

Adaptive Designs and Treatment Selection

Franz König

Werner Brannath Martin Posch Peter Bauer

Section of Medical Statistics
Medical University of Vienna

ROES Seminar 2007
Bern, 2007

Flexible (Adaptive) Versus Frequentist Trial

Classical frequentist trials

- details of design and analysis must be prefixed in advance (population, treatments, doses, main and secondary outcome variable(s), analysis strategy, sample sizes,...)
- lack of flexibility to react to information from inside or outside the trial

Flexible (adaptive) design

- allow for mid-trial design modifications based on all internal and external information gathered at interim analyses without compromising the type I error rate
- **To control the type I error rate, the design modifications need not be specified in advance.**

Flexible Trials and Drug Development

Full flexibility (unscheduled adaptivity)=

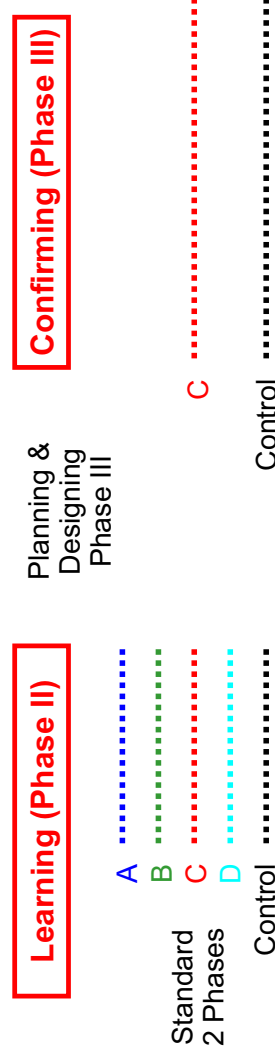
adapting design parameters without a (complete) specification of the adaptation rule.

- dealing with the unexpected
- dealing with the expected unpredictability of clinical trials

Flexible designs allow to...

- integrate Phase II and Phase III trials into a single trial - adaptive seamless designs (e.g., Bauer & Kieser 1999, Bretz et al 2006,...)
- formally integrate the data of exploratory and confirmatory phases
- speed up the drug development process
- react flexibly to unexpected events

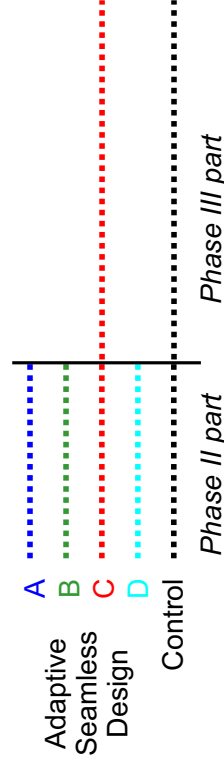
Separate Phase II and III Trials



- Conduct phase II trial.
- Plan phase III trial based on the information from phase II trial (which treatment, which number of patients, etc.).
- Conduct confirmatory phase III trial. **Demonstrate efficacy using ONLY phase III trial data.**

Adaptive Seamless Phase II + III Trials

Learning, Selecting and Confirming (Phase II & III)



- Conduct phase II trial as **internal part of a combined trial**.
- Plan phase III trial based on data from phase II part.
- Conduct phase III trial as **internal part of the same trial**.
- Demonstrate efficacy with data from **phase III + II part**.

Adaptive Seamless Phase II + III Trials

- Smaller time lag between phase II and phase III. Speeds up the drug development process.
- Allows us to use also the data from a (late) Phase II trial for the efficacy hypothesis testing. This saves resources (patients), costs and time.
- Improves quality of the drug development process by using the same study protocol (study plan) for the (late) phase II and (early) phase III trial.

Difficulties with Adaptive Seamless Phase II + III Trials

- Selecting the apparent “best” treatment at the end of the phase II part leads to a bias of the overall test statistic.
- Number of selected doses is a random number.
- Sample sizes will depend on the number of selected doses and the efficacy seen in the phase II part. This may further bias the overall test statistics.
- The dose selection process depends on several measurements (efficacy and safety) and is complex and typically unknown at the planning stage.
- Selection rule for doses and sample sizes for the phase III part **cannot** be laid down in detail in advance. Hence, the bias of the test statistic cannot be fully quantified!
- **Nevertheless, control of (multiple) type I error rate required!**

Approaches to Treatment Selection

- Methods for predefined selection rules
(THALL ET AL. 1988, 1989; STALLARD & TODD 2003, ...)
- **Flexible Two Stage Closed Tests**
(BAUER & KIESER 1999; HOMMEL 2001; POSCH ET AL. (2005); ...)
 - Do not require a predefined treatment and sample size selection rule.
 - Combine two methodology concepts:
Flexible Two Stage Tests and Closed Testing Principle.

