j)$	$d_1(\boldsymbol{T}_j)$
1	EECC	$1/6$	0.625	0.500
2	ECEC	$1/6$	0.750	0.000
3	CEEC	$1/6$	0.750	0.000
4	ECCE	$1/6$	0.750	0.000
5	CECE	$1/6$	0.750	0.000
6	CCEE	$1/6$	0.625	0.500
	average value:		0.708	0.167

$$d_1(\boldsymbol{T}_1) = d(CG(\boldsymbol{T}_1)) = \frac{USL_1 - CG(\boldsymbol{T}_1)}{USL_1 - TV_1} = \frac{0.75 - 0.625}{0.75 - 0.5} = 0.5$$

FP7 HEALTH 2013 - 602552

Correct guesses of PBR(4) for $N = 4$

- `randomizeR` was used for the evaluation (Schindler and Uschner, 2015).

j	$\boldsymbol{T}'_j$	$P(\boldsymbol{T}_j)$	$CG(\boldsymbol{T}_j)$	$d_1(\boldsymbol{T}_j)$
1	EECC	$1/6$	0.625	0.500
2	ECEC	$1/6$	0.750	0.000
3	CEEC	$1/6$	0.750	0.000
4	ECCE	$1/6$	0.750	0.000
5	CECE	$1/6$	0.750	0.000
6	CCEE	$1/6$	0.625	0.500
	average value:		0.708	0.167

$$\bar{d}_1(\boldsymbol{T}) = 1/6\,(0.5 + 0 + 0 + 0 + 0 + 0.5) = 0.167$$

Assessment of PBR(4) for $N = 4$

j	T'_j	$P(T_j)$	$d_1(T_j)$	$d_2(T_j)$	$d_3(T_j)$	$d_4(T_j)$	$\bar{d}(T_j)$
1	EECC	1/6	0.500	0.804	0.649	1.000	0.712
2	ECEC	1/6	0.000	1.000	0.668	1.000	0.000
3	CEEC	1/6	0.000	1.000	0.776	1.000	0.000
4	ECCE	1/6	0.000	1.000	0.776	1.000	0.000
5	CECE	1/6	0.000	1.000	0.961	1.000	0.000
6	CCEE	1/6	0.500	0.804	1.000	1.000	0.765
	average value:		0.167	0.935	0.805	1.000	0.246

Settings: $\vartheta = 1/4$, $\xi = 2.83$, $\alpha_0 = 0.05$, and $1 - \beta_0 = 0.8$.

$d_1(T) = d(CG(T))$ $\quad d_2(T) = d(\alpha_{TT}(T))$ $\quad d_3(T) = d(1 - \beta_{TT}(T))$ $\quad d_4(T) = d(1 - \beta_0(T))$

$$\bar{d}(T_1) = \sqrt[3]{d_1(T_1)} \cdot \sqrt[6]{d_2(T_1)} \cdot \sqrt[6]{d_3(T_1)} \cdot \sqrt[3]{d_4(T_1)}$$
$$= \sqrt[3]{0.500} \cdot \sqrt[6]{0.804} \cdot \sqrt[6]{0.649} \cdot \sqrt[3]{1.000}$$
$$= 0.712$$

Assessment of PBR(4) for $N = 4$

j	$\boldsymbol{T}'_j$	$P(\boldsymbol{T}_j)$	$d_1(\boldsymbol{T}_j)$	$d_2(\boldsymbol{T}_j)$	$d_3(\boldsymbol{T}_j)$	$d_4(\boldsymbol{T}_j)$	$\bar{d}(\boldsymbol{T}_j)$
1	EECC	$1/6$	0.500	0.804	0.649	1.000	0.712
2	ECEC	$1/6$	0.000	1.000	0.668	1.000	0.000
3	CEEC	$1/6$	0.000	1.000	0.776	1.000	0.000
4	ECCE	$1/6$	0.000	1.000	0.776	1.000	0.000
5	CECE	$1/6$	0.000	1.000	0.961	1.000	0.000
6	CCEE	$1/6$	0.500	0.804	1.000	1.000	0.765
	average value:		0.167	0.935	0.805	1.000	0.246

Settings: $\vartheta = 1/4$, $\xi = 2.83$, $\alpha_0 = 0.05$, and $1 - \beta_0 = 0.8$.
$d_1(\boldsymbol{T}) = d(CG(\boldsymbol{T}))$ $d_2(\boldsymbol{T}) = d(\alpha_{TT}(\boldsymbol{T}))$ $d_3(\boldsymbol{T}) = d(1 - \beta_{TT}(\boldsymbol{T}))$ $d_4(\boldsymbol{T}) = d(1 - \beta_0(\boldsymbol{T}))$

$$\varnothing \bar{d}(\boldsymbol{T}) = 1/6\,(0.712 + 0 + 0 + 0 + 0 + 0.765) = 0.246$$

FP7 HEALTH 2013 - 602552

Assessment of PBR(4) for $N = 4$

j	$\boldsymbol{T}'_j$	$P(\boldsymbol{T}_j)$	$d_1(\boldsymbol{T}_j)$	$d_2(\boldsymbol{T}_j)$	$d_3(\boldsymbol{T}_j)$	$d_4(\boldsymbol{T}_j)$	$\bar{d}(\boldsymbol{T}_j)$
1	EECC	1/6	0.500	0.804	0.649	1.000	0.712
2	ECEC	1/6	0.000	1.000	0.668	1.000	0.000
3	CEEC	1/6	0.000	1.000	0.776	1.000	0.000
4	ECCE	1/6	0.000	1.000	0.776	1.000	0.000
5	CECE	1/6	0.000	1.000	0.961	1.000	0.000
6	CCEE	1/6	0.500	0.804	1.000	1.000	0.765
	average value:		0.167	0.935	0.805	1.000	0.246

Settings: $\vartheta = 1/4$, $\xi = 2.83$, $\alpha_0 = 0.05$, and $1 - \beta_0 = 0.8$.
$d_1(\boldsymbol{T}) = d(CG(\boldsymbol{T}))$ $d_2(\boldsymbol{T}) = d(\alpha_{TT}(\boldsymbol{T}))$ $d_3(\boldsymbol{T}) = d(1 - \beta_{TT}(\boldsymbol{T}))$ $d_4(\boldsymbol{T}) = d(1 - \beta_0(\boldsymbol{T}))$

2/3 of the randomization sequences are undesired.

