Type I Error Rate Control in Adaptive Clinical Trials with Blinded Interim Analysis

Martin Posch and Magdalena Malina
Medical University of Vienna, Vienna, Austria

ROeS Seminar
Dornbirn, 2013

Blinded. And therefore unbiased?

Are adaptations based on blinded (interim) data unsuspicous to generate bias?

ICH E9: Statistical principles for clinical trials

European Medicines Agency

September 1998
CPMP/ICH/363/96

ICH Topic E 9
Statistical Principles for Clinical Trials

The statistical analysis plan should be reviewed and possibly updated as a result of the blind review of the data:

- Exclusion of subjects or data from the analysis sets;
- Possible transformations may also be checked, and outliers defined;
- Important covariates identified in other recent research may be added to the model;
- The use of parametric or non-parametric methods may be reconsidered.

ICH E9: Statistical principles for clinical trials

"Decisions made at this time should be described in the report, and should be distinguished from those made after the statistician has had access to the treatment codes, as blind decisions will generally introduce less potential for bias"

Adaptive Designs Reflection Paper

London, 18 October 2007
Doc. Ref. CHMP/EWP/2459/02

COMMITTEE FOR MEDICINAL PRODUCTS FOR HUMAN USE
(CHMP)

REFLECTION PAPER ON METHODOLOGICAL ISSUES IN CONFIRMATORY CLINICAL TRIALS PLANNED WITH AN ADAPTIVE DESIGN

- "Routinely breaking the blind should be avoided (...)"
- "Whenever possible, methods for blinded sample size reassessment that properly control the type I error should be used (...)"

Guideline on Data Monitoring Committees

European Medicines Agency
Pre-authorisation Evaluation of Medicines for Human Use

London, 27 July 2005
Doc. Ref. EMEA/CHMP/EWP/5872/03 Corr

COMMITTEE FOR MEDICINAL PRODUCTS FOR HUMAN USE
(CHMP)

GUIDELINE ON DATA MONITORING COMMITTEES

- "Such operating procedures should also describe how the integrity of the study with respect to preventing dissemination of unblinded study information is ensured."
- "As mentioned several times a possible bias in the future conduct of the clinical trial might be induced by the dissemination of unblinded treatment information seen by the DMC."

Adaptations based on blinded data in clinical trials

Change of Endpoint E.g., because of missing data

POSCH AND PROSCHAN, 2012

Change of Population E.g., improve recruitment targeting alternate sources of patients

Addition of an interim analysis Because of lagging recruitment

Change of analysis strategy Because of inappropriate distributional assumptions

Change of sample size e.g., because of low event rate or large variability

Adaptations based on blinded data in clinical trials

Change of Endpoint E.g., because of missing data

POSCH AND PROSCHAN, 2012

Change of Population E.g., improve recruitment targeting alternate sources of patients

Addition of an interim analysis Because of lagging recruitment

Change of analysis strategy Because of inappropriate distributional assumptions

Change of sample size e.g., because of low event rate or large variability

7

Sample Size Reassessment and Type I error inflation?

- Consider a clinical trial where an interim analysis is performed, where sample size for the remaining trial is reassessed based on blinded interim data.
- At the end of the trial a fixed sample test for the difference in means is performed.

What is the Type I error rate?

The Scenario

- Parallel group comparison with preplanned sample size N.
- Normally distributed primary endpoint X_i
- Test at $\alpha = 0.025$

$$H_0 : \mu_1 \leq \mu_0 \text{ against } H_1 : \mu_1 > \mu_0$$

- After $n = N/2$ observations per group an interim analysis is performed.
- The sample size is reassessed to $\tilde{N}$.
- In the final analysis the unadjusted z-test is performed at nominal level α (assuming σ is known).

Type I Error Rate after Sample Size Reassessment

- Unblinded interim analysis: max. type I error rate: 0.0625

  PROSCHAN & HUNSBERGER (1995)

  The type I error rate can be controlled by

Type I Error Rate after Sample Size Reassessment

- Unblinded interim analysis: max. type I error rate: 0.0625

  PROSCHAN & HUNSBERGER (1995)

  The type I error rate can be controlled by

  - restricted sample size reassessment rules

    CHEN, DEMETS & LAN (2004), MEHTA & POCOCK (2010)

10

Type I Error Rate after Sample Size Reassessment

- Unblinded interim analysis: max. type I error rate: 0.0625

  PROSCHAN & HUNSBERGER (1995)

  The type I error rate can be controlled by
  - restricted sample size reassessment rules

    CHEN, DEMETS & LAN (2004), MEHTA & POCOCK (2010)
  - adjusted tests

    BAUER & KÖHNE 1994, SCHÄFER & MÜLLER (2001), ...

Type I Error Rate after Sample Size Reassessment

- Unblinded interim analysis: max. type I error rate: 0.0625

  PROSCHAN & HUNSBERGER (1995)

  The type I error rate can be controlled by
  - restricted sample size reassessment rules

    CHEN, DEMETS & LAN (2004), MEHTA & POCOCK (2010)
  - adjusted tests

    BAUER & KÖHNE 1994, SCHÄFER & MÜLLER (2001), ...
- Blinded interim analysis and sample size recalculation with lumped variance estimate: No α-inflation for reasonable sample sizes (for the t-test).