Flexible Two Stage Test based on Combination Tests

(BAUER 1989, BAUER & KÖHNE 1994, ...)

Stage 1



Planning:

- Fix design (only) for Stage 1
- Fix combination function $C(p, q)$ and critical value c e.g. $C(p, q) = p \cdot q$

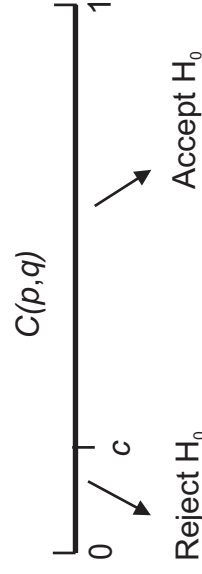
Stage 1:

- Compute p-value p from Stage 1 data
- Fix design for Stage 2 based on data from Stage 1

Planning of Stage 2



Stage 2



Stage 2:

- Compute p-value q from Stage 2 data.
- Reject H_0 iff $C(p, q) \leq c$.

Type I error control and combination functions

Type I error control

Type I error rate $\leq \alpha$ if we choose critical value c such that

$$P[C(p, q) \leq c] = \alpha$$

for independent and uniformly distributed p-values p and q .

- **Fisher product test:** $C(p, q) = p \cdot q$
(BAUER 1989, BAUER & KÖHNE, 1994)
- **Weighted inverse normal method:**
 $C(p, q) = \Phi(w_1 \Phi^{-1}(p) + w_2 \Phi^{-1}(q))$
(LEHMACHER & WASSMER, 1999)

Dose selection and efficacy testing

- Parallel group design with $k = 2$ dose groups and a control group (i.e., in total three parallel groups).
- Testing the one sided hypotheses

Dose 1 vs control: $H_{0,1} : \mu_1 \leq \mu_0$ vs. $H_{1,1} : \mu_1 > \mu_0$

Dose 2 vs control: $H_{0,2} : \mu_2 \leq \mu_0$ vs. $H_{1,2} : \mu_2 > \mu_0$

Dose selection and efficacy testing

- After Stage 1 we decide either to
 - go into Stage 2 with BOTH doses or
 - go into Stage 2 with only ONE dose.
- Selection rule unknown before end of Stage 1.
- Choice of sample sizes for Stage 2 depends on selected dose(s) and observed efficiency.
- Regulatory bodies ask for a level $\alpha = 0.025$ test of the intersection hypothesis

$$H_{0,1} \cap H_{0,2} : \mu_1, \mu_2 \leq \mu_0$$

Flexible Closed Test (BAUER & KIESER 1999, HOMMEL 2001)

- Use flexible two stage test for $H_{0,1} \cap H_{0,2}$, e.g. fix a combination test $C(p, q)$ at level α .
- At Stage 1 use a multiplicity adjusted p-value for p e.g. p-value of Šidak test

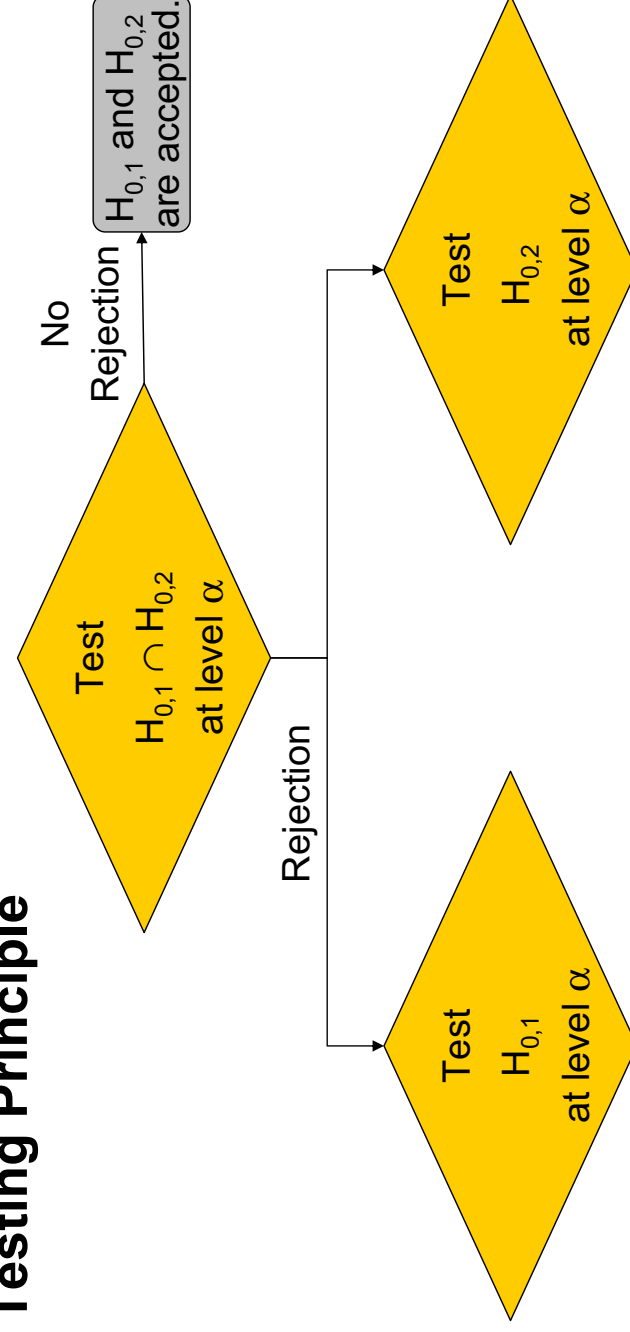
$$p = p_{12} = 1 - [1 - \min(p_1, p_2)]^2$$

