

# A Decision Theoretic Approach to Optimize Clinical Trial Designs for Targeted Therapies

Martin Posch

Section for Medical Statistics, Medical University of Vienna

ISCB 2016

Joint work with Thomas Ondra, Sebastian Jobjörnsson, Robert Beckman, Carl-Fredrik Burman, Franz König, Nigel Stallard



This project has received funding from the European Union's 7th Framework Programme for research, technological development and demonstration under the IDEAL Grant Agreement no 602552, and the InSPiRe Grant Agreement no 602144.

## Full Population $F$

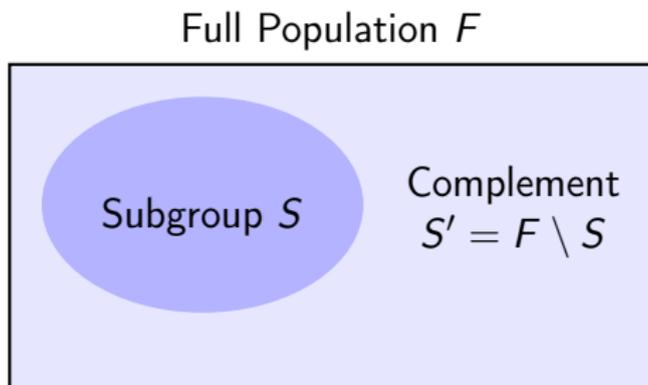


- Assume the treatment effects in the subgroups satisfy  $\delta_S \geq \delta_{S'}$ .
- The treatment effect in  $F$  is

$$\delta_F = \lambda_S \delta_S + (1 - \lambda_S) \delta_{S'}$$

where  $\lambda_S$  is the prevalence of subgroup  $S$ .

- Test of hypotheses  $H_F : \delta_F \leq 0$  and  $H_S : \delta_S \leq 0$ .

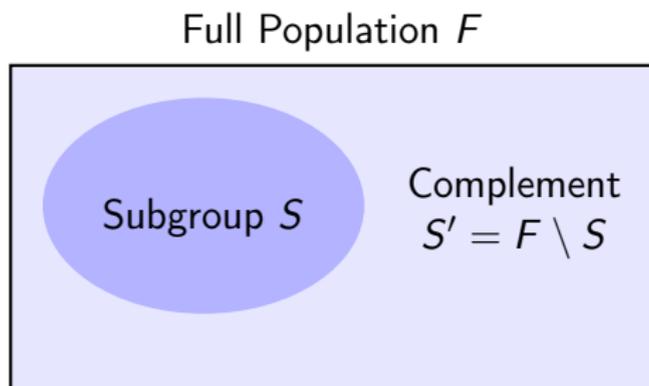


- Assume the treatment effects in the subgroups satisfy  $\delta_S \geq \delta_{S'}$ .
- The treatment effect in  $F$  is

$$\delta_F = \lambda_S \delta_S + (1 - \lambda_S) \delta_{S'}$$

where  $\lambda_S$  is the prevalence of subgroup  $S$ .

- Test of hypotheses  $H_F : \delta_F \leq 0$  and  $H_S : \delta_S \leq 0$ .

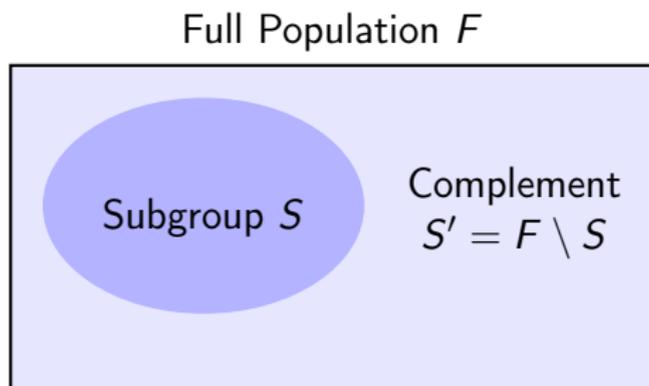


- Assume the treatment effects in the subgroups satisfy  $\delta_S \geq \delta_{S'}$ .
- The treatment effect in  $F$  is

$$\delta_F = \lambda_S \delta_S + (1 - \lambda_S) \delta_{S'}$$

where  $\lambda_S$  is the prevalence of subgroup  $S$ .

- Test of hypotheses  $H_F : \delta_F \leq 0$  and  $H_S : \delta_S \leq 0$ .