  FRIEDE & KIESER (2003)

Type I Error Rate after Sample Size Reassessment

- Unblinded interim analysis: max. type I error rate: 0.0625

  PROSCHAN & HUNSBERGER (1995)

  The type I error rate can be controlled by
  - restricted sample size reassessment rules

    CHEN, DEMETS & LAN (2004), MEHTA & POCOCK (2010)
  - adjusted tests

    BAUER & KÖHNE 1994, SCHÄFER & MÜLLER (2001), ...
- Blinded interim analysis and sample size recalculation with lumped variance estimate: No α-inflation for reasonable sample sizes (for the t-test).

  FRIEDE & KIESER (2003)

Is it in general true that in blinded interim analyses there is no information on the observed interim effect size and the sample size reassessment is conservative?

Effect Estimation Based on the Primary Endpoint Only

- E-M algorithm to estimate effect size.

  GOULD AND SHIH (1998)
- Permuted block randomization allows to partially unblind data.

  VAN DER MEULEN (2007)

Can these methods give reliable estimates?

Only under the alternative, for large effect sizes that are not realistic for clinical trials.

FRIEDE & KIESER 2002, MILLER, FRIEDE, KIESER (2009)

Effect Estimation Based on the Primary Endpoint Only

- E-M algorithm to estimate effect size.

  GOULD AND SHIH (1998)

- Permuted block randomization allows to partially unblind data.

  VAN DER MEULEN (2007)

Can these methods give reliable estimates?

Only under the alternative, for large effect sizes that are not realistic for clinical trials.

FRIEDE & KIESER 2002, MILLER, FRIEDE, KIESER (2009)

Can the observed treatment effect be estimated if additionally a biomarker endpoint is measured?

Unblinding based on Biomarker Information

- Assme it is a priori known that the drug has an impact on a biomarker (e.g., a safety endpoint, a laboratory parameter, or the level of experimental drug in the blood) such that for the biomarker Y there is a treatment effect $(\nu_0 \neq \nu_1)$.
- the primary endpoint X the null hypothesis holds $(\mu_0 = \mu_1)$
- primary endpoint and biomarker are stochastically independent (will be relaxed later)

Can we estimate the observed treatment effect in the primary endpoint without unblinding?

Partial unblinding of the blinded sample

- Treatment group indicator $g_i \in \{0, 1\}, i = 1, \ldots, n$,

Partial unblinding of the blinded sample

- Treatment group indicator $g_i \in \{0, 1\}, i = 1, \ldots, n$,
- The g_i are iid Bernoulli and $P(g_i = 1) = 1/2$.

Partial unblinding of the blinded sample

- Treatment group indicator $g_i \in \{0, 1\}, i = 1, \ldots, n$,
- The g_i are iid Bernoulli and $P(g_i = 1) = 1/2$.
- The unblinded interim effect estimate is

$$\bar{x}_1 - \bar{x}_0 = \frac{\sum_{i=1}^{n} g_i x_i}{n_1} - \frac{\sum_{i=1}^{n} (1 - g_i) x_i}{n_0},$$

  Where $n_1 = \sum_{i=1}^{n} g_i$ and $n_1 = n - n_0$.

Partial unblinding of the blinded sample

- Treatment group indicator $g_i \in \{0, 1\}, i = 1, \ldots, n$,
- The g_i are iid Bernoulli and $P(g_i = 1) = 1/2$.
- The unblinded interim effect estimate is

$$\bar{x}_1 - \bar{x}_0 = \frac{\sum_{i=1}^{n} g_i x_i}{n_1} - \frac{\sum_{i=1}^{n} (1 - g_i) x_i}{n_0},$$

  Where $n_1 = \sum_{i=1}^{n} g_i$ and $n_1 = n - n_0$.
- Assume from historic data the effect sizes ν_0, ν_1 and variance σ_y^2 of the biomarker Y are known.

Partial unblinding of the blinded sample

- Treatment group indicator $g_i \in \{0, 1\}, i = 1, \ldots, n$,
- The g_i are iid Bernoulli and $P(g_i = 1) = 1/2$.
- The unblinded interim effect estimate is

$$\bar{x}_1 - \bar{x}_0 = \frac{\sum_{i=1}^{n} g_i x_i}{n_1} - \frac{\sum_{i=1}^{n} (1 - g_i) x_i}{n_0},$$

  Where $n_1 = \sum_{i=1}^{n} g_i$ and $n_1 = n - n_0$.
- Assume from historic data the effect sizes ν_0, ν_1 and variance σ_y^2 of the biomarker Y are known.
- The treatment allocation can be estimated by

$$E(g_i|y_i) = \frac{\varphi_{\nu_1,\sigma_y}(y_i)}{\varphi_{\nu_1,\sigma_y}(y_i) + \varphi_{\nu_0,\sigma_y}(y_i)},$$

  where φ denotes the standard normal density.