$\Rightarrow$ Approach for $N = 4$ not usefull.

FP7 HEALTH 2013 - 602552

Investigated randomization procedures

PBR(k) (Permuted Block Randomization with block length k) Within each block half of the patients are assigend to E and C.

RPBR(k) (Randomized Permuted Block Randomization with maximal block length k) PBR with random block lengths $2, 4, .., k$.

CR Complete randomization is accomplished by tossing a fair coin.

BSD(a) (Big Stick Design) CR allow for imbalance within the limit a.

Comparison for $N = 50$

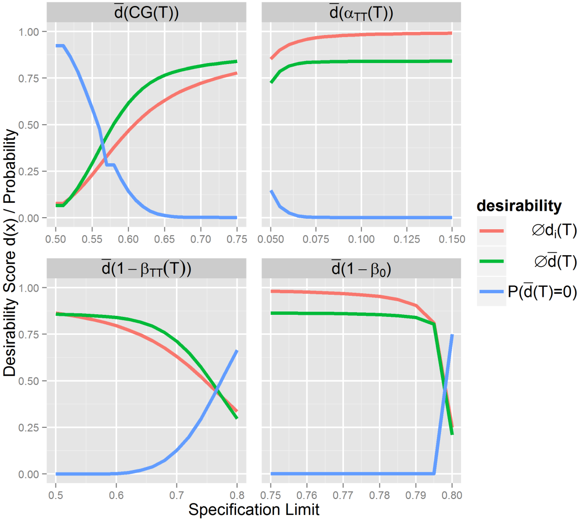
Settings: $\vartheta = 1/50$, $\xi = 0.40$, $\alpha_0 = 0.05$, and $1 - \beta_0 = 0.8$.
Results based on 100.000 simulations.

Design	$\varnothing\bar{d}(\boldsymbol{T})$ (sd)	$P(\bar{d}(\boldsymbol{T}) = 0)$
CR	0.5131 (0.388)	0.3534
RPBR(8)	0.6088 (0.081)	0.0011
PBR(8)	0.6759 (0.07)	0.0001
PBR(50)	0.7797 (0.181)	0.0408
BSD(4)	0.8400 (0.084)	0.0024

- BSD(4) has low probability of generating undesired randomization sequences.
- BSD(4) seems to be the best compromise between handling a time trend, the proportion of correct guesses, and the loss in power.

FP7 HEALTH 2013 - 602552

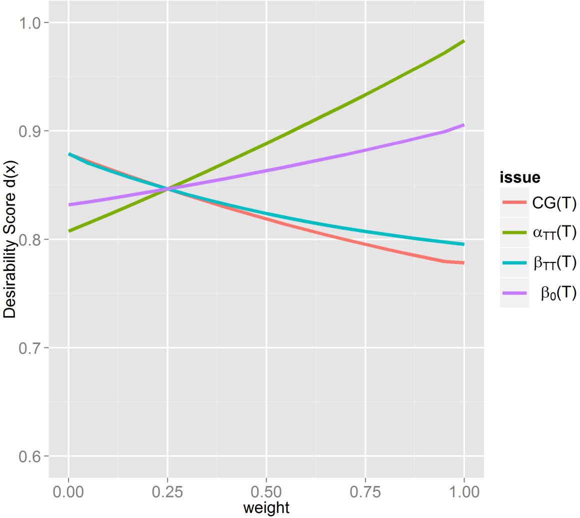
Analysis of the used USLs for BSD(4)



- Change of the desirability scores, when a specification limit convergences against the TV_i.

FP7 HEALTH 2013 - 602552

Analysis of the used weights for BSD(4)



- Change of the weight of an fixed issue. The other weights are splitted equally.

FP7 HEALTH 2013 - 602552

Conclusions

- Presented a framework for the scientific evaluation of randomization procedures dependent on arising demands.
- Evaluation should be part of the statistical trial and analysis plan.
- Other TVs, USLs, and weights for the investigated issues lead to different recommendations.
- Other randomization procedures can be implemented easily.
- Include other issues for measuring (further) demands.

This project has received funding from the European Union's 7th Framework Programme for research, technological development and demonstration under Grant Agreement no 602552.

FP7 HEALTH 2013 - 602552

References

Blackwell, D. and J. L. Hodges Jr. (1957). Design for the control of selection bias. *Annals of Mathematical Statistics 25*, 449–460.

Derringer, G. and R. Suich (1980). Simultaneous optimization of several response variables. *Journal of Quality Technology 12*, 214–219.

ICH E9 (1998). Statistical principles for clinical trials. *Current Step 4 version dated 5 Februrary 1998. Available from:* `http://www.ich.org`.

Rosenkranz, G. K. (2011). The impact of randomization on the analysis of clinical trials. *Statistics in Medicine 30*, 3475–3487.

Schindler, D. and D. Uschner (2015). *randomizeR: Randomization for clinical trials*. R package version 1.0.

Tamm, M. and R.-D. Hilgers (2014). Chronological bias in randomized clinical trials under different types of unobserved time trends. *Methods of Information in Medicine 53*, 501–510.

FP7 HEALTH 2013 - 602552