Assessment of randomization procedures based on single sequences under selection and chronological bias

Ralf-Dieter Hilgers, Simon Langer

Department of Medical Statistics, RWTH-Aachen University

Friday 10th July 2015

FP7 HEALTH 2013 - 602552

Table of contents

FP7 HEALTH 2013 - 602552

Introduction

- 21 out of 63 Orphan drug legislations involve open label studies (Joppi, 2013)
- *treatment comparisons should involve consideration of the potential contribution of bias to the p-value* (ICH E9)
- no recommendation to give scientific arguments for selection of randomization procedure
- no uniform performance of randomization procedures
- Clinical Scenario Evaluation (Benda 2010)

FP7 HEALTH 2013 - 602552

Introduction - Objective

Propose a tool for assessing the impact of selection bias as well as chronological bias on the type one error probability for a given randomization sequence (procedure) and thus enabling a scientific discussion of the appropriate choice of the randomization procedure

Setting - Model

Two arm parallel group design

$$y_i = \mu_A Z_i + \mu_B(1 - Z_i) + \tau_i + \epsilon_i, \quad 1 \leq i \leq n_A + n_B \tag{1}$$

- allocation

$$Z_i = \begin{cases} 1 & \text{if patient } i \text{ is allocated to group } A \\ 0 & \text{if patient } i \text{ is allocated to group } B \end{cases}$$

- μ_j expected response under treatment $j = A, B$
- τ_i denotes the fixed unobserved "bias" effect acting on the response of patient i
- errors ϵ_i iid $N(0, \sigma^2)$

Setting -Test Statistic

Aim: test the hypotheses $H_0 : \mu_A = \mu_B$ vs. $H_1 : \mu_A \neq \mu_B$

use t-Test (under misspecification)

$$T = \frac{\sqrt{\frac{n_A n_B}{n_A+n_B}}\,(\tilde{y}_A - \tilde{y}_B)}{\frac{1}{n_A+n_B-2}\left(\sum_{i=1}^{n} Z_i(y_i - \tilde{y}_A)^2 + \sum_{i=1}^{n}(1-Z_i)(y_i - \tilde{y}_B)^2\right)}$$

where $\tilde{y}_A = \frac{1}{n_A}\sum_{i=1}^{n} y_i Z_i$; $\tilde{y}_B = \frac{1}{n_B}\sum_{i=1}^{n} y_i(1-Z_i)$; $n = n_A + n_B$

Setting - Theorem

Theorem: Under $H_0 : \mu_A = \mu_B$ the type 1 error probability in (1) (under misspecification) for the allocation sequence $\mathbf{Z} = (Z_1, \ldots, Z_n)$ is

$$P\left(|T| > t_{n_A+n_B-2}(1-\alpha/2)\middle|\mathbf{Z}\right)$$
$$= F_{n_A+n_B-2,\theta_1,\theta_2}\left(t_{n_A+n_B-2}(\alpha/2)\right) + F_{n_A+n_B-2,-\theta_1,\theta_2}\left(t_{n_A+n_B-2}(\alpha/2)\right).$$

$F_{n_A+n_B-2,\theta_1,\theta_2}$ denotes the distribution function of the doubly non-central t-distribution with $n_A + n_B - 2$ degrees of freedom and parameters

$$\theta_1 = \frac{1}{\sigma}\sqrt{\frac{n_A n_B}{n_A + n_B}}\left(\tilde{\tau}_A - \tilde{\tau}_B\right) \quad \theta_2 = \frac{1}{\sigma^2}\left[\sum_{i=1}^{n}\tau_i^2 - n_A\tilde{\tau}_A^2 - n_B\tilde{\tau}_B^2\right]$$

where $\tilde{\tau}_A = \frac{1}{n_A}\sum_{i=1}^{n}\tau_i Z_i$; $\quad \tilde{\tau}_B = \frac{1}{n_B}\sum_{i=1}^{n}\tau_i(1-Z_i)$

Setting - Distribution of the Test Statistic (1)

Sketch of the proof:

for the given allocation vector $\mathbf{Z} = (Z_1, \ldots, Z_n)$

$$\tilde{y}_A - \tilde{y}_B \sim N\left(\mu_A - \mu_B + \tilde{\tau}_A - \tilde{\tau}_B, \sigma^2 \frac{n_A + n_B}{n_A n_B}\right)$$

where $\tilde{\tau}_A = \frac{1}{n_A} \sum_{i=1}^{n} \tau_i Z_i$; $\quad \tilde{\tau}_B = \frac{1}{n_B} \sum_{i=1}^{n} \tau_i (1 - Z_i)$