Examples

- Distribution of the biomarker: $\nu_0 = 0$, $\nu_1 = 2$, $\sigma_y = 1$:

X_i	Y_i	$E(g_i\|y_i)$
0.38	0.78	0.39
-0.55	2.00	0.88
-0.49	1.61	0.77
0.62	-0.68	0.03

- if $\nu_0 = \nu_1$ then $E(g_i|y_i) = .5$.
- If $|\nu_1 - \nu_0| \to \infty$ then $E(g_i|y_i) \to g_i$.

The larger the biomarker effect size $\nu_1 - \nu_0$, the better the data can be unblinded.

The blinded treatment effect estimate

$$\hat{\Delta}_B = \frac{\sum_{i=1}^{n} E(g_i|y_i)x_i}{\bar{n}_1} - \frac{\sum_{i=1}^{n}(1 - E(g_i|y_i))x_i}{\bar{n}_0},$$

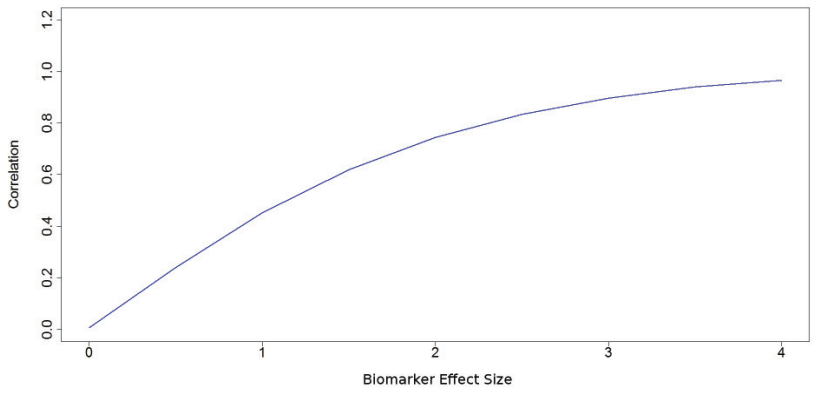
where $\bar{n}_1 = \sum_{i=1}^{n} E(g_i|y_i)$ and $\bar{n}_0 = 1 - \bar{n}_1$.

The blinded treatment effect estimate

$$\hat{\Delta}_B = \frac{\sum_{i=1}^{n} E(g_i|y_i)x_i}{\bar{n}_1} - \frac{\sum_{i=1}^{n}(1 - E(g_i|y_i))x_i}{\bar{n}_0},$$

where $\bar{n}_1 = \sum_{i=1}^{n} E(g_i|y_i)$ and $\bar{n}_0 = 1 - \bar{n}_1$.

Correlation of blinded and unblinded effect estimate $\rho(\bar{X}_1 - \bar{X}_0, \hat{\Delta}_b)$



What if we use in addition primary endpoint data?

- The primary endpoint and the biomarker are bivariate normal with correlation ρ in each group.

What if we use in addition primary endpoint data?

- The primary endpoint and the biomarker are bivariate normal with correlation ρ in each group.
- Given the blinded data X and Y

$$E(g_i|x_i, y_i) = \frac{\varphi_{\mu,\nu_1,\rho}(x_i, y_i)}{\varphi_{\mu,\nu_0,\rho}(x_i, y_i) + \varphi_{\mu,\nu_1,\rho}(x_i, y_i)},$$

where $\varphi_{\mu,\nu_0,\rho}$ denotes the bivariate normal density.

What if we use in addition primary endpoint data?

- The primary endpoint and the biomarker are bivariate normal with correlation ρ in each group.
- Given the blinded data X and Y

$$E(g_i|x_i, y_i) = \frac{\varphi_{\mu,\nu_1,\rho}(x_i, y_i)}{\varphi_{\mu,\nu_0,\rho}(x_i, y_i) + \varphi_{\mu,\nu_1,\rho}(x_i, y_i)},$$

  where $\varphi_{\mu,\nu_0,\rho}$ denotes the bivariate normal density.
- if $\rho = 0$ then $E(g_i|x_i, y_i) = E(g_i|y_i)$.

What if we use in addition primary endpoint data?

- The primary endpoint and the biomarker are bivariate normal with correlation ρ in each group.
- Given the blinded data X and Y

$$E(g_i|x_i, y_i) = \frac{\varphi_{\mu,\nu_1,\rho}(x_i, y_i)}{\varphi_{\mu,\nu_0,\rho}(x_i, y_i) + \varphi_{\mu,\nu_1,\rho}(x_i, y_i)},$$

  where $\varphi_{\mu,\nu_0,\rho}$ denotes the bivariate normal density.
- if $\rho = 0$ then $E(g_i|x_i, y_i) = E(g_i|y_i)$.
- if $\rho = 1$ and $\nu_0 \neq \nu_1$, then $E(g_i|x_i, y_i) = g_i$.

16

What if we use in addition primary endpoint data?

