Multiple Type-I Error Control in Response Adaptive Phase II/III Designs with Treatment Selection

Georg Gutjahr, Martin Posch, and Werner Brannath

Medical University of Vienna
Department of Medical Statistics

March 12, 2010

Research supported by FWF P21763

Topic of this talk

Two stage design for comparing multiple treatments against a single control (e.g., Berry et al., 2001)

1. Response-adaptive randomization in stage 1
2. Treatment selection between stage 1 and stage 2
3. Block randomization between
   selected treatment and control in stage 2

Controversial: selection of the critical values by simulation under global null hypothesis (Posch, Maurer, and Bretz, 2010)

Problem: control of the familywise error rate in the strong sense

Outline

- The design

  A multi-armed trial ...

  ... with treatment selection ...

  ... and response-adaptive allocation

- Test that guarantees strong control of the FWE rate
- Simulation results

Setup

Comparison of two treatments versus a single control.

Treatment A	Responses $\sim$ Normal $(\mu + \delta_A, 1)$
Treatment B	Responses $\sim$ Normal $(\mu + \delta_B, 1)$
Control C	Responses $\sim$ Normal $(\mu, 1)$

SETUP

Comparison of two treatments versus a single control.

Treatment A	Responses $\sim$ Normal $(\mu + \delta_A, 1)$
Treatment B	Responses $\sim$ Normal $(\mu + \delta_B, 1)$
Control C	Responses $\sim$ Normal $(\mu, 1)$

Closure principle

$$\begin{array}{ccc} & H_{\mathrm{AB}} & \\ \diagup & & \diagdown \\ H_{\mathrm{A}} & & H_{\mathrm{B}} \end{array}$$

SETUP

Comparison of two treatments versus a single control.

Treatment A	Responses $\sim$ Normal $(\mu + \cancel{\delta_A}, 1)$
Treatment B	Responses $\sim$ Normal $(\mu + \cancel{\delta_B}, 1)$
Control C	Responses $\sim$ Normal $(\mu, 1)$

Closure principle

$$\begin{array}{ccc} & \boxed{H_{\mathrm{AB}}} & \\ / & & \backslash \\ H_{\mathrm{A}} & & H_{\mathrm{B}} \end{array}$$

Setup

Comparison of two treatments versus a single control.

Treatment A	Responses $\sim$ Normal $(\mu + \cancel{\delta_A}, 1)$
Treatment B	Responses $\sim$ Normal $(\mu + \delta_B, 1)$
Control C	Responses $\sim$ Normal $(\mu, 1)$

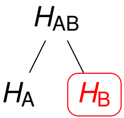
Closure principle

$$\begin{tikzcd} & H_{\mathrm{AB}} \arrow[dl, dash] \arrow[dr, dash] & \\ \boxed{H_{\mathrm{A}}} & & H_{\mathrm{B}} \end{tikzcd}$$

SETUP

Comparison of two treatments versus a single control.

Treatment A	Responses $\sim$ Normal $(\mu + \delta_A, 1)$
Treatment B	Responses $\sim$ Normal $(\mu + \cancel{\delta_B}, 1)$
Control C	Responses $\sim$ Normal $(\mu, 1)$



Outline

- The design

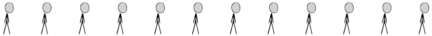
  A multi-armed trial ...

  ... with treatment selection ...

  ... and response-adaptive allocation

- Test that guarantees strong control of the FWE rate
- Simulation results

Multi-armed trial with fixed sample sizes

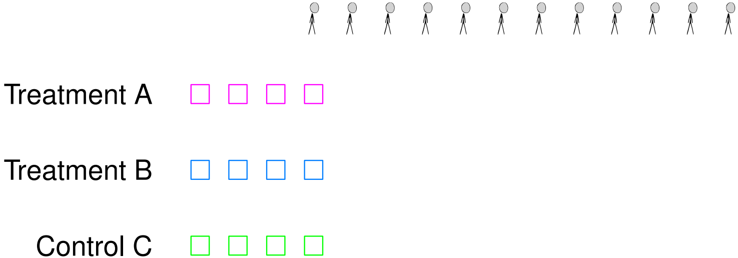


Treatment A

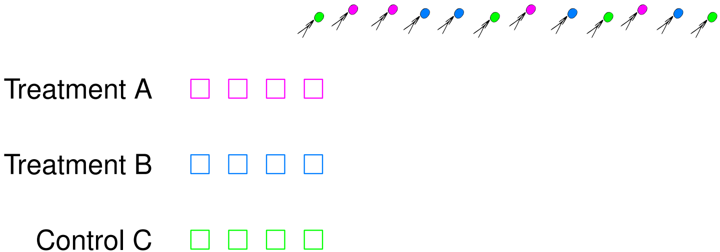
Treatment B

Control C

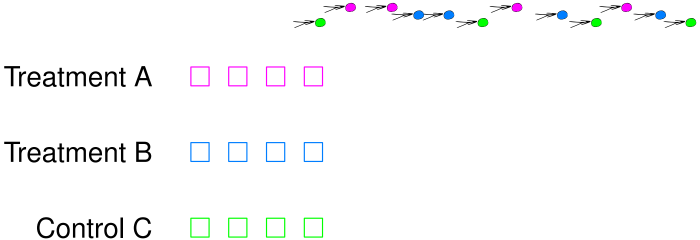
Multi-armed trial with fixed sample sizes



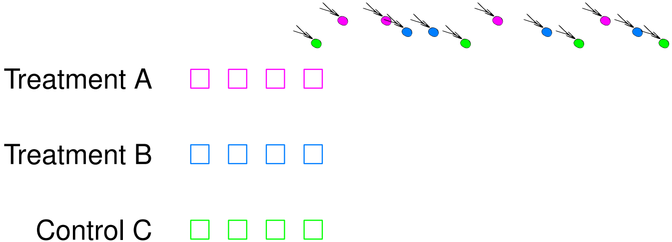
Multi-armed trial with fixed sample sizes



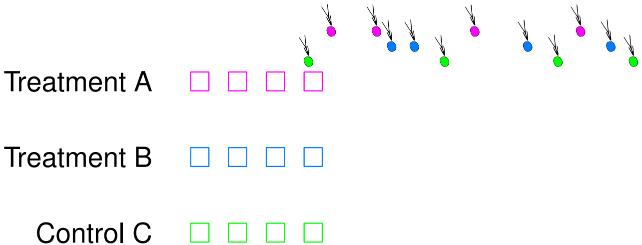
Multi-armed trial with fixed sample sizes



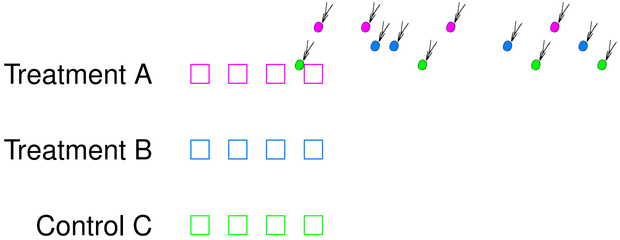
Multi-armed trial with fixed sample sizes