Setting - Distribution of the Test Statistic (2)

for the allocation vector $\mathbf{Z} = (Z_1, \ldots, Z_n)$

$$\sum_{i=1}^{n} Z_i (y_i - \tilde{y}_A)^2 + \sum_{i=1}^{n} (1 - Z_i)(y_i - \tilde{y}_B)^2 \sim \chi_{n_A + n_B - 2}(\theta_2)$$

with non-centrality parameter

$$\begin{aligned} \theta_2 &= \frac{1}{\sigma} \left(\sum_{i=1}^{n} Z_i (\tau_i - \tilde{\tau}_A)^2 + \sum_{i=1}^{n} (1 - Z_i)(\tau_i - \tilde{\tau}_B)^2 \right) \\ &= \frac{1}{\sigma} \left[\sum_{i=1}^{n} \tau_i^2 - n_A \tilde{\tau}_A^2 - n_B \tilde{\tau}_B^2 \right] \end{aligned}$$

Setting - Distribution of the Test Statistic (3)

Thus T follows a doubly non-central t distribution with $n_A + n_B - 2$ degrees of freedom and non-centrality parameters (Johnson, Kotz, Balakrishnan, 1995, Robins, 1948)

$$\theta_1 = \frac{1}{\sigma}\sqrt{\frac{n_A n_B}{n_A + n_B}}\left(\mu_A - \mu_B + \tilde{\tau}_A - \tilde{\tau}_B\right) = \frac{1}{\sigma}\sqrt{\frac{n_A n_B}{n_A + n_B}}\left(\tilde{\tau}_A - \tilde{\tau}_B\right)$$

$$\theta_2 = \frac{1}{\sigma^2}\left[\sum_{i=1}^{n} \tau_i^2 - n_A\tilde{\tau}_A^2 - n_B\tilde{\tau}_B^2\right]$$

using the properties of the distribution (Kocherlakota, 1991)

$$F_{\nu,\theta_1,\theta_2}(t) = 1 - F_{\nu,-\theta_1,\theta_2}(-t)$$

Assessment - Randomization Procedures

CR Complete randomization is accomplished by tossing a fair coin, so the probability that patient i will receive treatment 1 is always $\frac{1}{2}$

RAR Random Allocation rule, fix total sample size N. Randomize so that half the patients receive treatment 1

PBR (Permuted Block Randomization) Implementation of RAR within k Blocks of size $m_s, 1 \leq s \leq k$

BSD(a) (Big Stick design) CR allow for imbalance within a limit a

FP7 HEALTH 2013 - 602552

Assessment - Selection Bias Model

With $N_j(i-1)$ is the number of treatment j assignments after $(i-1)$ assignments and

$$p_A(i-1) = \frac{n_A - N_A(i-1)}{(n_A + n_B) - (N_A(i-1) + N_B(i-1))}, \quad n_A = n_B$$

the selection biasing policy (Tamm, 2012), according to convergence strategy (Blackwell, 1957), $q \in [\frac{1}{2}, 1]$

$$E(y_i) = \begin{cases} \mu_A Z_i + \mu_B(1 - Z_i) + \eta & \text{if } p_A(i-1) > q \\ \mu_A Z_i + \mu_B(1 - Z_i) & \text{if } 1 - q \leq p_A(i-1) \leq q \\ \mu_A Z_i + \mu_B(1 - Z_i) - \eta & \text{if } p_A(i-1) < 1 - q \end{cases}$$

Then

$$\tau_i = \eta \left[\mathbb{1}_{(q,1]}(p_A(i-1)) - \mathbb{1}_{[0,1-q)}(p_A(i-1)) \right]$$

FP7 HEALTH 2013 - 602552

Setting - Selection Bias

Table: Empirical type 1 error probability by different cut-off values for a two sided t-test with $2n = 40; \alpha = 0.05, \beta = 0.2$ and selection effect $\eta = \frac{\delta}{2} = 0.45$ under PBR

Cutoff	Type 1 Error rate		
q	PBR(4)	PBR(8)	PBR(10)
1/2	0.191	0.134	0.118
2/3	0.145	0.134	0.086
1	0.050	0.050	0.048

using SAS with 10 000 replications

Assessment - Selection Bias

Table: Empirical type 1 error probability of a two sided t-test with $2n$; $\alpha = 0.05$, $\beta = 0.2$ and selection bias effect $\eta = \frac{\delta(2n)}{2}$