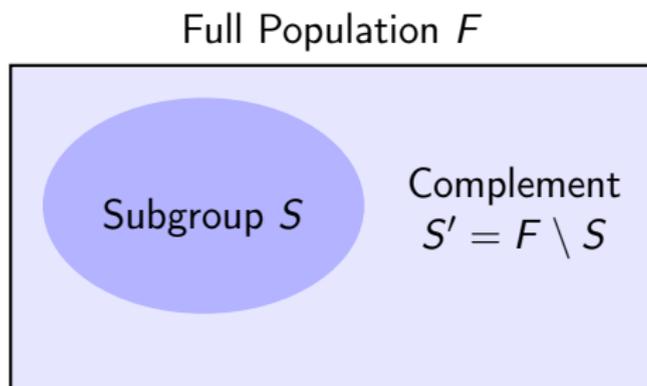
- Assume the treatment effects in the subgroups satisfy  $\delta_S \geq \delta_{S'}$ .
- The treatment effect in  $F$  is

$$\delta_F = \lambda_S \delta_S + (1 - \lambda_S) \delta_{S'}$$

where  $\lambda_S$  is the prevalence of subgroup  $S$ .

- Test of hypotheses  $H_F : \delta_F \leq 0$  and  $H_S : \delta_S \leq 0$ .

# Patient Populations



- Assume the treatment effects in the subgroups satisfy  $\delta_S \geq \delta_{S'}$ .
- The treatment effect in  $F$  is

$$\delta_F = \lambda_S \delta_S + (1 - \lambda_S) \delta_{S'}$$

where  $\lambda_S$  is the prevalence of subgroup  $S$ .

- Test of hypotheses  $H_F : \delta_F \leq 0$  and  $H_S : \delta_S \leq 0$ .

# Which clinical trial design to choose?

## Classical Design:

Recruitment from population  $F$ .  
No Biomarker is determined.  
Test of  $H_F$ .

## Stratification Design:

Recruitment from population  $F$ .  
Stratified randomization by Biomarker.  
Test of  $H_F$  and  $H_S$ .

## Enrichment Design:

Recruitment only from population  $S$ .  
Test of  $H_S$

Biomarker  
Designs

# Which clinical trial design to choose?

## Classical Design:

Recruitment from population  $F$ .  
No Biomarker is determined.  
Test of  $H_F$ .

## Stratification Design:

Recruitment from population  $F$ .  
Stratified randomization by Biomarker.  
Test of  $H_F$  and  $H_S$ .

## Enrichment Design:

Recruitment only from population  $S$ .  
Test of  $H_S$

Biomarker  
Designs

# Which clinical trial design to choose?

## Classical Design:

Recruitment from population  $F$ .  
No Biomarker is determined.  
Test of  $H_F$ .

## Stratification Design:

Recruitment from population  $F$ .  
Stratified randomization by Biomarker.  
Test of  $H_F$  and  $H_S$ .

## Enrichment Design:

Recruitment only from population  $S$ .  
Test of  $H_S$

Biomarker  
Designs

# Which clinical trial design to choose?

## Classical Design:

Recruitment from population  $F$ .  
No Biomarker is determined.  
Test of  $H_F$ .

## Stratification Design:

Recruitment from population  $F$ .  
Stratified randomization by Biomarker.  
Test of  $H_F$  and  $H_S$ .

## Enrichment Design:

Recruitment only from population  $S$ .  
Test of  $H_S$

} Biomarker  
Designs

# Testing Procedures for Parallel Group Comparison of Means

## Classical Design:

$H_F$  is tested with a z-test.

## Stratification Design:

- $H_S$  and  $H_F$  are tested with a closed Spiessens-Debois (2010) test at levels  $\alpha_S, \alpha_F$ . If a hypothesis is rejected, the other is tested at level  $\alpha$ .
- To reject  $H_F$ , also the consistency condition

$$p_S \leq \tau_S \text{ and } p_{S'} \leq \tau_{S'},$$

for parameters  $\tau_S, \tau_{S'}$ , must be satisfied.

## Enrichment Design:

$H_S$  is tested with a z-test.

# Testing Procedures for Parallel Group Comparison of Means

## Classical Design:

$H_F$  is tested with a z-test.

## Stratification Design:

- $H_S$  and  $H_F$  are tested with a closed Spiessens-Debois (2010) test at levels  $\alpha_S, \alpha_F$ . If a hypothesis is rejected, the other is tested at level  $\alpha$ .
- To reject  $H_F$ , also the consistency condition

$$p_S \leq \tau_S \text{ and } p_{S'} \leq \tau_{S'},$$

for parameters  $\tau_S, \tau_{S'}$ , must be satisfied.