- The primary endpoint and the biomarker are bivariate normal with correlation ρ in each group.
- Given the blinded data X and Y

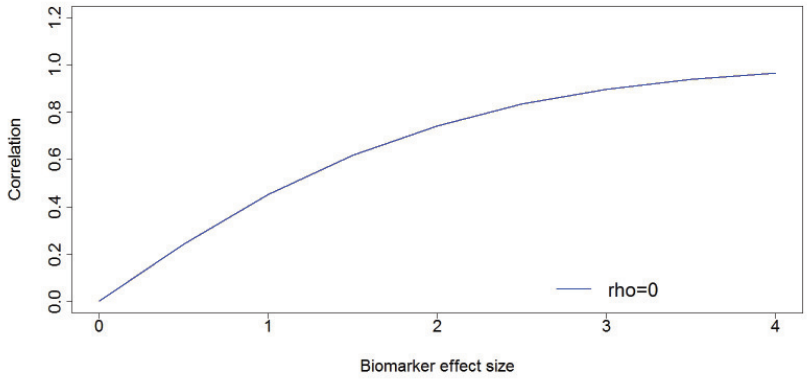
$$E(g_i|x_i, y_i) = \frac{\varphi_{\mu,\nu_1,\rho}(x_i, y_i)}{\varphi_{\mu,\nu_0,\rho}(x_i, y_i) + \varphi_{\mu,\nu_1,\rho}(x_i, y_i)},$$

  where $\varphi_{\mu,\nu_0,\rho}$ denotes the bivariate normal density.
- if $\rho = 0$ then $E(g_i|x_i, y_i) = E(g_i|y_i)$.
- if $\rho = 1$ and $\nu_0 \neq \nu_1$, then $E(g_i|x_i, y_i) = g_i$.

The larger the correlation, the better the data can be unblinded.

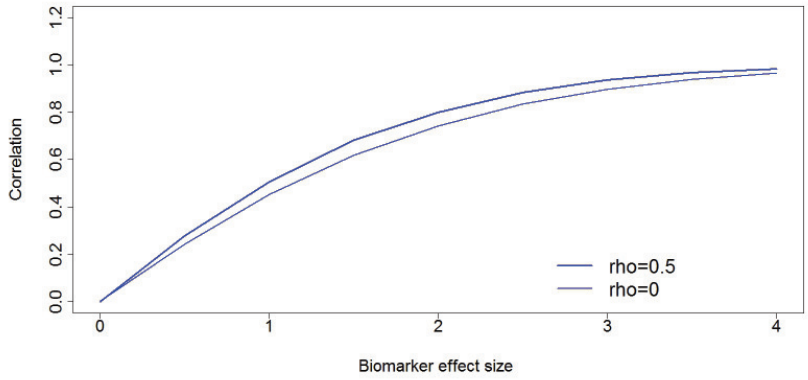
Blinded effect estimate for correlated biomarkers

Correlation between blinded and unblinded effect estimate $\rho(\bar{X}_1 - \bar{X}_0, \hat{\Delta}_b)$



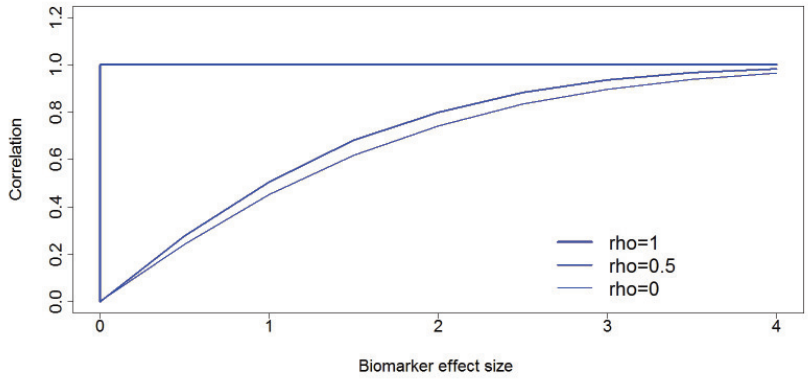
Blinded effect estimate for correlated biomarkers

Correlation between blinded and unblinded effect estimate $\rho(\bar{X}_1 - \bar{X}_0, \hat{\Delta}_b)$



Blinded effect estimate for correlated biomarkers

Correlation between blinded and unblinded effect estimate $\rho(\bar{X}_1 - \bar{X}_0, \hat{\Delta}_b)$



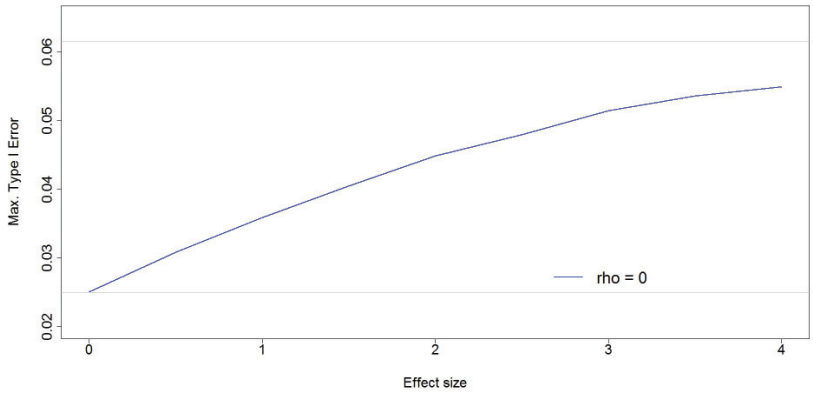
Maximum Type I Error Rate after Blinded Sample Size Reassessment