2n	$\delta(2n)$	CR	RAR	PBR(4)	BSD (2)
8	2.381	0.058	0.102	0.141	0.064
20	1.325	0.054	0.082	0.177	0.075
32	1.024	0.055	0.072	0.188	0.083
40	0.909	0.053	0.071	0.195	0.088

using R with 100 000 replications

Assessment - Selection Bias

Tamm (2012) found:

- empirical type I error increases with smaller blocksize under PBR
- empirical type I error decreases with higher q
- elevation are substantial even for $q = 2/3$

Assessment - Chronological Bias Tamm (2014)

- long recruitment time in Rare Diseases, (EMA, 2006)
  - changes in population characteristics
  - learning effect in therapy / surgical experience (Hopper, 2007)
  - change in diagnosis (FDA, 2011), etc.
- special form of accidential bias, when considering a time-heterogeneous covariate

$$\tau_i = \lambda \begin{cases} \frac{i}{n_A+n_B} & \text{linear time trend} \\ \mathbb{1}_{i \geq c}(i) & \text{stepwise trend} \\ \log(\frac{i}{n_A+n_B}) & \text{log trend} \end{cases}$$

Assessment - Chronological Bias Tamm (2014)

Tamm (2014) found, that

- using t-test in presence of time trends results in conservative test decisions under permuted block randomizations even for small time trends; Blocked ANOVA should be used
- medium block sizes already achieve a good reduction of the inflated type I error rate in worst-case scenarios

Assessment - Selection and chronological bias for special sequences

all Sequences assume $n_A = n_B = n/2$

Sequence	Bias	θ_1	θ_2	type 1 error probability
Alternating 10101...	Selection	$-\frac{\eta}{\sigma}\sqrt{n}/2$	0	0.074
	Time Trend	$\frac{\lambda}{\sigma}\sqrt{n}/2$	$\frac{\lambda^2}{\sigma^2}n(n^2-4)/12$	0.000
Separating 111...000...	Selection	$-\frac{\eta}{\sigma}\sqrt{1/n}$	$\frac{\eta^2}{\sigma^2}(n-2)/n$	0.050
	Time Trend	$\frac{\lambda}{\sigma}\sqrt{n^3}/4$	$\frac{\lambda^2}{\sigma^2}n(n^2-4)/48$	0.999
avoid TT 1001 1001..	Selection	$-\frac{\eta}{\sigma}\sqrt{n}/2$	$\frac{\eta^2}{\sigma^2}$	0.050
	Time Trend	0	$\frac{\lambda^2}{\sigma^2}n(-1+n^2)/12$	0.000

type 1 error probability setting: $n_A = n_B = 6, \eta = \frac{\delta(12)}{4} = 0.45, \lambda = \sigma = 1$