MULTI-ARMED TRIAL WITH FIXED SAMPLE SIZES



Multi-armed trial with fixed sample sizes

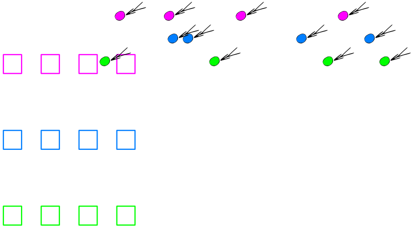


Multi-armed trial with fixed sample sizes

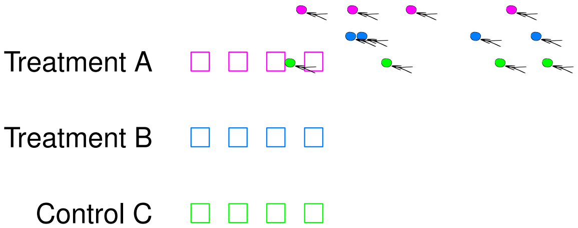
Treatment A

Treatment B

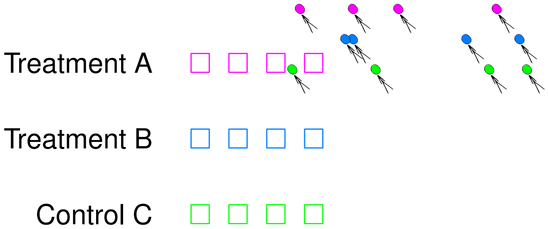
Control C



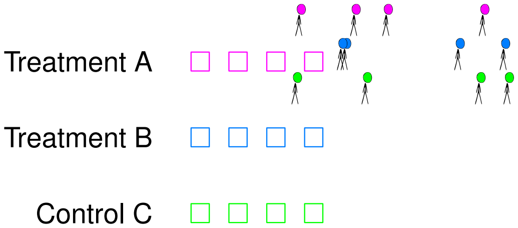
Multi-armed trial with fixed sample sizes



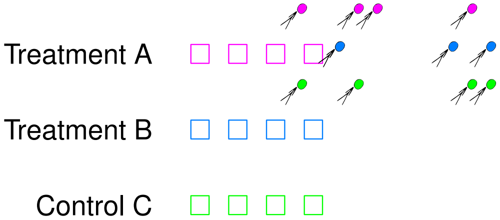
Multi-armed trial with fixed sample sizes



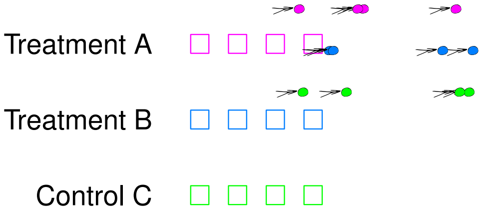
MULTI-ARMED TRIAL WITH FIXED SAMPLE SIZES



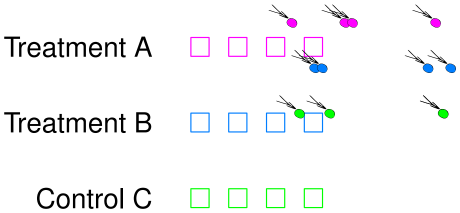
MULTI-ARMED TRIAL WITH FIXED SAMPLE SIZES



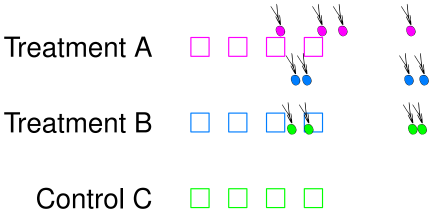
Multi-armed trial with fixed sample sizes



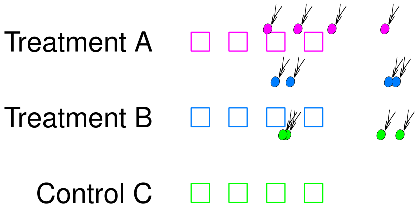
Multi-armed trial with fixed sample sizes



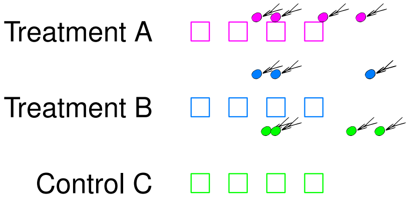
Multi-armed trial with fixed sample sizes



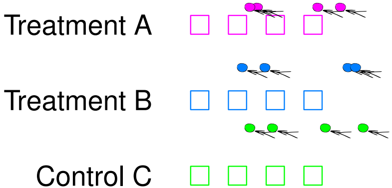
Multi-armed trial with fixed sample sizes



Multi-armed trial with fixed sample sizes



MULTI-ARMED TRIAL WITH FIXED SAMPLE SIZES

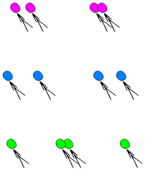


Multi-armed trial with fixed sample sizes

Treatment A

Treatment B

Control C

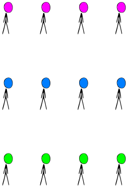


MULTI-ARMED TRIAL WITH FIXED SAMPLE SIZES

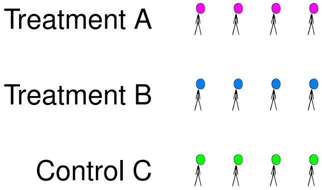
Treatment A

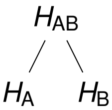
Treatment B

Control C



MULTI-ARMED TRIAL WITH FIXED SAMPLE SIZES



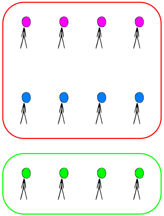


Multi-armed trial with fixed sample sizes

Treatment A

Treatment B

Control C

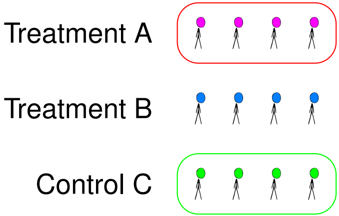


Pool observations

H_{AB}

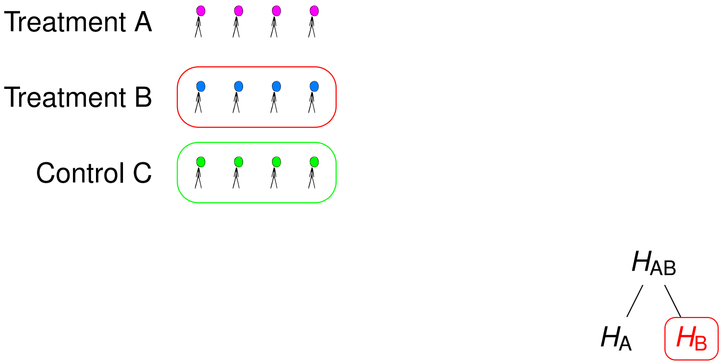
H_A H_B

MULTI-ARMED TRIAL WITH FIXED SAMPLE SIZES



$$\begin{array}{ccc} & H_{AB} & \\ \boxed{H_A} & & H_B \end{array}$$

Multi-armed trial with fixed sample sizes



Outline

- The design

  A multi-armed trial . . .

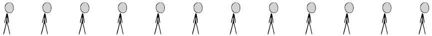
  . . . with treatment selection . . .