- Determine the “worst case” sample size reassessment rule, that maximizes the type I error rate, given the blinded effect estimate.
- In the interim analysis, choose the $\tilde{N}$ that maximizes

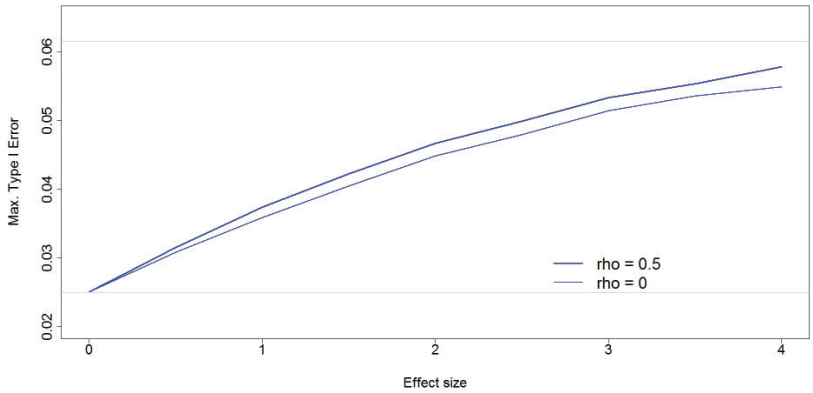
$$P_{H_0}\left(Z_{\tilde{N}} > z_{1-\alpha} | \hat{\Delta}_b\right),$$

where $Z_{\tilde{N}}$ is the z-test statistics of the test with reassessed sample size $\tilde{N}$.

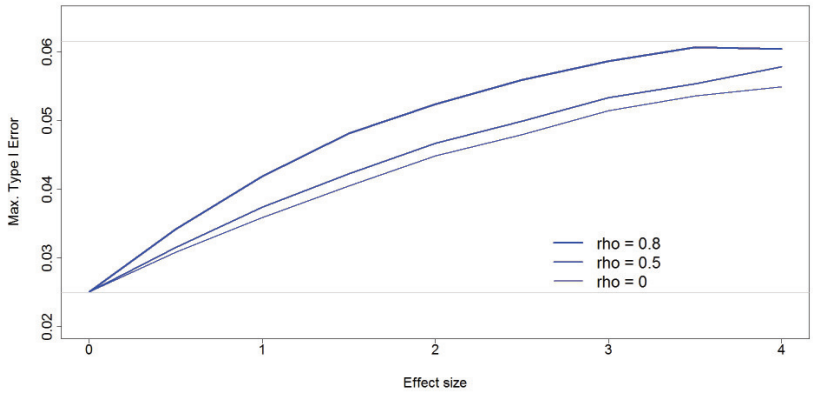
Maximum Type I Error Rate



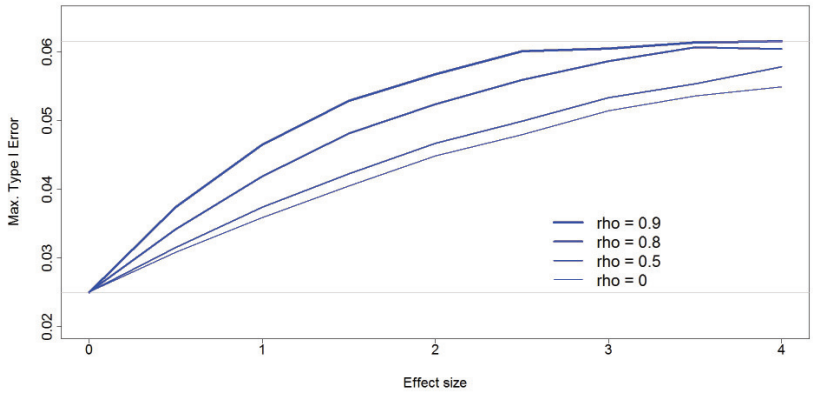
Maximum Type I Error Rate



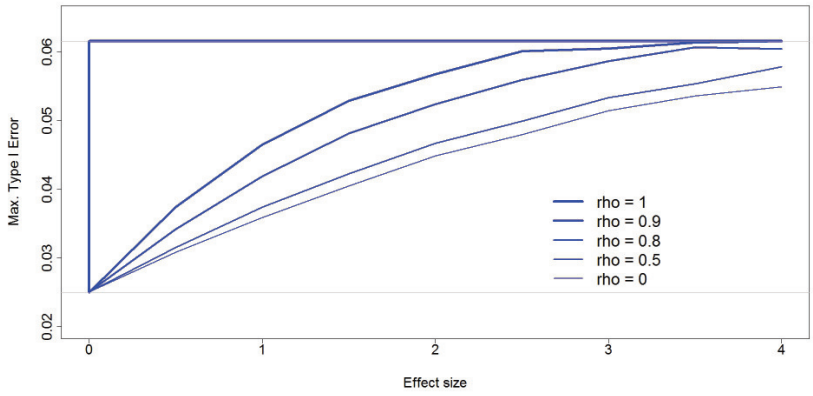
Maximum Type I Error Rate



Maximum Type I Error Rate



Maximum Type I Error Rate



What if we use in addition information on blocking?