- At Stage 2 use the p-value for the selected dose(s):
 - If we select only ONE dose, say dose 1, we use $q = q_1$
 - If we select BOTH doses we use e.g. Šidak test

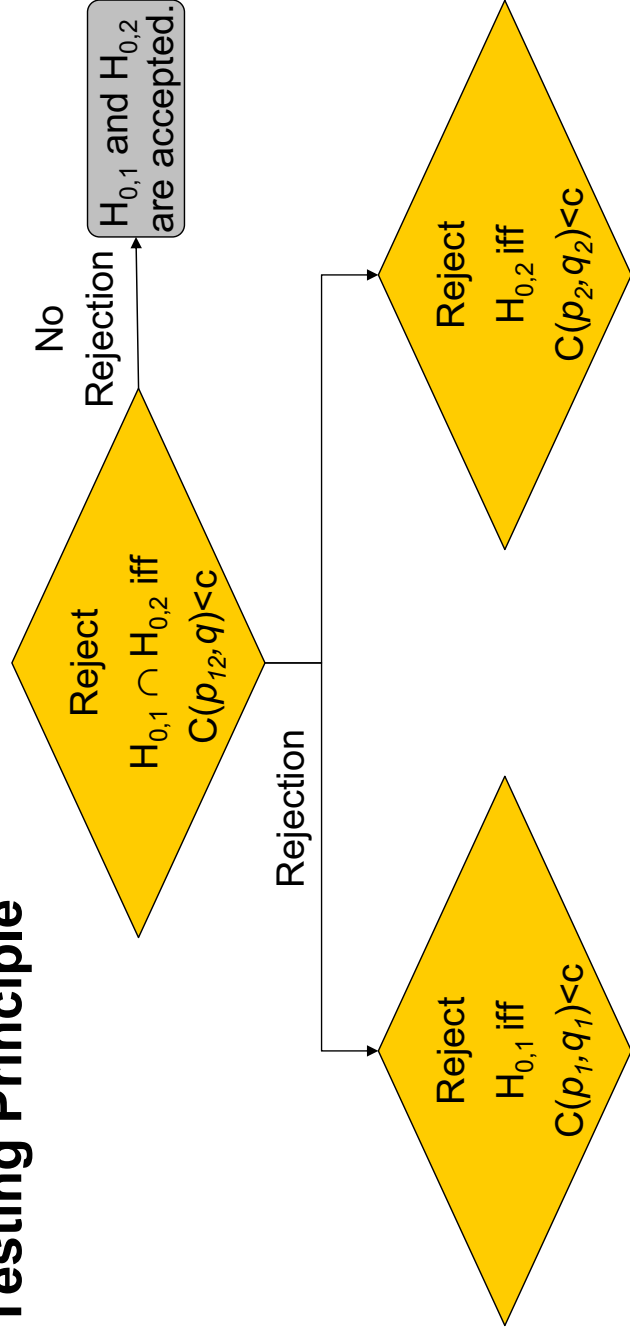
$$q = q_{12} = 1 - [1 - \min(q_1, q_2)]^2$$

- In both cases reject $H_{0,1} \cap H_{0,2}$ iff $C(p, q) \leq c$.