  . . . and response-adaptive allocation

- Test that guarantees strong control of the FWE rate

- Simulation results

Design with treatment selection

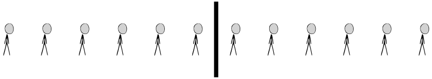


Treatment A

Treatment B

Control C

Design with treatment selection

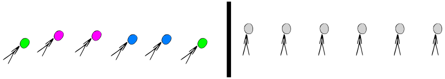


Treatment A

Treatment B

Control C

Design with treatment selection

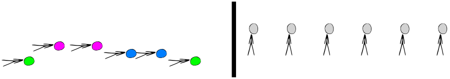


Treatment A

Treatment B

Control C

Design with treatment selection

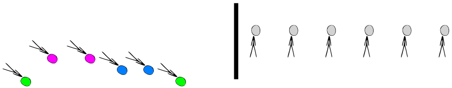


Treatment A

Treatment B

Control C

Design with treatment selection

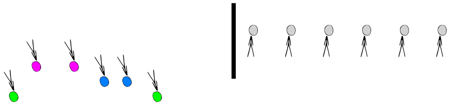


Treatment A

Treatment B

Control C

Design with treatment selection

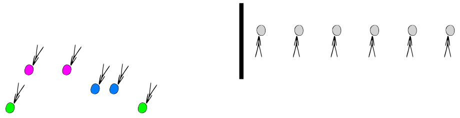


Treatment A

Treatment B

Control C

Design with treatment selection

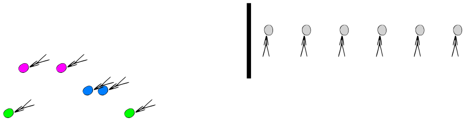


Treatment A

Treatment B

Control C

DESIGN WITH TREATMENT SELECTION



Treatment A

Treatment B

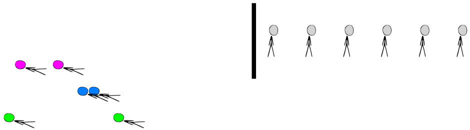
Control C

DESIGN WITH TREATMENT SELECTION

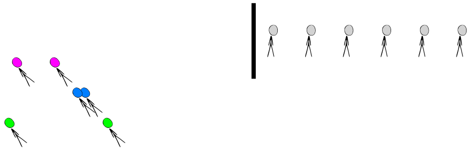
Treatment A

Treatment B

Control C



DESIGN WITH TREATMENT SELECTION

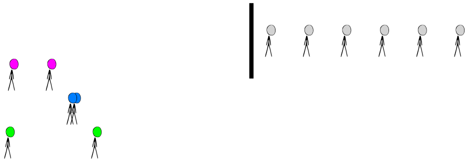


Treatment A

Treatment B

Control C

Design with treatment selection

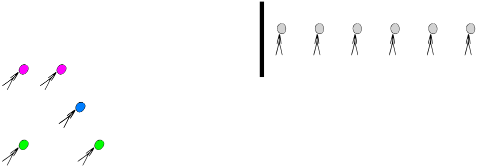


Treatment A

Treatment B

Control C

Design with treatment selection

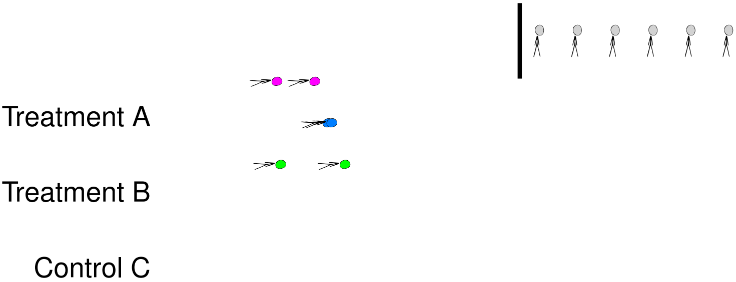


Treatment A

Treatment B

Control C

Design with treatment selection

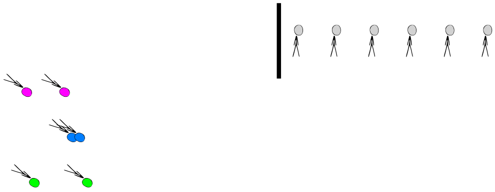