## Enrichment Design:

$H_S$  is tested with a z-test.

## Classical Design:

$H_F$  is tested with a z-test.

## Stratification Design:

- $H_S$  and  $H_F$  are tested with a closed Spiessens-Debois (2010) test at levels  $\alpha_S, \alpha_F$ . If a hypothesis is rejected, the other is tested at level  $\alpha$ .
- To reject  $H_F$ , also the consistency condition

$$p_S \leq \tau_S \text{ and } p_{S'} \leq \tau_{S'},$$

for parameters  $\tau_S, \tau_{S'}$ , must be satisfied.

## Enrichment Design:

$H_S$  is tested with a z-test.

# Testing Procedures for Parallel Group Comparison of Means

## Classical Design:

$H_F$  is tested with a z-test.

## Stratification Design:

- $H_S$  and  $H_F$  are tested with a closed Spiessens-Debois (2010) test at levels  $\alpha_S, \alpha_F$ . If a hypothesis is rejected, the other is tested at level  $\alpha$ .
- To reject  $H_F$ , also the consistency condition

$$p_S \leq \tau_S \text{ and } p_{S'} \leq \tau_{S'},$$

for parameters  $\tau_S, \tau_{S'}$ , must be satisfied.

## Enrichment Design:

$H_S$  is tested with a z-test.

## Classical Design:

$H_F$  is tested with a z-test.

## Stratification Design:

- $H_S$  and  $H_F$  are tested with a closed Spiessens-Debois (2010) test at levels  $\alpha_S, \alpha_F$ . If a hypothesis is rejected, the other is tested at level  $\alpha$ .
- To reject  $H_F$ , also the consistency condition

$$p_S \leq \tau_S \text{ and } p_{S'} \leq \tau_{S'},$$

for parameters  $\tau_S, \tau_{S'}$ , must be satisfied.

## Enrichment Design:

$H_S$  is tested with a z-test.

# Optimizing Clinical trial designs

- When is a **biomarker (BM) design beneficial** compared to a classical design?
- When to choose **stratified, when an enrichment design**?
- Which **sample size**?
- Which **significance levels**  $\alpha_F$  and  $\alpha_S$  for  $H_F$  and  $H_S$  in the weighted multiple test for the stratified design are optimal?

We apply a **utility based approach**, (cf. Beckman et al., 2011; Rosenblum et al., 2014; Graf et al., 2015), to model the expected utilities of a particular trial design from a **sponsor's** and a **public health** view.

# The Utility for a design $d$

$$U(d) = \underbrace{-C(d)}_{\text{Cost}} + \underbrace{\begin{cases} \varphi_{F,d} & \text{if } H_F \text{ is rejected} \\ \varphi_{S,d} & \text{if only } H_S \text{ is rejected} \\ 0 & \text{if no hypothesis is rejected} \end{cases}}_{\text{Reward}} .$$

## Sponsor view

$$\varphi_{F,d} = N \cdot r_F \cdot (\hat{\delta}_{F,d} - \mu_F)^+$$

$$\varphi_{S,d} = \lambda_S \cdot N \cdot r_S \cdot (\hat{\delta}_{S,d} - \mu_S)^+$$

- $N$  ... number of future patients (market size).
- $r_F, r_S$  ... revenue parameters.
- $\hat{\delta}_{F,d}, \hat{\delta}_{S,d}$  ... efficacy estimates.
- $\mu_F, \mu_S$  ... clinically relevant thresholds.

## Public health view

$$\varphi_{F,d} = N \cdot r_F \cdot (\delta_F - \mu_F)$$

$$\varphi_{S,d} = \lambda_S \cdot N \cdot r_S \cdot (\delta_S - \mu_S)$$

- $\delta_S, \delta_F$  ... true effect sizes.

## Sponsor view

$$\varphi_{F,d} = N \cdot r_F \cdot (\hat{\delta}_{F,d} - \mu_F)^+$$

$$\varphi_{S,d} = \lambda_S \cdot N \cdot r_S \cdot (\hat{\delta}_{S,d} - \mu_S)^+$$

- $N$  ... number of future patients (market size).
- $r_F, r_S$  ... revenue parameters.
- $\hat{\delta}_{F,d}, \hat{\delta}_{S,d}$  ... efficacy estimates.
- $\mu_F, \mu_S$  ... clinically relevant thresholds.