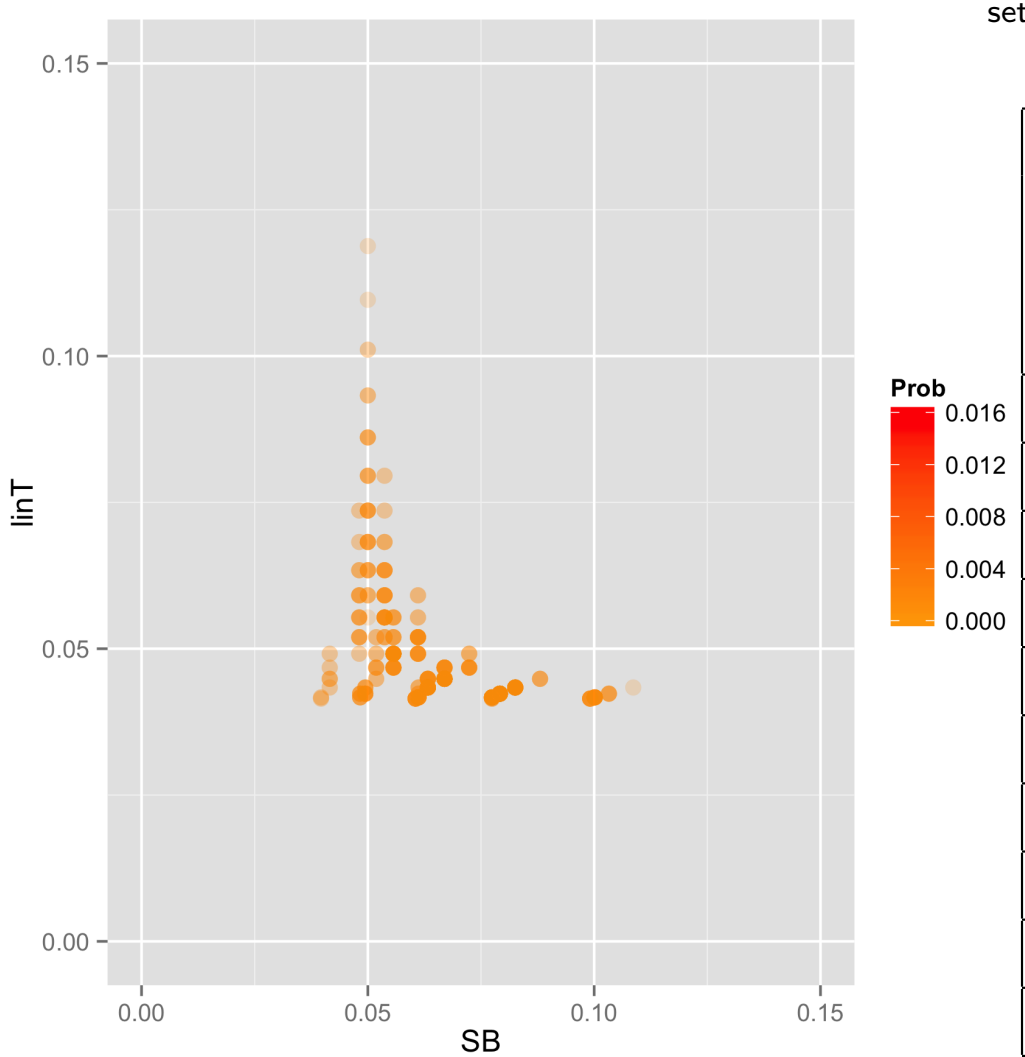
Assessment - Joint additive Bias

weighted additive (selection and chronological) bias model

$$\tau_i = \underbrace{\lambda \frac{i}{n_A + n_B}}_{\text{time trend}} + \underbrace{\eta \, \text{sign}(N_A(i-1) - N_B(i-1))}_{\text{selection bias with } q=\frac{1}{2}}$$

- weights via definition of λ and η
- different shape of time trend can be incorporated
- relaxed version of selection bias possible
- multiplicative could also be done

Assessment - Assessment of Bias for RAR (N=12)