- Consider blocked randomization with block size $k = 2$

What if we use in addition information on blocking?

- Consider blocked randomization with block size $k = 2$
- $P\{(g_1, g_2) = (1, 0)\} = P\{(g_1, g_2) = (0, 1)\} = 1/2$.

What if we use in addition information on blocking?

- Consider blocked randomization with block size $k = 2$
- $P\{(g_1, g_2) = (1, 0)\} = P\{(g_1, g_2) = (0, 1)\} = 1/2$.
- E.g., for the subjects in the first block, given the values of X and Y

$$P\{(g_1, g_2) = (1, 0) | x_1, x_2, y_1, y_2\} = \frac{\varphi_{\mu,\nu_1,\rho}(x_1, y_1)\varphi_{\mu,\nu_0,\rho}(x_2, y_2)}{\varphi_{\mu,\nu_0,\rho}(x_1, y_1)\varphi_{\mu,\nu_1,\rho}(x_2, y_2) + \varphi_{\mu,\nu_1,\rho}(x_1, y_1)\varphi_{\mu,\nu_0,\rho}(x_2, y_2)}.$$

What if we use in addition information on blocking?

- Consider blocked randomization with block size $k = 2$
- $P\{(g_1, g_2) = (1, 0)\} = P\{(g_1, g_2) = (0, 1)\} = 1/2$.
- E.g., for the subjects in the first block, given the values of X and Y

$$P\{(g_1, g_2) = (1, 0) | x_1, x_2, y_1, y_2\} = \frac{\varphi_{\mu,\nu_1,\rho}(x_1, y_1)\varphi_{\mu,\nu_0,\rho}(x_2, y_2)}{\varphi_{\mu,\nu_0,\rho}(x_1, y_1)\varphi_{\mu,\nu_1,\rho}(x_2, y_2) + \varphi_{\mu,\nu_1,\rho}(x_1, y_1)\varphi_{\mu,\nu_0,\rho}(x_2, y_2)}.$$

- Similarly, for larger block sizes (need to sum over all possible permutations).

20

What if we use in addition information on blocking?

- Consider blocked randomization with block size $k = 2$
- $P\{(g_1, g_2) = (1, 0)\} = P\{(g_1, g_2) = (0, 1)\} = 1/2$.
- E.g., for the subjects in the first block, given the values of X and Y

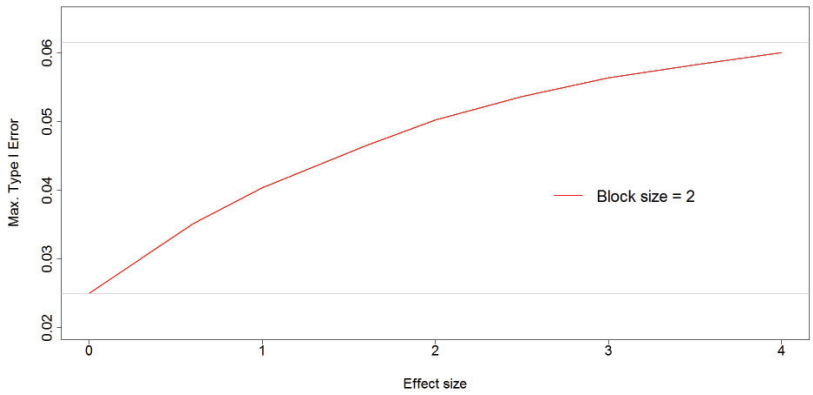
$$P\{(g_1, g_2) = (1, 0) | x_1, x_2, y_1, y_2\} = \frac{\varphi_{\mu,\nu_1,\rho}(x_1, y_1)\varphi_{\mu,\nu_0,\rho}(x_2, y_2)}{\varphi_{\mu,\nu_0,\rho}(x_1, y_1)\varphi_{\mu,\nu_1,\rho}(x_2, y_2) + \varphi_{\mu,\nu_1,\rho}(x_1, y_1)\varphi_{\mu,\nu_0,\rho}(x_2, y_2)}.$$

- Similarly, for larger block sizes (need to sum over all possible permutations).

The smaller the block size, the better the data can be unblinded.