Design with treatment selection

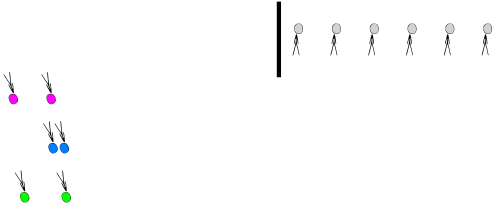
Treatment A

Treatment B

Control C



Design with treatment selection

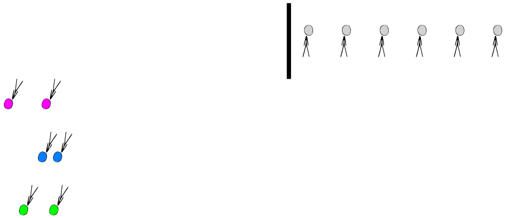


Treatment A

Treatment B

Control C

Design with treatment selection

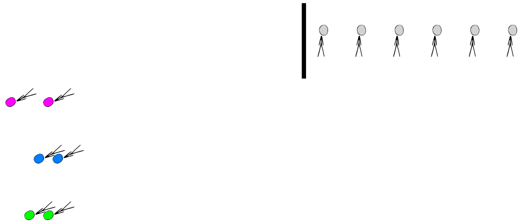


Treatment A

Treatment B

Control C

Design with treatment selection

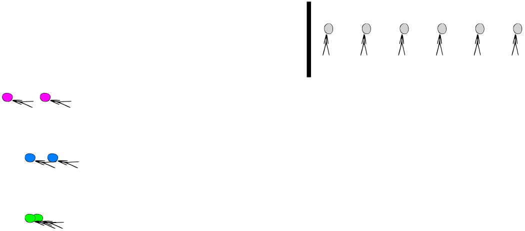


Treatment A

Treatment B

Control C

Design with treatment selection

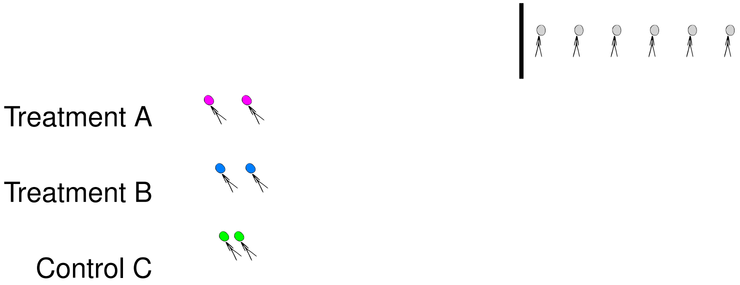


Treatment A

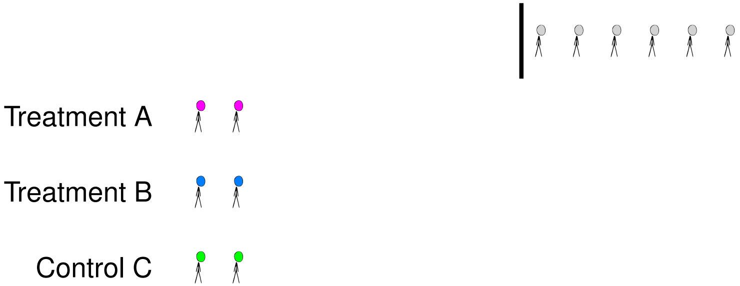
Treatment B

Control C

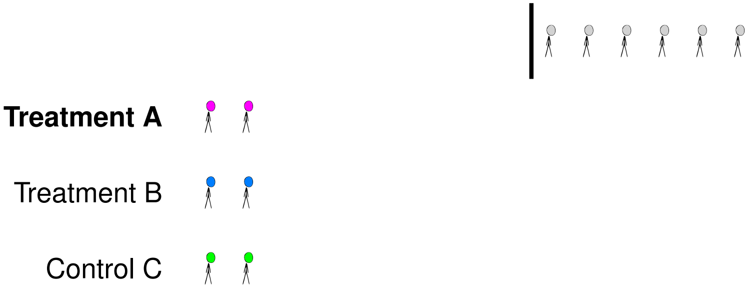
Design with treatment selection



Design with treatment selection



DESIGN WITH TREATMENT SELECTION

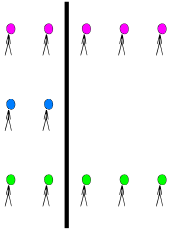


DESIGN WITH TREATMENT SELECTION

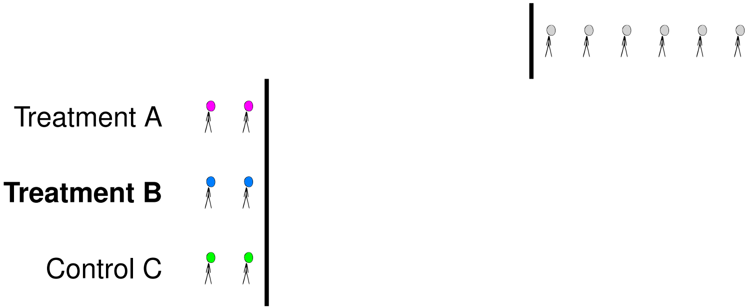
Treatment A

Treatment B

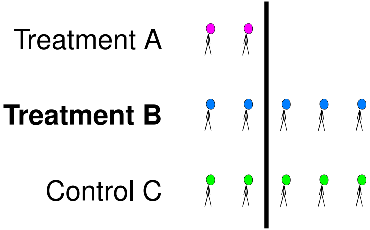
Control C



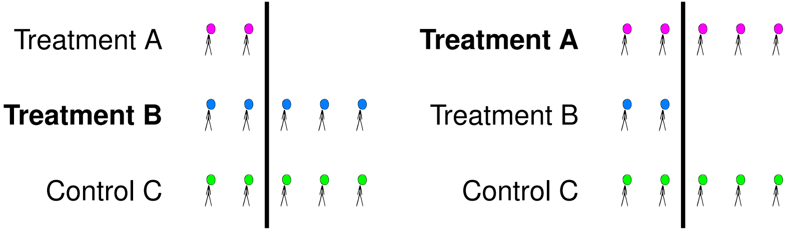
Design with treatment selection



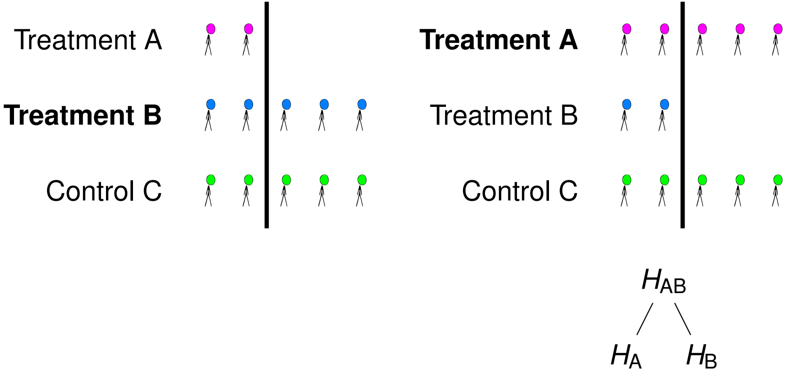
Design with treatment selection



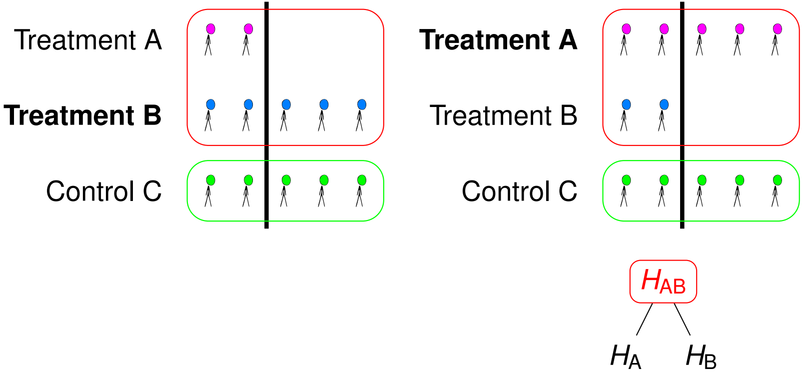
Design with treatment selection



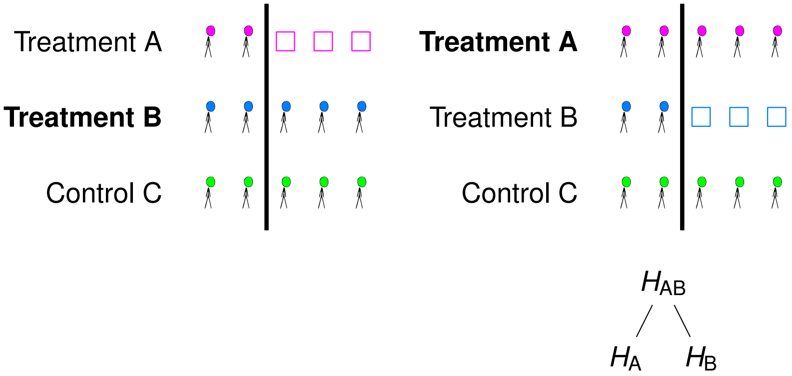
Design with treatment selection



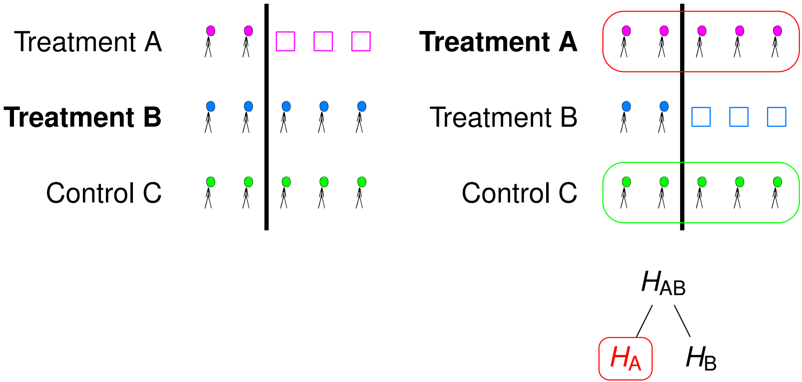
Design with treatment selection



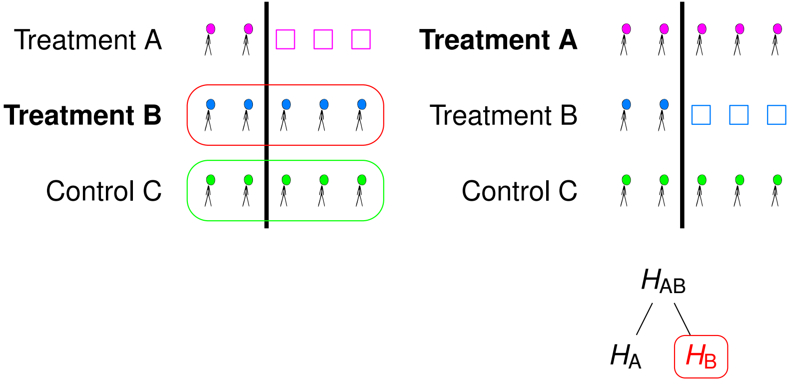
DESIGN WITH TREATMENT SELECTION



Design with treatment selection



DESIGN WITH TREATMENT SELECTION



Outline

- The design

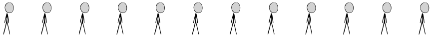
    A multi-armed trial ...

