Randomization Tests in the Presence of Selection Bias

D. Uschner[1] R.-D. Hilgers[1] W.F. Rosenberger[2] N. Heussen[1]

[1]RWTH Aachen University

[2]George Mason University Fairfax VA

22nd August 2016

This research is part of the IDeAl project http://www.ideal.rwth-aachen.de and has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under Grant Agreement no 602552.

FP7 HEALTH 2013 - 602552

Motivation and Outline

- In small population groups, the assumptions of parametric tests are not fulfilled.
- Randomization tests yield a non-parametric alternative, that relies only on the assumption of randomization.
- Problem: Randomization is susceptible to third order selection bias.

Research Question and Tasks

Does randomization based inference yield valid inferences when selection bias is present?

- Develop model for selection bias in randomization tests.
- Analyze the properties of randomization tests when selection bias is present.

FP7 HEALTH 2013 - 602552

Notation

Let the randomization sequence $T = (T_1, \ldots, T_N)$ be a random vector with T_i taking values

$$t_i = \begin{cases} 0 & \text{if patient } i \text{ is allocated to } C \\ 1 & \text{if patient } i \text{ is allocated to } E, \end{cases}$$

and let the imbalance after i patients be denoted by

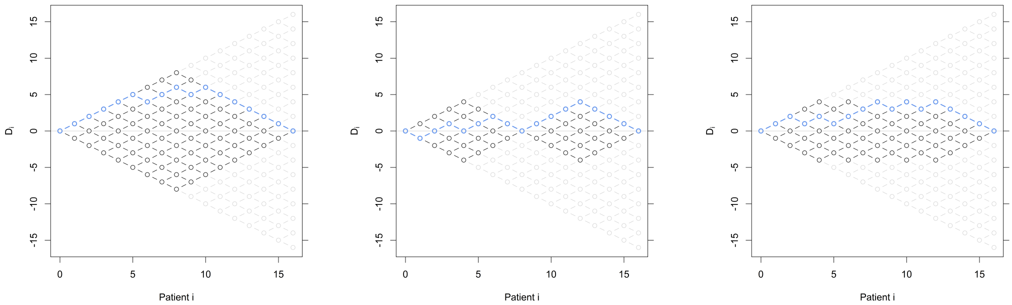
$$D_i = \sum_{j=1}^{i} (2 \cdot T_i - 1).$$

The set of all randomization sequences of a randomization procedure $\mathcal{M}$ is denoted by

$$\Omega_{\mathcal{M}} \subseteq \Omega := \{0, 1\}^N.$$

FP7 HEALTH 2013 - 602552

Randomization procedures



(a) Ω_{RAR} (b) $\Omega_{PBR(8)}$ (c) $\Omega_{MP(4)}$

Rosenberger and Lachin, 2016

FP7 HEALTH 2013 - 602552

Randomization based inference

Null hypothesis

The response y_i of each patient i is independent of the treatment he received:

$$H_0 : y_{i|E} = y_{i|C}$$

- Difference in means test statistic is a measure for the extremeness of the result:

$$S(t) = \frac{N}{2} \sum_{i=1}^{N} (t_i \cdot y_i - (1 - t_i) \cdot y_i)\,.$$

- H_0 is rejected if the probability to observe a more extreme value of the test statistic is lower than 5%:

$$p = \sum_{t \in \Omega} \mathbb{1}(|S(t)| \geq |S(t_{obs})|) \cdot \mathbb{P}(T = t) < 0.05.$$

FP7 HEALTH 2013 - 602552

Randomization based inference

Instructive example

Let $N = 8$ patients be allocated with RAR. Assume $y_{obs} = (1, 6, 7, 2, 8, 4, 3, 5)$ and $t_{obs} = (1, 1, 0, 1, 0, 0, 1, 0)$. The observed test statistic is

$$S(t_{obs}) = -3,$$

This yields a p-value of

$$p = \frac{|\{t : |S(t)| \geq |-3|\}|}{70} = \frac{8}{70} > 0.05.$$

$\Rightarrow$ the null hypothesis cannot be rejected.

	t	$P(T = t)$	$S(t)$
1	1 1 1 1 0 0 0 0	$\frac{1}{70}$	-1
2	1 1 1 0 1 0 0 0	$\frac{1}{70}$	2
3	1 1 0 1 1 0 0 0	$\frac{1}{70}$	-0.5
4	1 0 1 1 1 0 0 0	$\frac{1}{70}$	0
5	0 1 1 1 1 0 0 0	$\frac{1}{70}$	2.5
$\vdots$	$\vdots$	$\vdots$	$\vdots$
17	1 1 0 1 0 0 1 0	$\frac{1}{70}$	-3
$\vdots$	$\vdots$	$\vdots$	$\vdots$
70	0 0 0 0 1 1 1 1	$\frac{1}{70}$	1