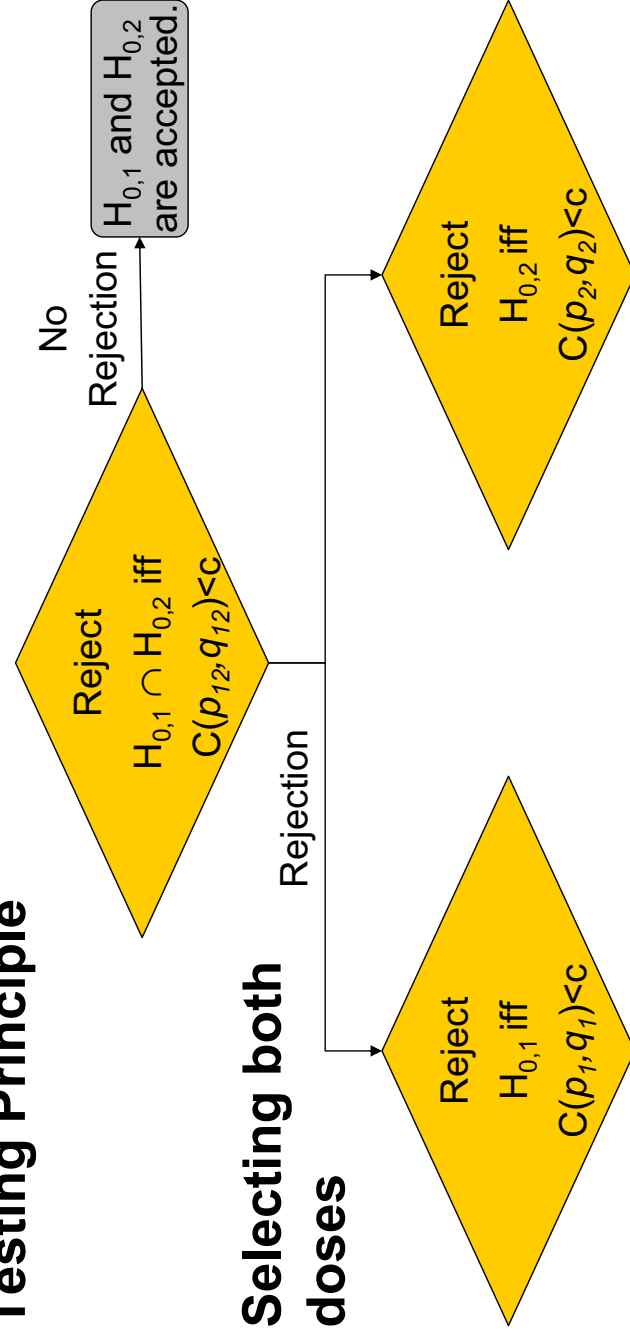
The Closed Testing Principle



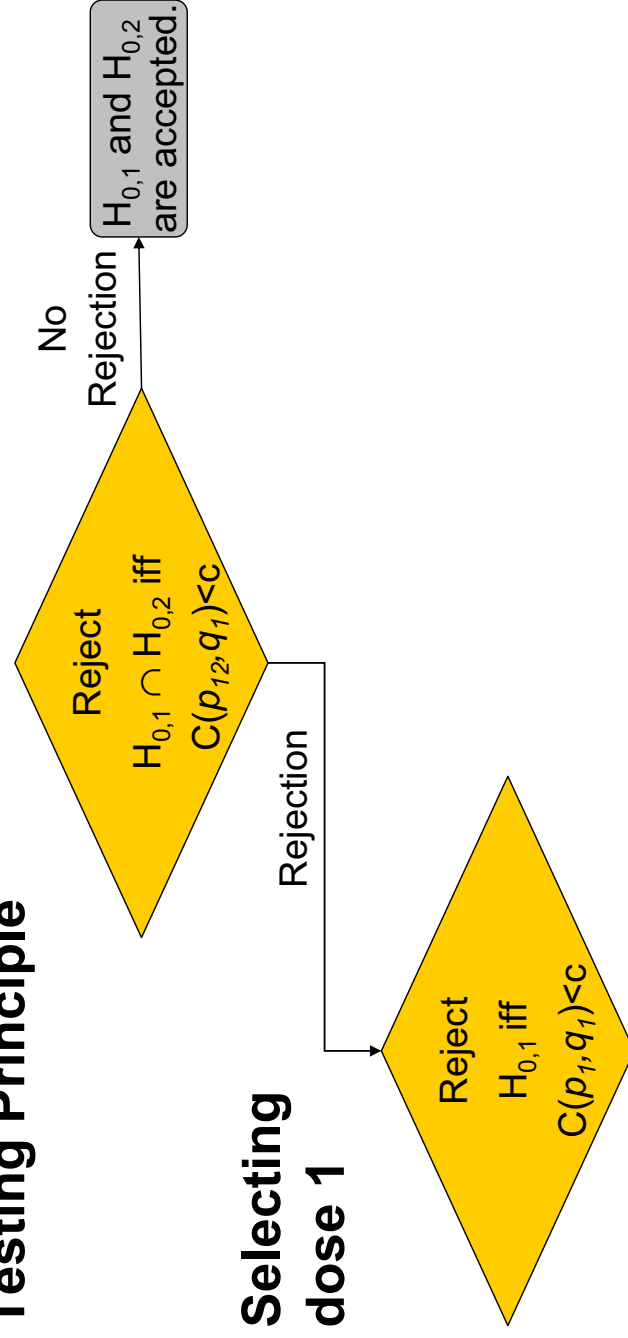
Adaptive Closed Testing Principle



Adaptive Closed Testing Principle



Adaptive Closed Testing Principle



Some Comments

- The procedure controls the familywise error rate
- Applying the closed testing approach, any number of treatments can be continued to the second stage and new treatments can be added.
- Instead of the Šidak adjusted p-values one can use
 - Dunnett adjusted p-values
 - Trend tests
 - Likelihood ratio tests
 - Simes tests
- Other adaptations can be performed: Endpoints, Population...
- Flexible closed tests can be based on the conditional error function, e.g., Adaptive Dunnett tests for treatment selection (KÖNIG ET AL, 2007).

Confidence Intervals

Confidence Intervals

- Assume an adaptive test of the one sided hypotheses

$$H_i : \mu_j - \mu_0 \leq 0 \quad \text{against} \quad H'_i : \mu_j - \mu_0 > 0, \quad i = 1, 2.$$

has been performed.

- For all parameter vectors $\theta = (\theta_1, \theta_2)$ define

$$H_i(\theta_j) : \mu_j - \mu_0 \leq \theta_j \quad \text{against} \quad H'_i(\theta_j) : \mu_j - \mu_0 > \theta_j, \quad i = 1, 2,$$