    ... with treatment selection ...

    ... and response-adaptive allocation

- Test that guarantees strong control of the FWE rate
- Simulation results

Response-adaptive allocation

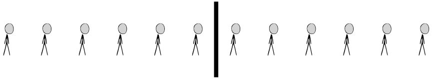


Treatment A

Treatment B

Control C

Response-adaptive allocation

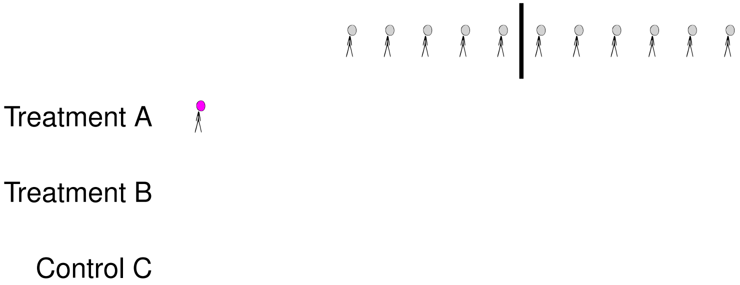


Treatment A

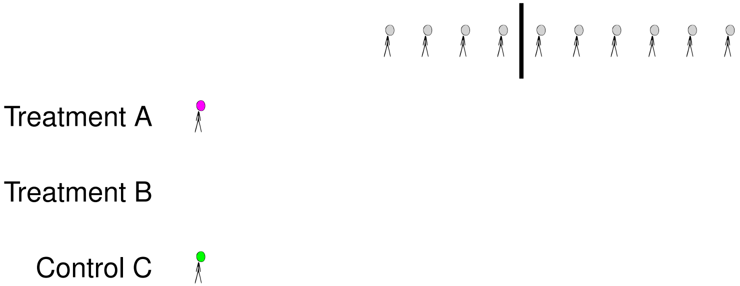
Treatment B

Control C

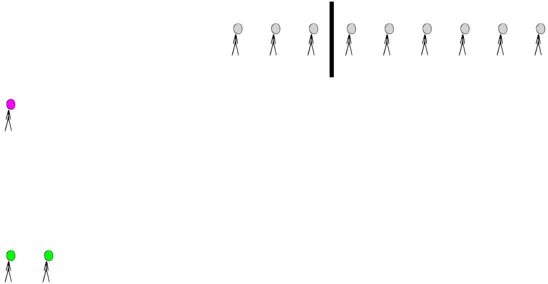
RESPONSE-ADAPTIVE ALLOCATION



RESPONSE-ADAPTIVE ALLOCATION



Response-adaptive allocation

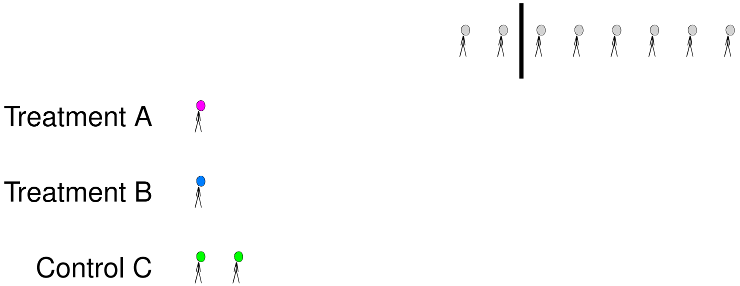


Treatment A

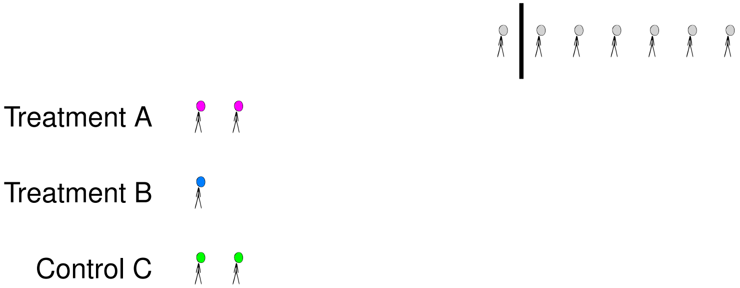
Treatment B

Control C

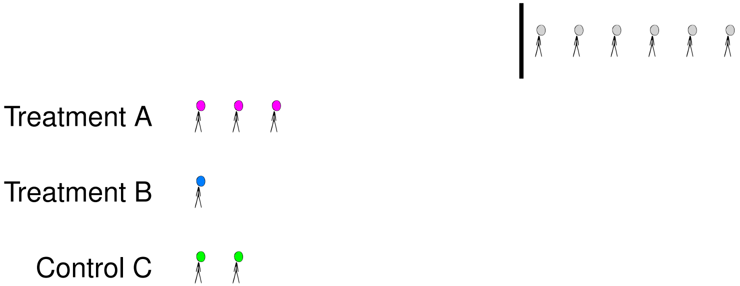
RESPONSE-ADAPTIVE ALLOCATION



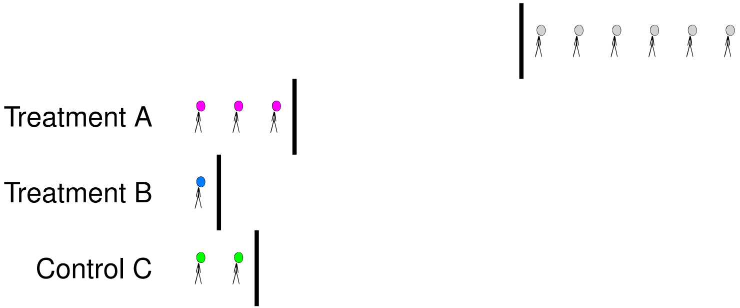
Response-adaptive allocation



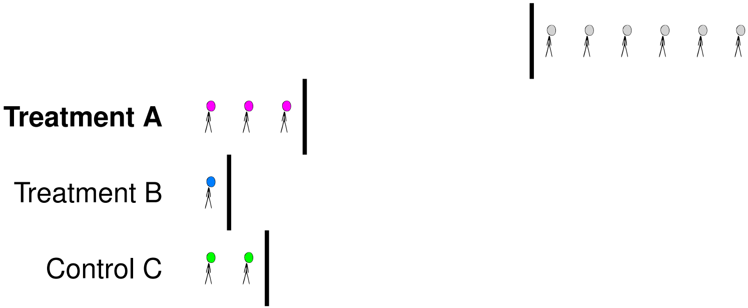
Response-adaptive allocation



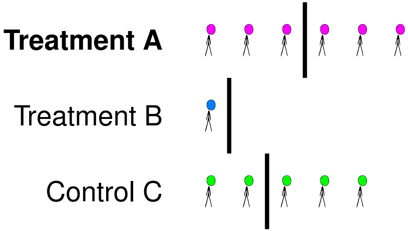
Response-adaptive allocation



Response-adaptive allocation



Response-adaptive allocation



Outline

- The design

  A multi-armed trial . . .

  . . . with treatment selection . . .

  . . . and response-adaptive allocation

- Test that guarantees strong control of the FWE rate
- Simulation results

Construction of the test (idea)

- For each hypothesis, we prespecify an hypothetical artificial auxiliary design with fixed-sample sizes.
- We view the actual response-adaptive design as a data-dependent modification of the auxiliary design.
- When we switch from stage 1 to stage 2, we change the sample sizes of the auxiliary designs by using the conditional invariance principle (Brannath et al., 2007).
- **Complication: Conditional distributions depend on nuisance parameter—unknown mean μ in control group**.