Table: Distribution of the test statistic S

Instructive example

Let $N = 8$ patients be allocated with RAR.
Assume $y_{obs} = (1, 6, 7, 2, 8, 4, 3, 5)$ and $t_{obs} = (1, 1, 0, 1, 0, 0, 1, 0)$. The observed test statistic is

$$S(t_{obs}) = -3,$$

This yields a p-value of

$$p = \frac{|\{t : |S(t)| \geq |-3|\}|}{70} = \frac{8}{70} > 0.05.$$

$\Rightarrow$ the null hypothesis cannot be rejected.

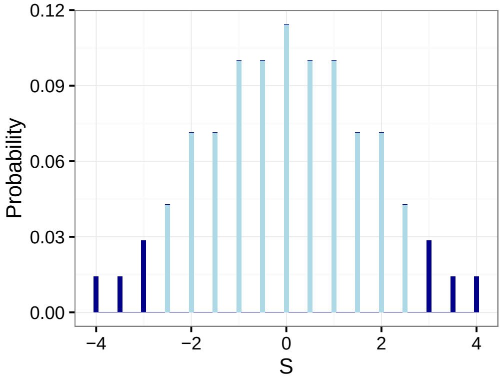


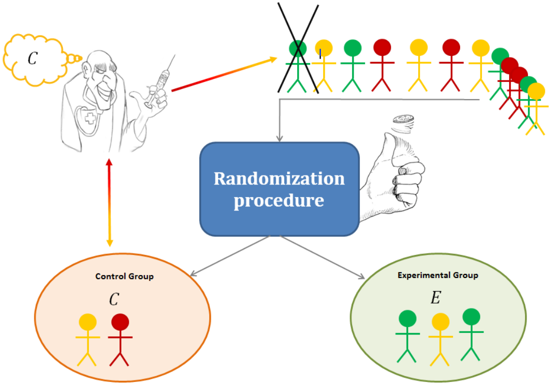
Figure: Distribution of the test statistic S

FP7 HEALTH 2013 - 602552

Third order selection bias (Berger, 2005)

Setting

- Trial is randomized.
- Randomization is restricted.
- Patient population is not homogeneous.
- Past allocations can be unmasked.
- Investigator favors one treatment.



FP7 HEALTH 2013 - 602552

Result: Model for the influence of third order selection bias

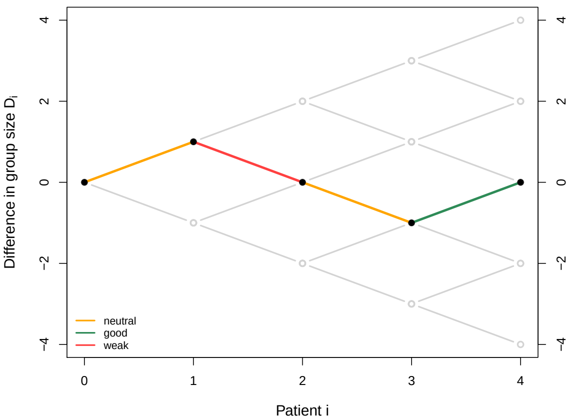
Biasing Policy

Following the convergence strategy proposed by Blackwell and Hodges (1957), we introduce the following model for the biased responses:

$$\tilde{y}_i = \begin{cases} y_i + \eta & D_{i-1}(t_{obs}) < 0 \\ y_i & D_{i-1}(t_{obs}) = 0 \\ y_i - \eta & D_{i-1}(t_{obs}) > 0 \end{cases}$$

$$\Leftrightarrow \tilde{y}_i = y_i - \eta \cdot \mathrm{sign}(D_{i-1}(t_{obs}))$$

with η the strength of the selection bias.