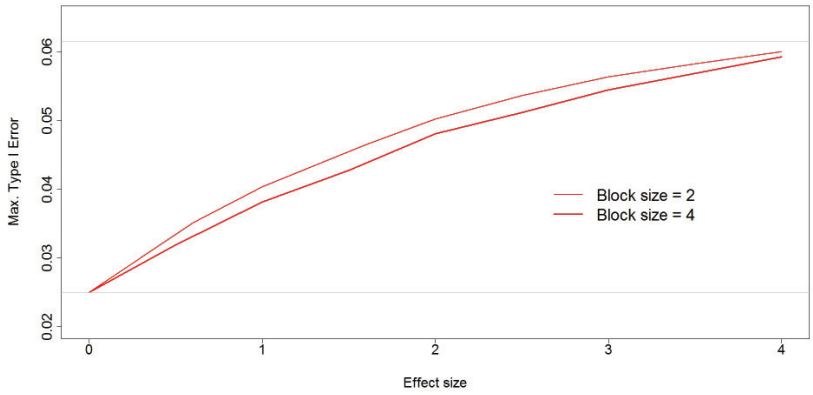
Maximum Type I Error Rate

Different Block Sizes - Independent Biomarker



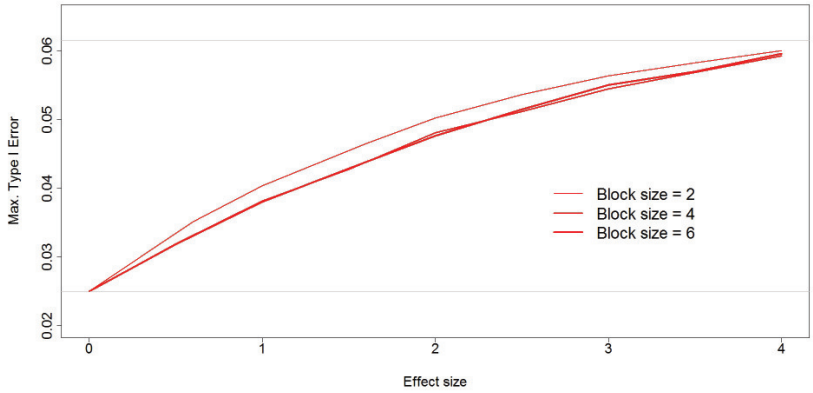
Maximum Type I Error Rate

Different Block Sizes - Independent Biomarker



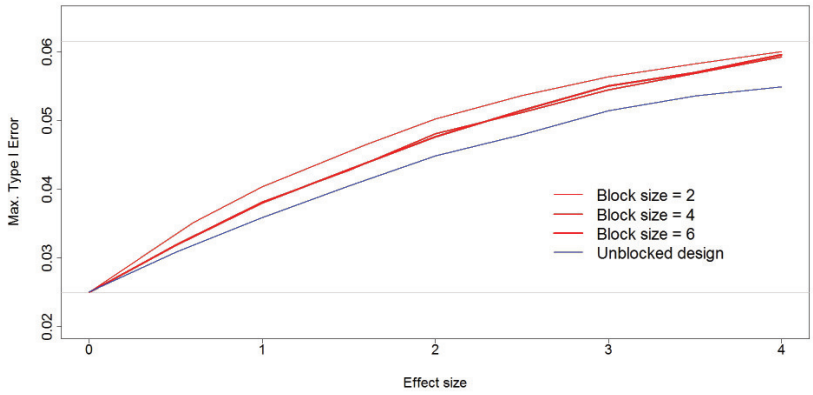
Maximum Type I Error Rate

Different Block Sizes - Independent Biomarker



Maximum Type I Error Rate

Different Block Sizes - Independent Biomarker

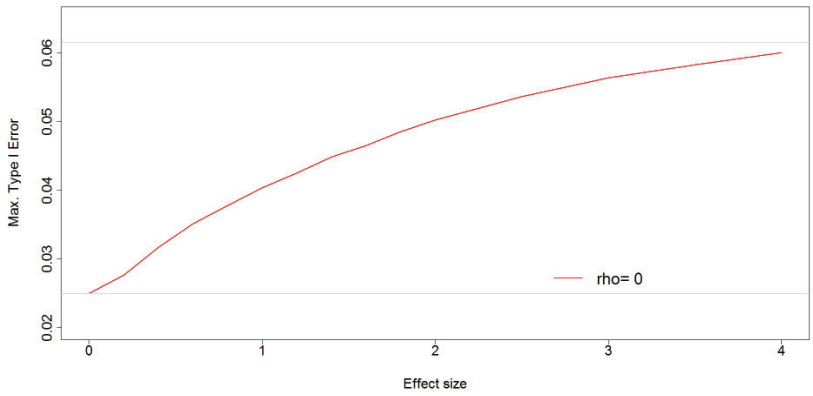


Maximum Type I Error Rate

Blocked Randomization (Block Size 2) and Correlated Biomarker

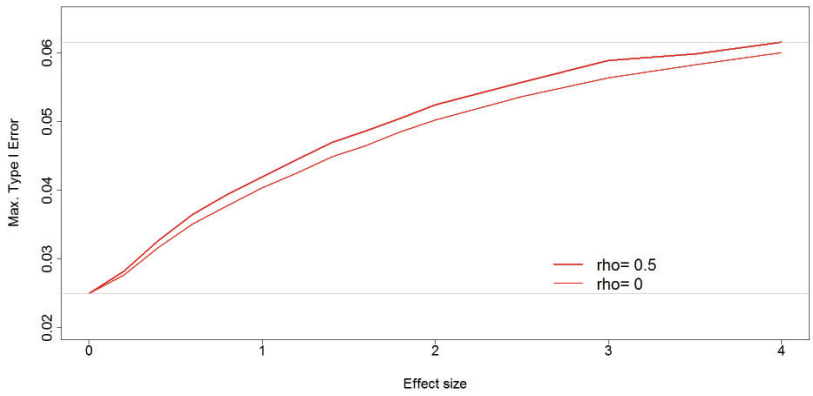
Maximum Type I Error Rate

Blocked Randomization (Block Size 2) and Correlated Biomarker



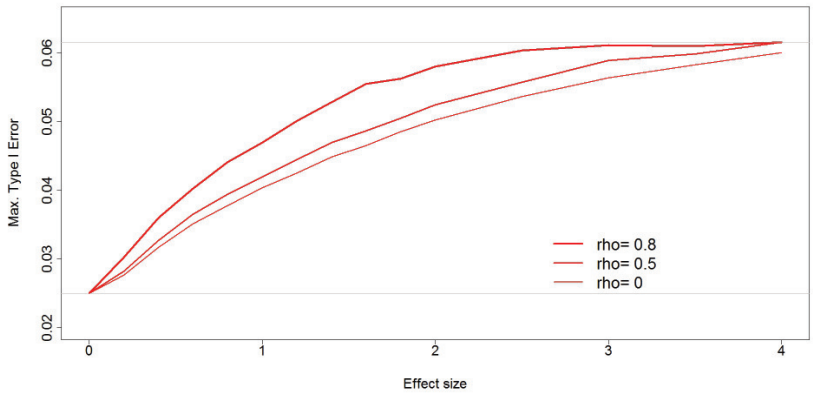