Conditional invariance principle

Consider an interim analysis during a clinical trial.

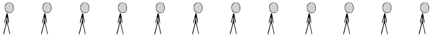
- Data D before interim analysis (first stage data).
- Test statistic $Z = X + Y$ with statistics
  X calculated from first stage data and
  Y calculated from second stage data.
- Data dependent design modification.

If we calculate from the modified second trial stage a statistic $\tilde{Y}$ so that under the null hypothesis $\tilde{Y} \,|\, D = Y \,|\, D$ in distribution, then in place of Z we can use the modified test statistic

$$\tilde{Z} = X + \tilde{Y}.$$

.

AUXILIARY DESIGN

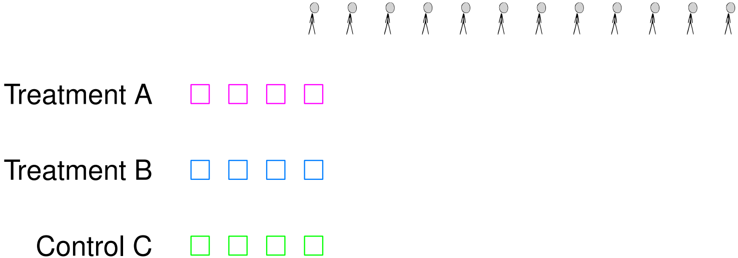


Treatment A

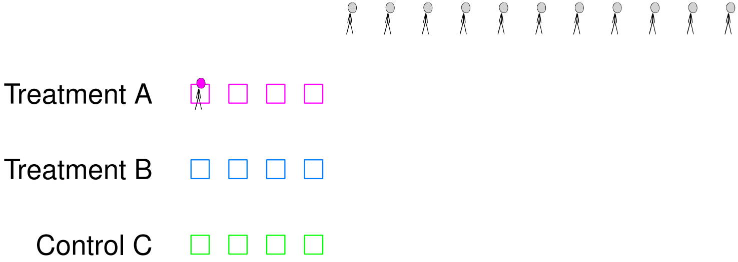
Treatment B

Control C

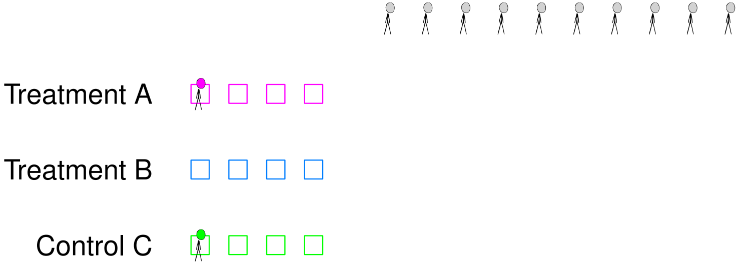
Auxiliary design



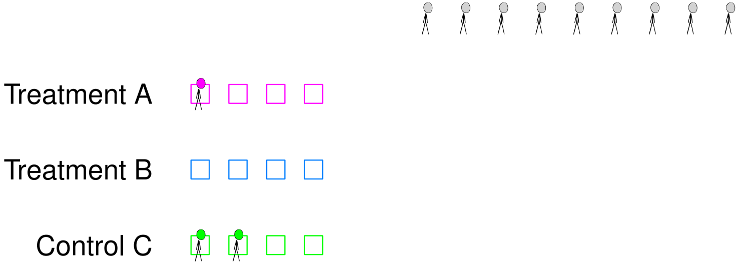
Auxiliary design



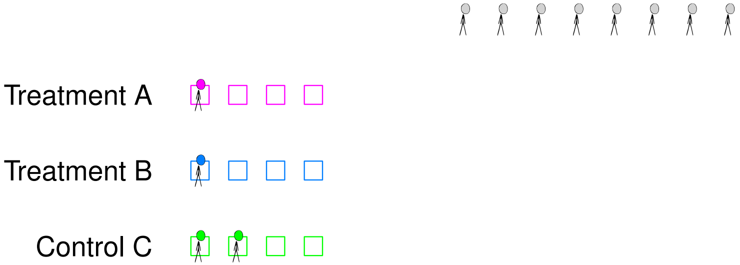
Auxiliary design



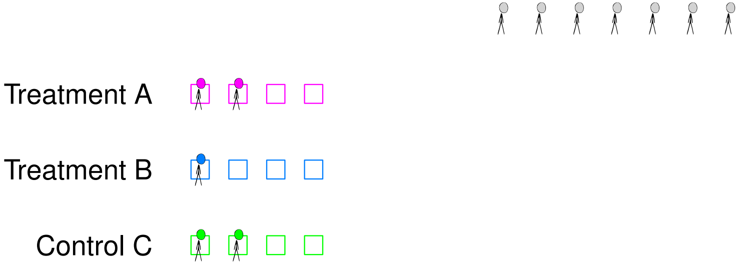
Auxiliary design



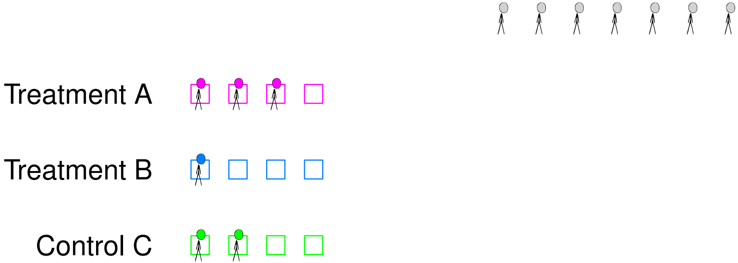
AUXILIARY DESIGN



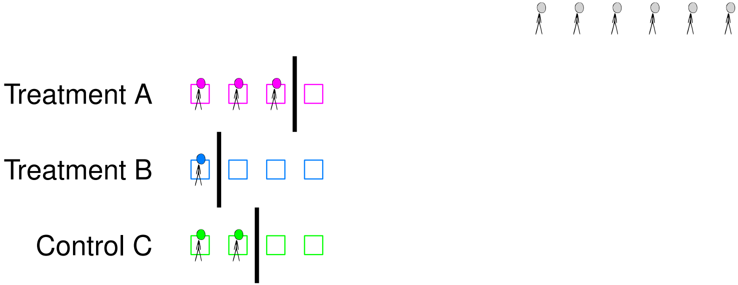
AUXILIARY DESIGN



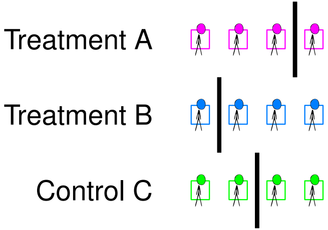
AUXILIARY DESIGN



AUXILIARY DESIGN

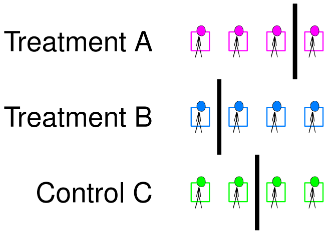


Auxiliary design



Always fill-up sample sizes

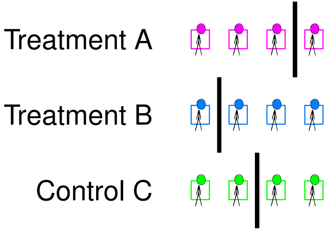
AUXILIARY DESIGN



Always fill-up sample sizes

Per-group sample sizes unchanged

Auxiliary design

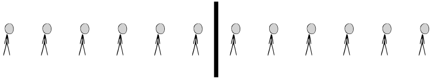


Always fill-up sample sizes

Per-group sample sizes unchanged

Order of allocation does not change test statistic

Modification of the auxiliary design