FP7 HEALTH 2013 - 602552

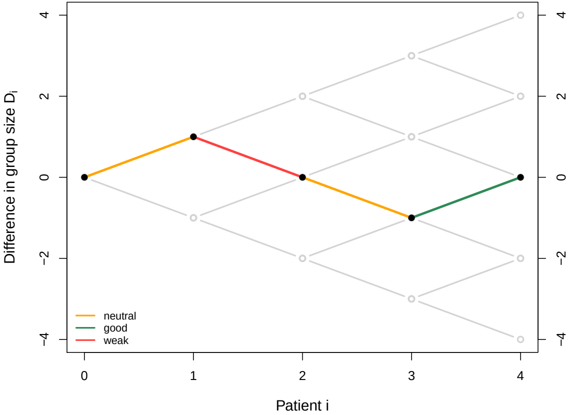
Result: Model for the influence of third order selection bias

Biasing Policy

If $N = 4$ and $t_{obs} = (1, 0, 0, 1)$ and third order selection bias is present with selection effect $\eta > 0$, we will observe the biased responses

$$\begin{aligned}\tilde{y}_1 &= y_1\\ \tilde{y}_2 &= y_2 - \eta\\ \tilde{y}_3 &= y_3\\ \tilde{y}_4 &= y_4 + \eta\end{aligned}$$

instead of the unbiased responses $y_1, \ldots y_N$.



FP7 HEALTH 2013 - 602552

Simulation study to investigate the effect of selection bias

Aim: Estimate type-I-error probability for increasing selection effect and different randomization procedures.

Simulation settings

- Sample size $N = 16$
- Randomization procedure $\mathcal{M} \in \{\text{MP}(4), \text{PBR}(8), \text{RAR}\}$
- In each setting, conducted $r = 60,000$ randomization tests.
- For each randomization test, generated $t_{obs} \in \Omega_{\mathcal{M}}$ and $\tilde{y}_{obs} = (\tilde{y}_1, \ldots, \tilde{y}_N)$, with $\tilde{y}_i = y_i - \eta \cdot \text{sign}(D_{i-1}(t_{tobs}))$ and y_i realization of $Y_i \sim \mathcal{N}(0, 1)$.
- The type-I-error rate is the proportion of p-values lower than 5%.

FP7 HEALTH 2013 - 602552

Result: Type-I-error rate elevated by selection bias

- Type I error rate elevated with increasing selection effect η.
- Effect differs with the randomization procedure.
- Permuted Block randomization is most susceptible to selection bias.

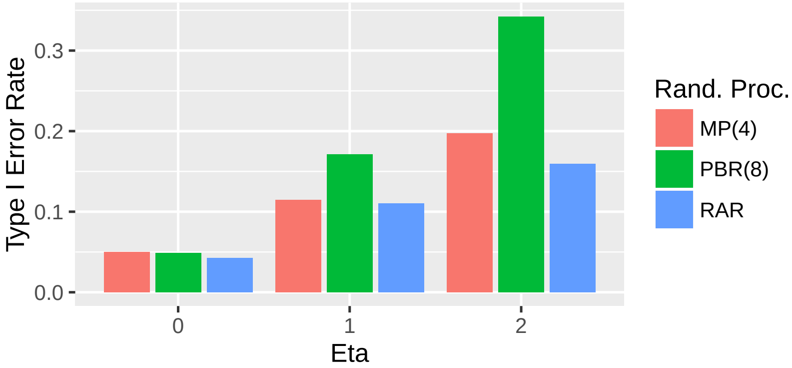


Figure: Type-I-error rate of the randomization test

FP7 HEALTH 2013 - 602552

Conditional testing approach to achieve adjusted test

Lemma

If $D_N(t) = 0$ for all $t \in \Omega_{\mathcal{M}}$, the bias of the responses leads to a shift of the test statistic

$$S(t, \tilde{y}) = S(t, y) - \frac{2\eta}{N} \cdot shift(t),$$

where the shift can be expressed as

$$shift(t) = \sum_{i=1}^{N} (2 \cdot t_i - 1) \cdot \mathrm{sign}(D_{i-1}(t_{obs}))$$

Particularly, $shift(t_{obs}) = \#rto(t_{obs}) := |\{i \in \{1, \ldots, N-1\} : D_{i-1}(t_{obs}) = 0\}|$.

FP7 HEALTH 2013 - 602552

Adjusted randomization test

Cox (1982): Condition randomization on any aspects of the treatment arrangements which there is reason to think relevant.