and let $H_{12}(\theta)$ denote the corresponding intersection hypothesis.

- For each θ denote the first stage p-value for $H_{12}(\theta)$ by $p_{12}(\theta)$.

A Confidence Region

Assume, that treatment 1 is selected in the interim analysis.

- Let $q_1(\theta_1)$ denote the second stage p-value for $H_1(\theta_1)$ which is also a p-value for $H_{12}(\theta)$.
- A $100\%(1 - \alpha)$ confidence region for $\mu_i - \mu_0$, $i = 1, 2$ is given by all vectors θ such that

$$C(p_{12}(\theta), q_1(\theta_1)) \geq c$$

- In general, this is not a cross product of confidence intervals!
- We need to embed the confidence region in a rectangle.

Adjusted p-values

- Define for each hypothesis stage wise adjusted p-values

$$p_i^{\text{adj}}(\theta_i) = \sup_{\xi \in \mathbb{R}^2, \xi_j \leq \theta_j} p_{12}(\xi).$$

- e.g., for the Šidak test

$$p_i^{\text{adj}}(\theta_i) = 1 - [1 - p_i(\theta_i)]^2.$$

- If, e.g., only Treatment 1 is continued

$$q_1^{\text{adj}}(\theta_1) = q_1(\theta_1), \quad \text{and} \quad q_2^{\text{adj}}(\theta_2) = 1.$$

Simultaneous adaptive Confidence Bounds

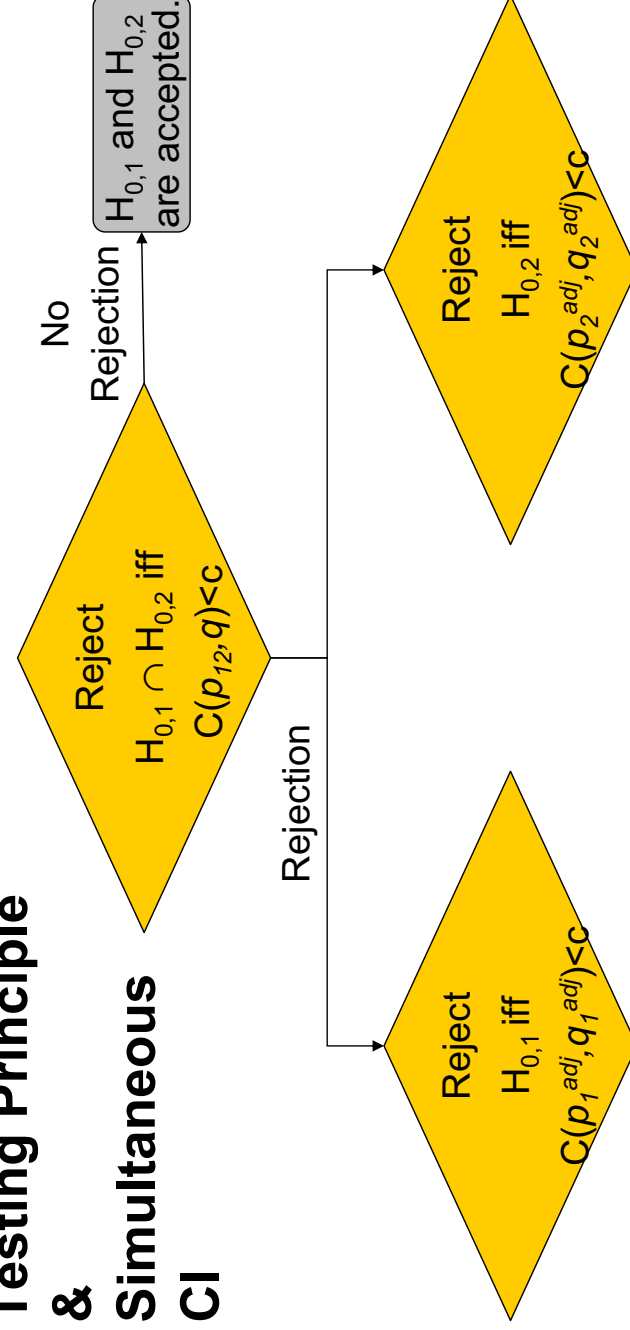
The one-sided confidence intervals for θ_i , $i = 1, 2$ are given by