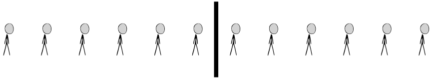


Treatment A

Treatment B

Control C

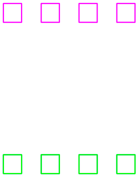
Modification of the auxiliary design



Treatment A

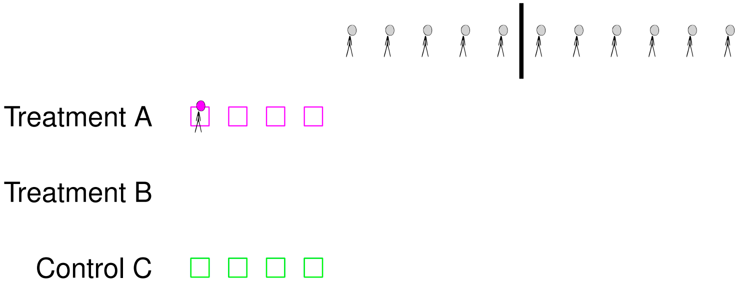
Treatment B

Control C

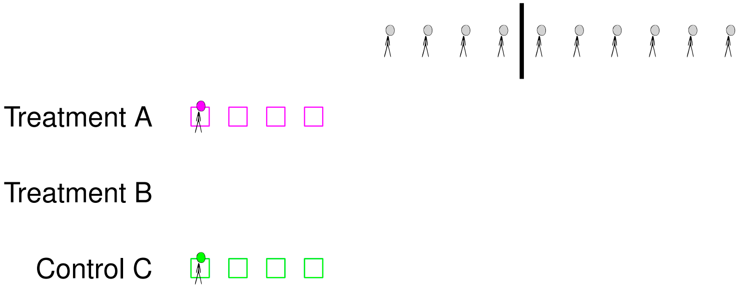


Auxiliary design for H_A

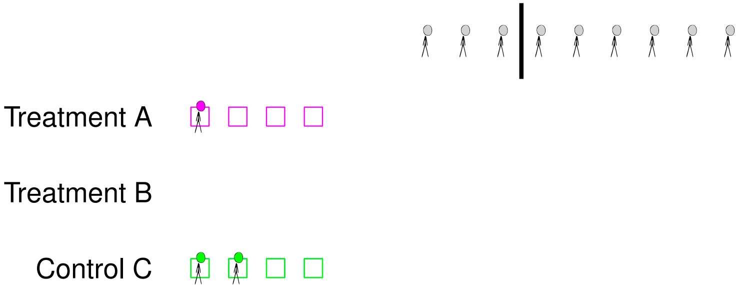
MODIFICATION OF THE AUXILIARY DESIGN



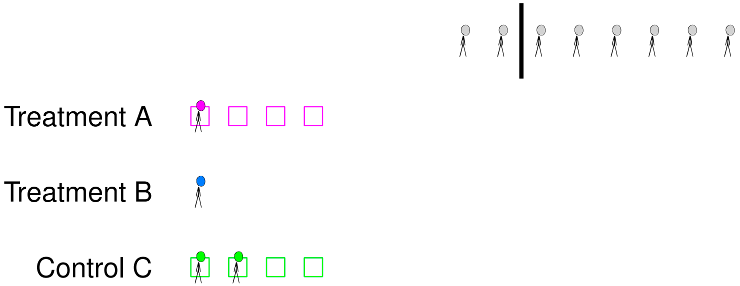
Modification of the auxiliary design



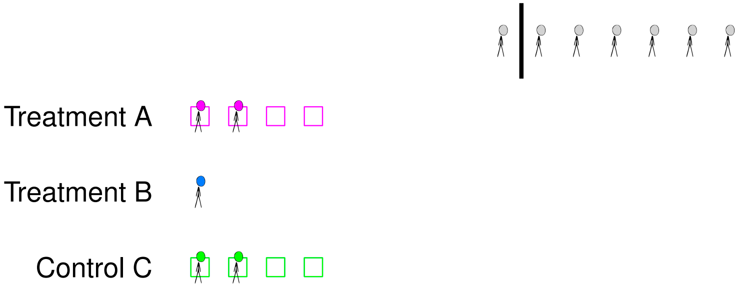
Modification of the auxiliary design



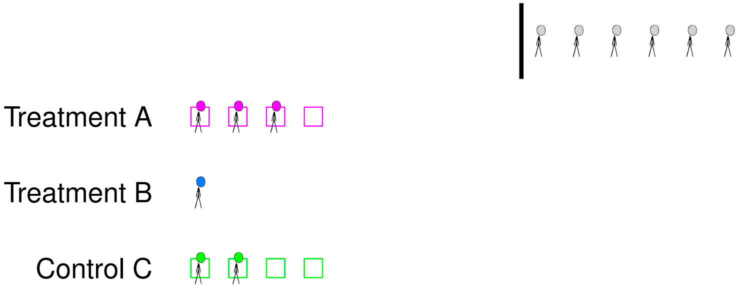
Modification of the auxiliary design



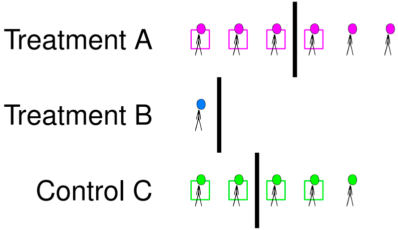
Modification of the auxiliary design



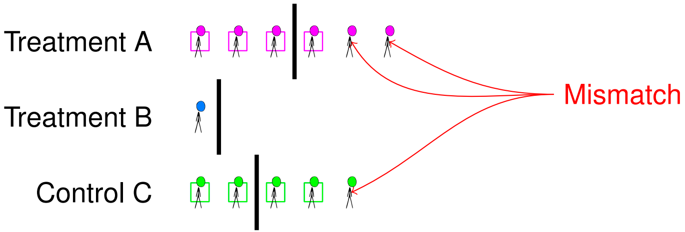
Modification of the auxiliary design



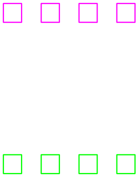
Modification of the auxiliary design



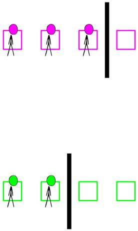
Modification of the auxiliary design



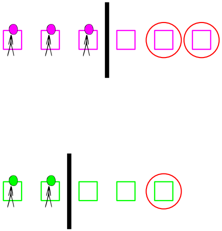
THE PROBLEM



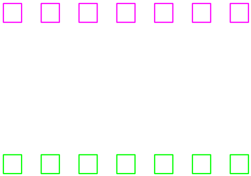
The problem



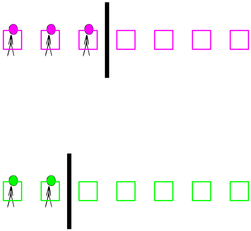
THE PROBLEM



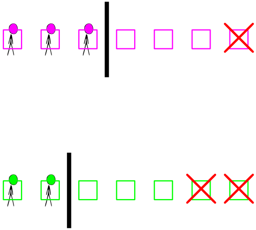
THE PROBLEM



THE PROBLEM



The problem



Test statistic in auxiliary design

Decomposition of auxiliary test statistic:

$$Z = \underbrace{\left(\frac{S_{T,1}}{w_T} - \frac{S_{C,1}}{w_C}\right)}_{X \text{ from first stage}} + \underbrace{\left(\frac{S_{T,2}}{w_T} - \frac{S_{C,2}}{w_C}\right)}_{Y \text{ from second stage}}.$$

with

- Sums S in treatment and control in first and second stage,
- Weights w that depend on auxiliary sample sizes.