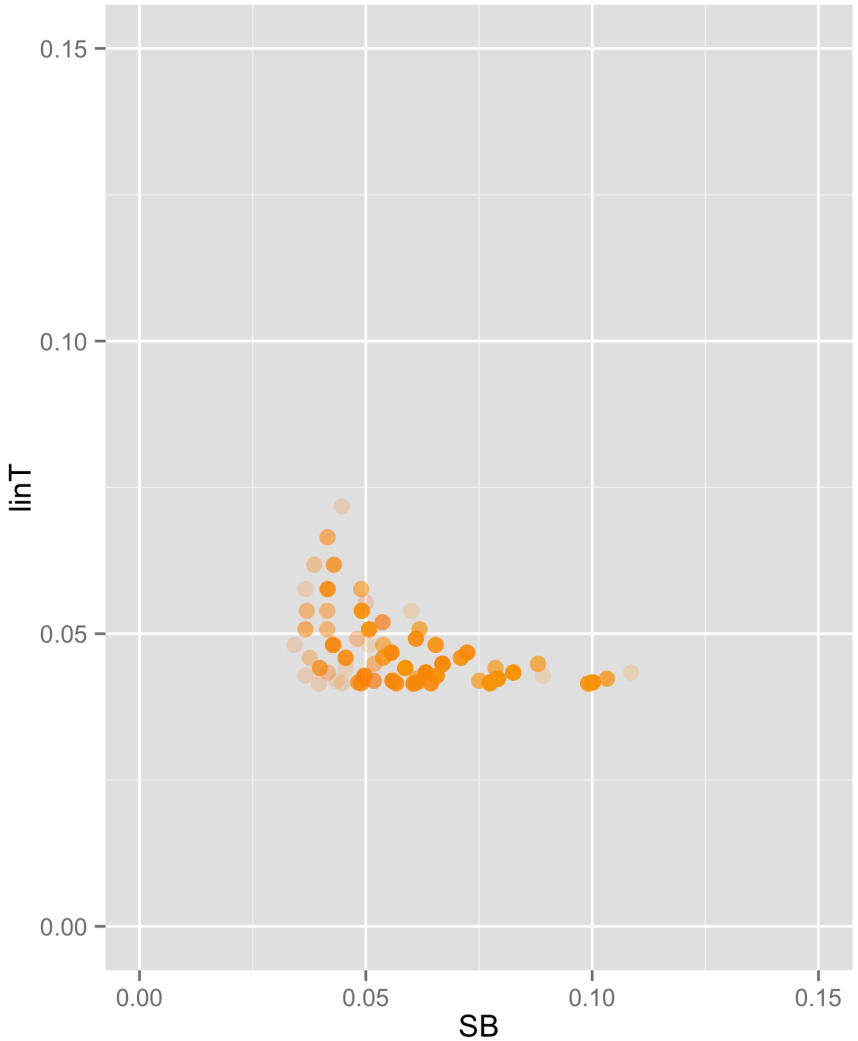


setting: $n_A = n_B = 6, \eta = \frac{\delta(12)}{4} = 0.45, \lambda = \sigma = 1$

Measure	p value	
	Selection Bias	Linear Time Trend
min	0.0397	0.0415
x05	0.0481	0.0414
x25	0.0518	0.0423
x50	0.0611	0.0449
mean	0.0635	0.0500
x75	0.0724	0.0520
x95	0.0991	0.0736
max	0.1086	0.1188
sd	0.0151	0.0115

FP7 HEALTH 2013 - 602552

Assessment - Assessment of Bias for BSD(2) (N=12)



setting: $n_A = n_B = 6, \eta = \frac{\delta(12)}{4} = 0.45, \lambda = \sigma = 1$

Measure	p value	
	Selection Bias	Linear Time Trend
min	0.0342	0.0415
x05	0.0370	0.0415
x25	0.0483	0.0420
x50	0.0537	0.0441
mean	0.0557	0.0462
x75	0.0619	0.0481
x95	0.0792	0.0576
max	0.1086	0.0717
sd	0.0129	0.0059

FP7 HEALTH 2013 - 602552

Assessment - Comparison RAR and BSD(2) (N=12)