$$I_i = \{\theta_i \mid C(p_i^{\text{adj}}(\theta_i), q_i^{\text{adj}}(\theta_i)) \geq c\}.$$

POSCH ET AL. 2005

- The confidence bound may be inconsistent with the test decision.
- Note that this is not specific to flexible designs but also arises in stepwise multiple testing problems in fixed sample designs.
- To make the test decision consistent with the confidence bound: we could use the adjusted p-values $p_i^{\text{adj}}(\theta_i = 0)$ and $q_i^{\text{adj}}(\theta_i = 0)$ also in flexible closed testing procedure.

Adaptive Closed Testing Principle & Simultaneous CI



Comments

- The confidence intervals are in general strictly conservative.
- The 50% confidence bounds are conservative point estimates (with non-positive median bias).

Point Estimation

Point Estimation

- 1 Univariate Estimates of all Treatments
- 2 Estimate of the Selected Treatment

Univariate Estimates

- 1st stage means: $\bar{X}_0, \bar{X}_1, \bar{X}_2$
- 2nd stage means: $\bar{Y}_0, \bar{Y}_1, \bar{Y}_2$
- 2nd stage sample size:

$$\tilde{n}_{2i} = \begin{cases} n_2 & \text{if treatment } i \text{ is selected} \\ 0 & \text{otherwise} \end{cases}$$

- Consider for each treatment the overall mean.

$$\bar{Z}_j = \tilde{t}_j \bar{X}_j + (1 - \tilde{t}_j) \bar{Y}_j,$$

where $\tilde{t}_j = n_1 / (n_1 + \tilde{n}_2)$.

- The Bias and Mean Squared Error (MSE) depend on the treatment selection rule and on $\mu_1 - \mu_2$.

Univariate Bias and MSE

- For the bias we have

$$E(\bar{Z}_j - \mu_j) = E[\tilde{t}_j(\bar{X}_j - \mu_j)] < 0.4\sigma / \sqrt{n_1}$$

(Brannath et al. 2005)

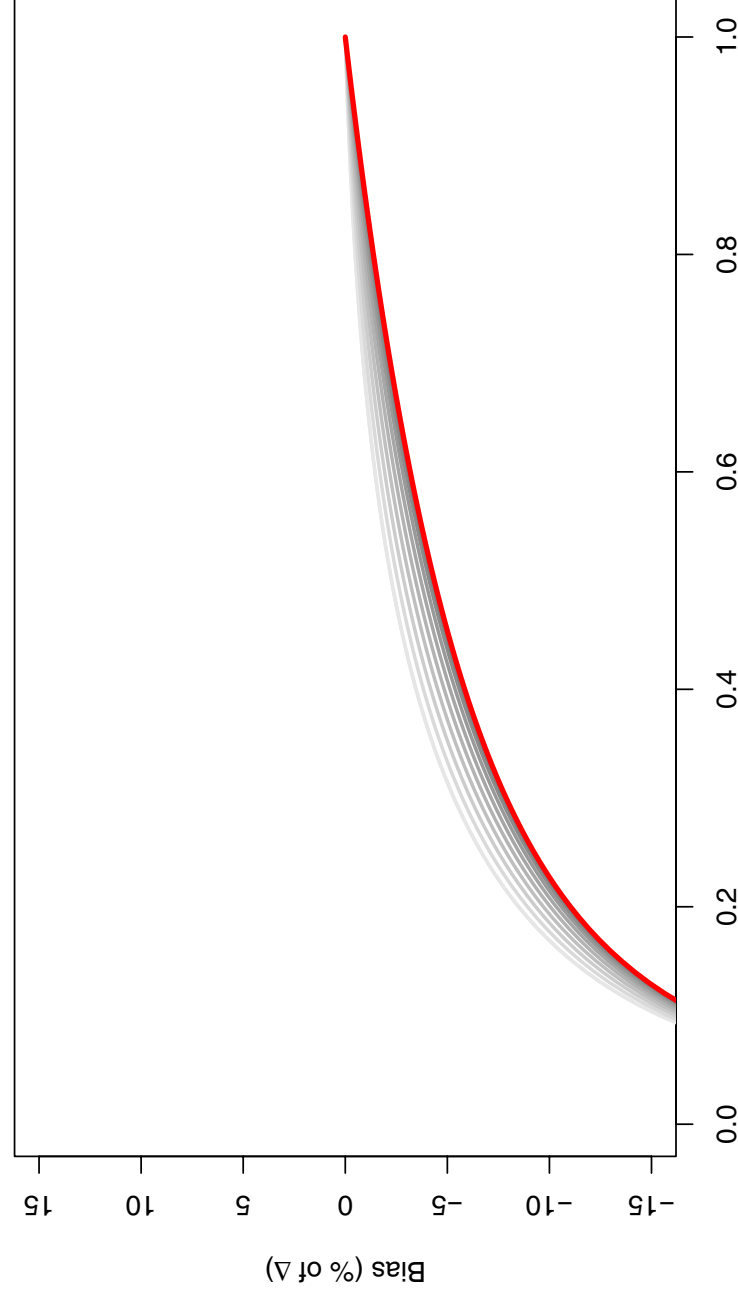
- For the rule that selects the treatment with the highest interim effect estimate the bias is given by