Auxiliary second stage statistic

We know the conditional null distribution given interim data

$$\frac{S_{T,2}}{w_T} - \frac{S_{C,2}}{w_C} \,\Big|\, \text{Interim data} \sim \text{Normal}\left(\mu \frac{n_T}{w_T} - \mu \frac{n_C}{w_C},\ \ \frac{n_T}{w_T^2} + \frac{n_C}{w_C^2} \right)$$

with n_T and n_C the auxiliary second stage sample sizes.

Auxiliary second stage statistic

We know the conditional null distribution given interim data

$$\frac{S_{T,2}}{w_T} - \frac{S_{C,2}}{w_C} \,\Big|\, \text{Interim data} \sim \text{Normal}\left(\mu \frac{n_T}{w_T} - \mu \frac{n_C}{w_C},\;\; \frac{n_T}{w_T^2} + \frac{n_C}{w_C^2}\right)$$

with n_T and n_C the auxiliary second stage sample sizes.

Let $\tilde{n}_T$ and $\tilde{n}_C$ denote the modified second stage sample sizes, and $\tilde{S}_{T,2}$ and $\tilde{S}_{C,2}$ the corresponding sums.

Auxiliary second stage statistic

We know the conditional null distribution given interim data

$$\frac{S_{T,2}}{w_T} - \frac{S_{C,2}}{w_C} \,\Big|\, \text{Interim data} \sim \text{Normal}\left(\mu \frac{n_T}{w_T} - \mu \frac{n_C}{w_C}, \;\; \frac{n_T}{w_T^2} + \frac{n_C}{w_C^2} \right)$$

with n_T and n_C the auxiliary second stage sample sizes.

Let $\tilde{n}_T$ and $\tilde{n}_C$ denote the modified second stage sample sizes, and $\tilde{S}_{T,2}$ and $\tilde{S}_{C,2}$ the corresponding sums.

Idea: Find the right weights $\tilde{w}_T$ and $\tilde{w_C}$ to match cond distribution:

$$\frac{\tilde{S}_{T,2}}{\tilde{w}_T} - \frac{\tilde{S}_{C,2}}{\tilde{w}_C} \,\Big|\, \text{Interim data} \sim \text{Normal}\left(\mu \frac{\tilde{n}_T}{\tilde{w}_T} - \mu \frac{\tilde{n}_C}{\tilde{w}_C}, \;\; \frac{\tilde{n}_T}{\tilde{w}_T^2} + \frac{\tilde{n}_C}{\tilde{w}_C^2} \right)$$

Conditions for the weights

- To match the mean:

$$\frac{\tilde{n}_T}{\tilde{w}_T} - \frac{\tilde{n}_C}{\tilde{w}_C} = \frac{n_T}{w_T} - \frac{n_C}{w_C}$$

- To match the variance:

$$\frac{\tilde{n}_T}{\tilde{w}_T^2} + \frac{\tilde{n}_C}{\tilde{w}_C^2} = \frac{n_T}{w_T^2} + \frac{n_C}{w_C^2}$$

Solution

$$\tilde{w}_T = \frac{\tilde{n}_T + \tilde{n}_C}{n_T w_T - n_C w_C + \sqrt{\frac{\tilde{n}_C}{\tilde{n}_T}\left((\tilde{n}_T + \tilde{n}_C)(n_T w_T^2 + n_C w_C^2) - (n_T w_T - n_C w_C)^2\right)}}$$

$$\tilde{w}_C = \frac{\tilde{n}_T + \tilde{n}_C}{n_C w_C - n_T w_T + \sqrt{\frac{\tilde{n}_T}{\tilde{n}_C}\left((\tilde{n}_T + \tilde{n}_C)(n_T w_T^2 + n_C w_C^2) - (n_T w_T - n_C w_C)^2\right)}}$$

PROPERTIES

It is easy to show that

- if stage 1 is not larger than stage 2,
  then weights are always positive.

- power of the test
  does not depend on the nuisance parameter.

Outline

- The design

  A multi-armed trial . . .

  . . . with treatment selection . . .

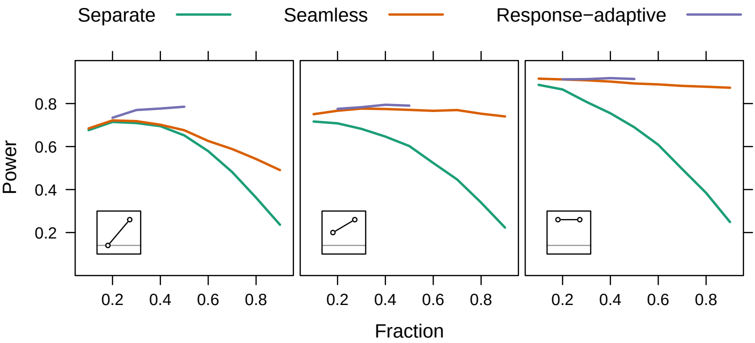
  . . . and response-adaptive allocation

- Test that guarantees strong control of the FWE rate
- Simulation results

SIMULATION

- Simulation for 2 and 4 treatments against single control
- Sample sizes chosen so that 0.8 power for most effective treatment against control with z-test
- Bayesian response-adaptive randomization rule between treatments
- Block randomization between treatments and control (no response-adaptive randomization for control)
- 10^4 replications

Two treatments against control



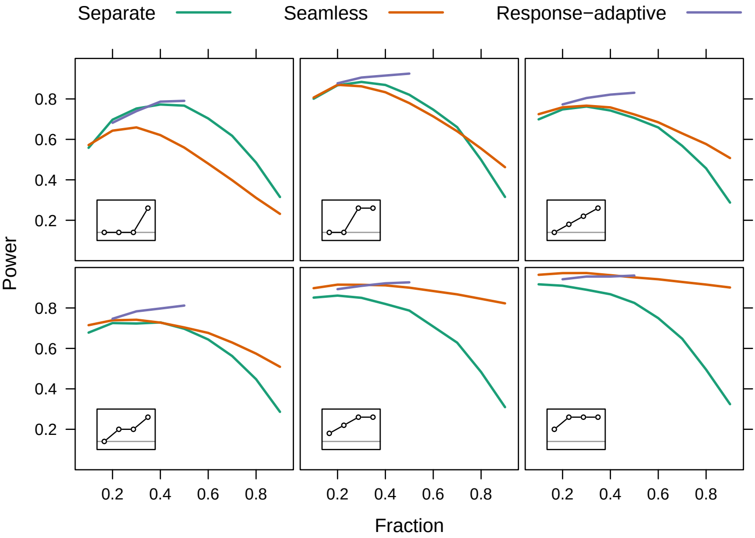
Two treatments against control

	Power		
			
Weak FWE by simulation	.82	.83	.90
Strong FWE control	.79	.83	.91

Two treatments against control

	Power		
			
Weak FWE by simulation	.82	.83	.90
Strong FWE control	.79	.83	.91
Strong FWE control & naive	.79	.83	.91

Four treatments against control



SUMMARY

The design

- Response-adaptive design with treatment selection

The test

- Closure principle—test each intersection hypothesis before rejecting an elementary hypothesis
- For each hypothesis, prespecify an auxiliary design
- Modify sample sizes of the auxiliary designs when switching from stage 1 to stage 2
- To modify sample sizes, match conditional distributions for every value of the nuisance parameter