Figure: RAR

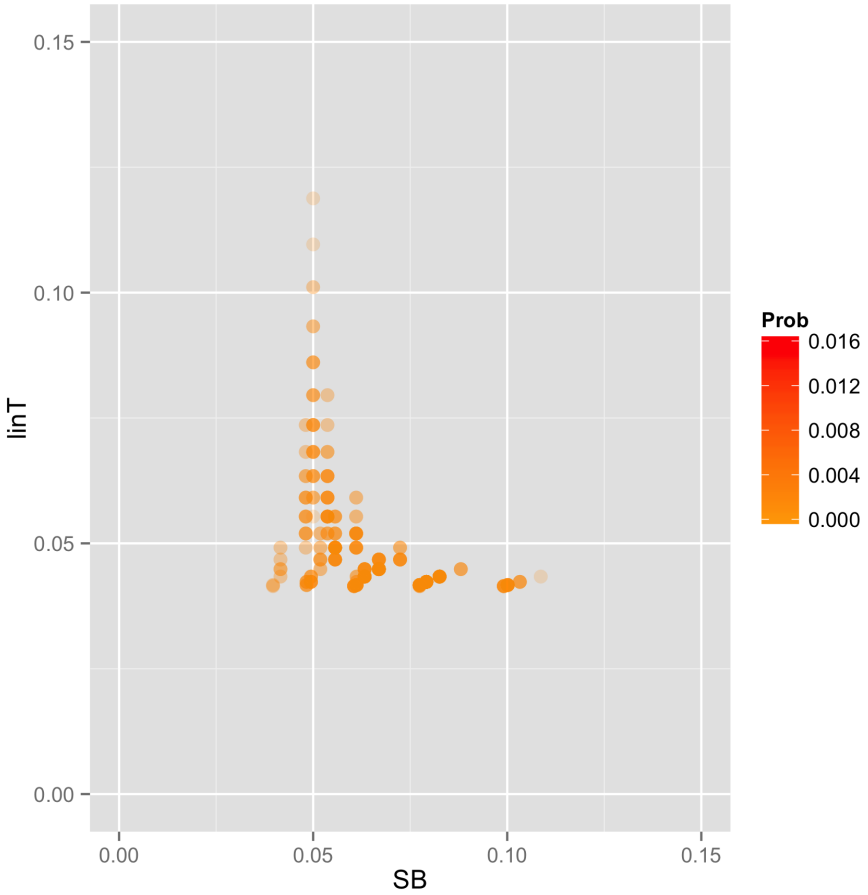
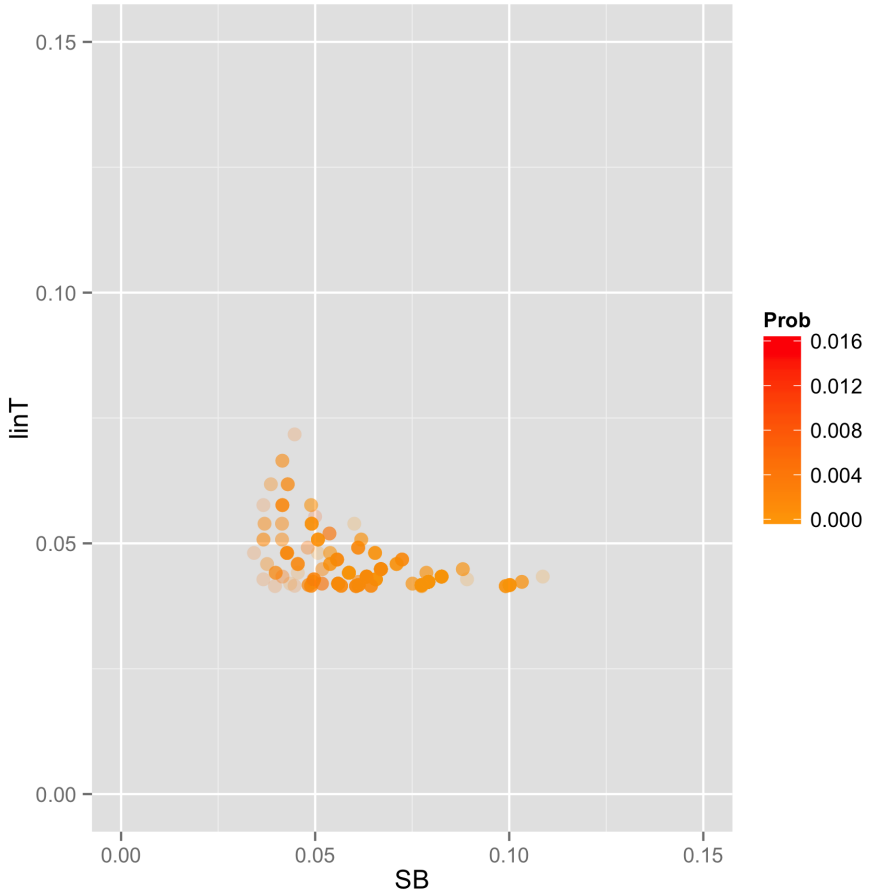


Figure: BSD(2)



setting: $n_A = n_B = 6, \eta = \frac{\delta(12)}{4} = 0.45, \lambda = \sigma = 1$

FP7 HEALTH 2013 - 602552

Conclusion

- presented a framework for scientific evaluation of randomization procedures on bias
- other randomization procedures are easy to implement
- this evaluation should be part of the TSAP
- dichotomous endpoint / survival endpoint could be evaluated by simulation *(work under progress by Marcia Rückbeil)*
- Cave: this is the wrong test (!) but mirrors the practical situation, use randomization tests to reflect the randomzation argument
- R package *(randomizeR)* coming soon, see website (http://www.ideal.rwth-aachen.de)

FP7 HEALTH 2013 - 602552

References

- Balakrishnan N, Kotz S, Johnson NL. Wiley: New York, 1995.
- Blackwell D, Hodges J. *Annals of Mathematical Statistics* 1957; **25**:449-460.
- ICH E9 (1998). Statistical principles for clinical trials.
- Joppi R, Bertele' V, Garattini S. (2013). *European Journal of Clinical Pharmacology* 2013; **69**:1009-1024.
- Kennes, L. N., Cramer, E., Hilgers, R. E., and Heussen, N. (2011). *Statistics in Medicine* 2011; **30**:2573-2581.
- Langer, S. (2012). *Master Thesis Mathematics RWTH Aachen*
- Rosenberger W, Lachin J. Wiley: New York, 2002.
- Tamm M, Hilgers RD. *Methods of Information in Medicine* 2012; **51**:138-143.
- Tamm M, Hilgers RD. *Methods of Information in Medicine* 2014; **53**:501-510.

FP7 HEALTH 2013 - 602552