$$E(Z_1 - \mu_1) = -\frac{\sigma}{2\sqrt{\pi n_1}}(1 - t) \exp^{-n_1(\mu_1 - \mu_2)^2 / (4\sigma^2)},$$

where $t = n_1 / (n_1 + n_2)$.

Univariate Bias of the Effect Estimate for μ_1

$$\mu_1 - \mu_0 = 1.0\Delta, \quad \mu_2 - \mu_0 = \Delta$$



r : Proportion of Total Sample Size in Pilot Study

Estimate of the Selected Treatment

- Typically, at the end of the trial one is only interested in the selected treatment
- Let $S \in \{1, 2\}$ denote the index of the selected treatment. Define the selection bias:

$$E[(\bar{Z}_S - \mu_S)]$$

and the selection mean squared error by

$$E[(\bar{Z}_S - \mu_S)^2]$$

Selection Bias and MSE

Consider the rule that selects the treatment with the highest interim effect estimate in the interim analysis.

- The selection bias is given by

$$E(Z_S - \mu_S) = \frac{\sigma t}{\sqrt{\pi n_1}} \exp[-n_1(\mu_1 - \mu_2)^2 / (4\sigma^2)].$$

It has opposite sign than the the univariate bias!

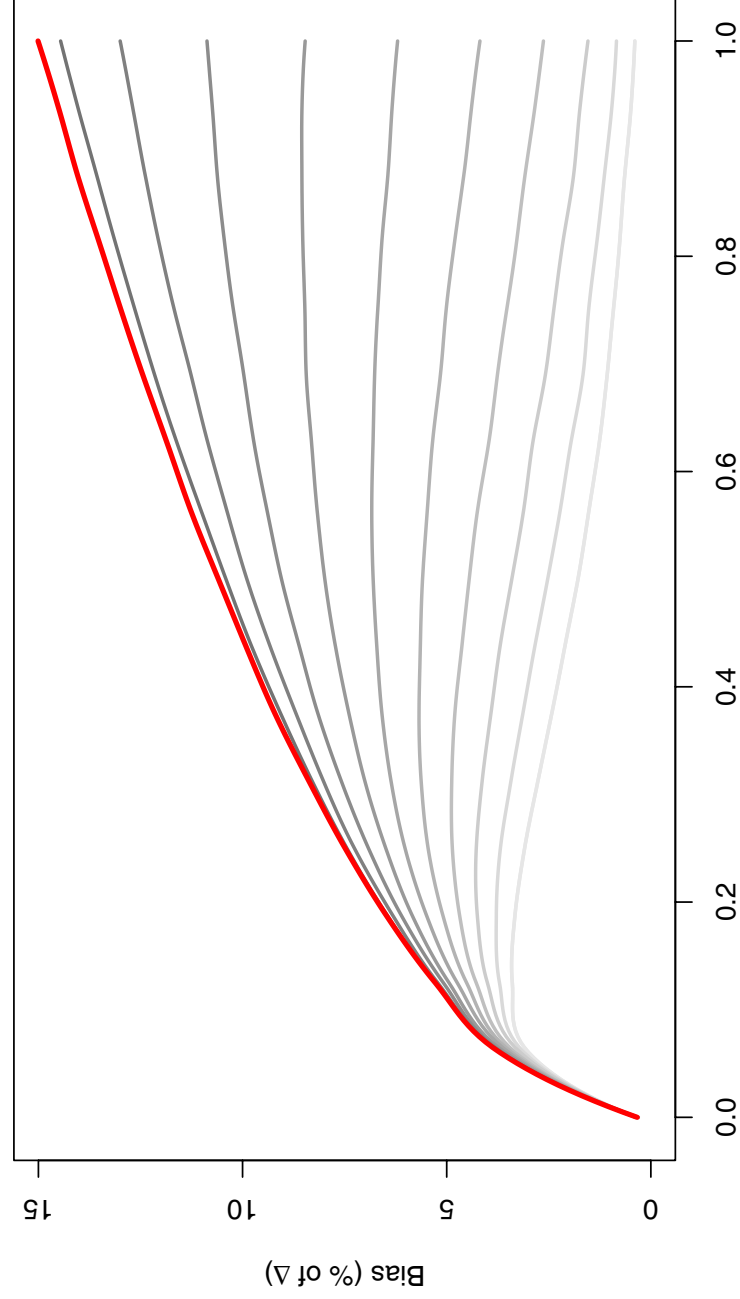
- The selection mean squared error is

$$MSE_{sel} = \frac{\sigma^2}{n}.$$

- One can show (POSCH ET AL. 2005) that for all selection rules that depend on the interim mean difference $\bar{X}_1 - \bar{X}_2$ the selection mean squared error is σ^2/n .

Bias of the effect estimate for the selected Treatment

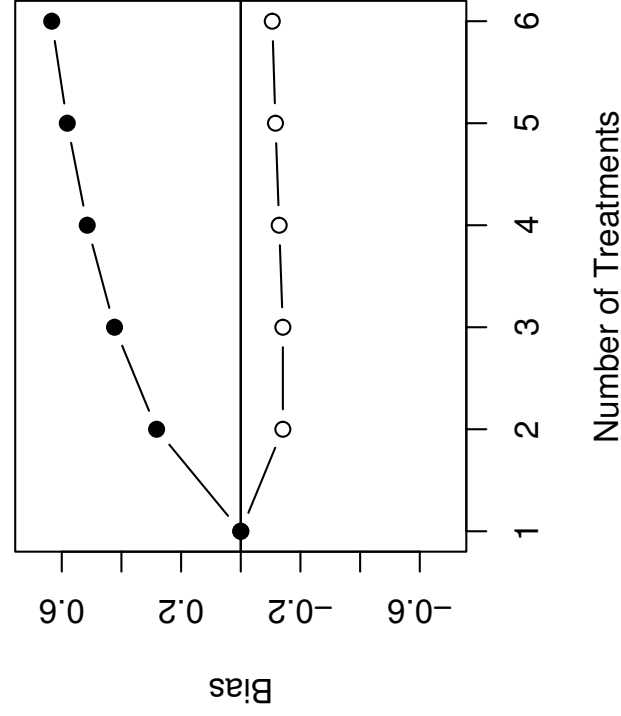
$$\mu_1 - \mu_0 = 1.0\Delta, \quad \mu_2 - \mu_0 = \Delta$$



r: Proportion of Total Sample Size in Pilot Study

Bias: More than two treatments

Example: All μ_j equal. Select the treatment with the highest interim estimate, $n_1 = n_2$. Bias in units of $\sigma / \sqrt{n_1}$.