Theorem

Conditioning on $shift(t_{obs})$ yields an unbiased test with p-value

$$p_{adj} = \sum_{t \in \Omega} \mathbb{1}(|S(t)| \geq |S(t_{obs})|) \cdot \mathbb{P}(T = t | shift(t) = shift(t_{obs}))$$

where the conditional probability can be computed by re-weighting:

$$\mathbb{P}(T = t | shift(t) = shift(t_{obs})) = \frac{\mathbb{P}(T = t) \cdot \mathbb{1}(shift(t) = shift(t_{obs})) \cdot |\Omega_{\mathcal{M}}|}{|\{t \in \Omega_{\mathcal{M}} : shift(t) = shift(t_{obs})\}|}.$$

FP7 HEALTH 2013 - 602552

Conditional testing approach to achieve adjusted test

Instructive Example

Condition on $shift(t) = shift(t_{obs})$:

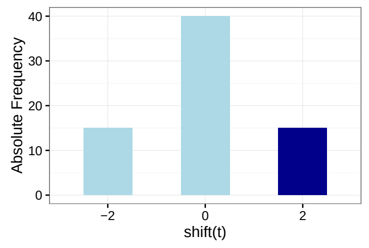


Figure: Frequency of shift

	t	$P(T=t)$	$S(t)$	$shift(t)$
1	1 1 1 1 0 0 0 0	1/70	-1	0
2	1 1 1 0 1 0 0 0	1/70	2	0
⋮	⋮	⋮	⋮	⋮
15	0 0 1 1 1 1 0 0	1/70	1.5	-2
16	1 1 1 0 0 0 1 0	1/70	-0.5	2
17	1 1 0 1 0 0 1 0	1/70	-3	2
18	1 0 1 1 0 0 1 0	1/70	-2.5	2
19	0 1 1 1 0 0 1 0	1/70	0	0
⋮	⋮	⋮	⋮	⋮
70	0 0 0 0 1 1 1 1	1/70	1	0

Table: Randomization distribution

FP7 HEALTH 2013 - 602552

Conditional testing approach to achieve adjusted test

Instructive Example

Condition on $shift(t) = shift(t_{obs})$

$\Rightarrow p_{adj} = 0.2$!

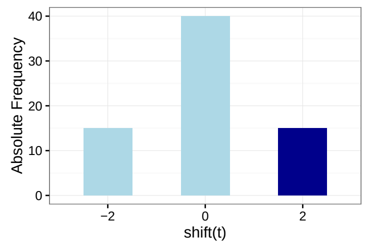


Figure: Frequency of shift

	t	$P(T=t)$	$S(t)$	$shift(t)$
1	1 1 1 1 0 0 0 0	0	-1	0
2	1 1 1 0 1 0 0 0	0	2	0
⋮	⋮	⋮	⋮	⋮
15	0 0 1 1 1 1 0 0	0	1.5	-2
16	1 1 1 0 0 0 1 0	1/15	-0.5	2
17	1 1 0 1 0 0 1 0	1/15	-3	2
18	1 0 1 1 0 0 1 0	1/15	-2.5	2
19	0 1 1 1 0 0 1 0	0	0	0
⋮	⋮	⋮	⋮	⋮
70	0 0 0 0 1 1 1 1	0	1	0

Table: Randomization distribution

FP7 HEALTH 2013 - 602552

Result: Adjusted test maintains type-I-error rate

- The adjusted test maintains the type-I-error rate for increasing η.
- The type-I-error rate is equal to the case $\eta = 0$ of the unadjusted test.
- Result does not depend on the randomization procedure.

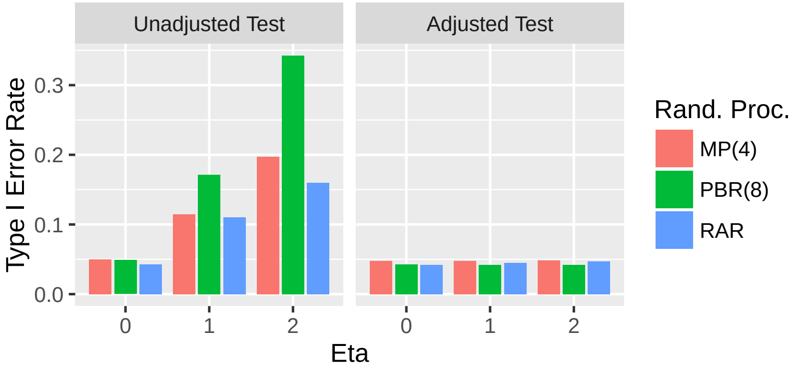


Figure: Type-I-error rate of the randomization test

FP7 HEALTH 2013 - 602552

Conclusions & Perspectives

- Randomization based inference generally does not control selection bias.
- Adjusted randomization test maintains type-I-error rate.
  - $\Rightarrow$ Conduct adjusted randomization test when a trial is suspected to be influenced by selection bias.
- Randomization based inference provides a valid alternative to parametric tests
- Small population group trials may benefit from the non-parametric analysis
- Possible extensions: Other test statistics, other types of bias, unbiased estimators.

FP7 HEALTH 2013 - 602552