Maximum Type I Error Rate

Blocked Randomization (Block Size 2) and Correlated Biomarker



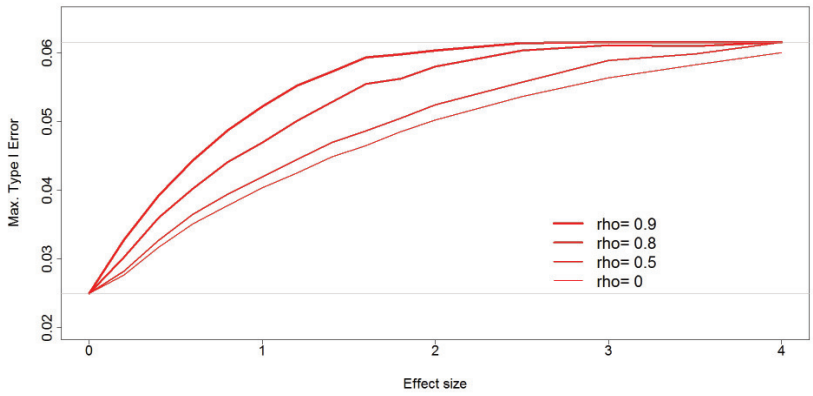
Maximum Type I Error Rate

Blocked Randomization (Block Size 2) and Correlated Biomarker



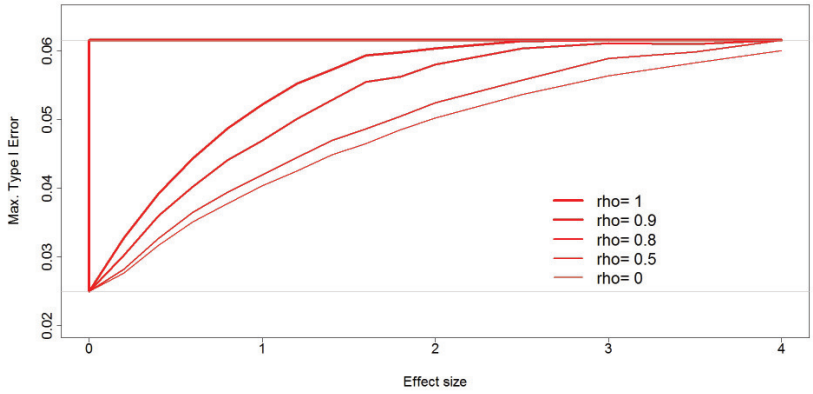
Maximum Type I Error Rate

Blocked Randomization (Block Size 2) and Correlated Biomarker



Maximum Type I Error Rate

Blocked Randomization (Block Size 2) and Correlated Biomarker



Discussion

- Assumption of known biomarker distribution can be relaxed
- Extension to more than one biomarker

Conclusion: Even blinded data reviews can reveal information on the treatment effect and may lead to biased tests.

Should sponsors' access to blinded data be more strictly regulated?