○ univariate bias ● selection bias

Summary & Comments

- \bar{Z}_I : negative bias
- \bar{Z}_S : positive bias
- For fixed selection rule “select the treatment with highest interim effect”:
 - Unbiased estimate for μ_S (Cohen and Sackowitz, 1989). Also unbiased conditional on the treatment selected. MSE larger than for the “naive” estimate.
 - Bias adjusted estimates for μ_S STALLARD & TODD, 2005, SHEN, 2001.

Selected References I

- Bauer P, Kieser M. Combining different phases in the development of medical treatments within a single trial. *Statistics in Medicine* 1999; **18**:1833–1848.
- Brannath W., König F, Bauer P. Multiplicity and flexibility in clinical trials. *Submitted*.
- Bretz F, Schmidli H, König F, Racine A, Maurer W. Confirmatory seamless phase II/III clinical trials with hypotheses selection at interim: General concepts. *Biometrical Journal* 2006; **48**(4): 623-34.
- Hommel G. Adaptive modifications of hypotheses after an interim analysis. *Biometrical Journal* 2001; **43**:581–589.
- König, F., Brannath, W., Bretz, F., Posch, M. Adaptive Dunnett tests for treatment selection. *Statistics in Medicine*.
- Müller HH, Schäfer H. Adaptive group sequential designs for clinical trials: Combining the advantages of adaptive and of classical group sequential approaches. *Biometrics* 2001; **57**:886–891.
- Posch M, König F, Branson M, Brannath W, Dunger-Baldauf C, Bauer P. Testing and estimation in flexible group sequential designs with adaptive treatment selection. *Statistics in Medicine* 2005; **30**:3697-3714.