## Public health view

$$\varphi_{F,d} = N \cdot r_F \cdot (\delta_F - \mu_F)$$

$$\varphi_{S,d} = \lambda_S \cdot N \cdot r_S \cdot (\delta_S - \mu_S)$$

- $\delta_S, \delta_F$  ... true effect sizes.

## Sponsor view

$$\varphi_{F,d} = N \cdot r_F \cdot (\hat{\delta}_{F,d} - \mu_F)^+$$

$$\varphi_{S,d} = \lambda_S \cdot N \cdot r_S \cdot (\hat{\delta}_{S,d} - \mu_S)^+$$

- $N$  ... number of future patients (market size).
- $r_F, r_S$  ... revenue parameters.
- $\hat{\delta}_{F,d}, \hat{\delta}_{S,d}$  ... efficacy estimates.
- $\mu_F, \mu_S$  ... clinically relevant thresholds.

## Public health view

$$\varphi_{F,d} = N \cdot r_F \cdot (\delta_F - \mu_F)$$

$$\varphi_{S,d} = \lambda_S \cdot N \cdot r_S \cdot (\delta_S - \mu_S)$$

- $\delta_S, \delta_F$  ... true effect sizes.

## Sponsor view

$$\varphi_{F,d} = N \cdot r_F \cdot (\hat{\delta}_{F,d} - \mu_F)^+$$

$$\varphi_{S,d} = \lambda_S \cdot N \cdot r_S \cdot (\hat{\delta}_{S,d} - \mu_S)^+$$

- $N$  ... number of future patients (market size).
- $r_F, r_S$  ... revenue parameters.
- $\hat{\delta}_{F,d}, \hat{\delta}_{S,d}$  ... efficacy estimates.
- $\mu_F, \mu_S$  ... clinically relevant thresholds.

## Public health view

$$\varphi_{F,d} = N \cdot r_F \cdot (\delta_F - \mu_F)$$

$$\varphi_{S,d} = \lambda_S \cdot N \cdot r_S \cdot (\delta_S - \mu_S)$$

- $\delta_S, \delta_F$  ... true effect sizes.

## Sponsor view

$$\varphi_{F,d} = N \cdot r_F \cdot (\hat{\delta}_{F,d} - \mu_F)^+$$

$$\varphi_{S,d} = \lambda_S \cdot N \cdot r_S \cdot (\hat{\delta}_{S,d} - \mu_S)^+$$

- $N$  ... number of future patients (market size).
- $r_F, r_S$  ... revenue parameters.
- $\hat{\delta}_{F,d}, \hat{\delta}_{S,d}$  ... efficacy estimates.
- $\mu_F, \mu_S$  ... clinically relevant thresholds.

## Public health view

$$\varphi_{F,d} = N \cdot r_F \cdot (\delta_F - \mu_F)$$

$$\varphi_{S,d} = \lambda_S \cdot N \cdot r_S \cdot (\delta_S - \mu_S)$$

- $\delta_S, \delta_F$  ... true effect sizes.

## Sponsor view

$$\varphi_{F,d} = N \cdot r_F \cdot (\hat{\delta}_{F,d} - \mu_F)^+$$

$$\varphi_{S,d} = \lambda_S \cdot N \cdot r_S \cdot (\hat{\delta}_{S,d} - \mu_S)^+$$

- $N$  ... number of future patients (market size).
- $r_F, r_S$  ... revenue parameters.
- $\hat{\delta}_{F,d}, \hat{\delta}_{S,d}$  ... efficacy estimates.
- $\mu_F, \mu_S$  ... clinically relevant thresholds.

## Public health view

$$\varphi_{F,d} = N \cdot r_F \cdot (\delta_F - \mu_F)$$

$$\varphi_{S,d} = \lambda_S \cdot N \cdot r_S \cdot (\delta_S - \mu_S)$$

- $\delta_S, \delta_F$  ... true effect sizes.

- Classical Design

$$C_{\text{setup}} + 2n c_{\text{per-patient}}.$$

- Stratified Design

$$C_{\text{setup}} + C_{\text{BM development}} + 2n(c_{\text{per-patient}} + C_{\text{BM determination}}).$$

- Enrichment Design

$$C_{\text{setup}} + C_{\text{BM development}} + 2n\left(c_{\text{per-patient}} + \frac{C_{\text{BM determination}}}{\lambda_S}\right).$$

# Trial Costs $C(d)$

- Classical Design

$$c_{\text{setup}} + 2n c_{\text{per-patient}}.$$

- Stratified Design

$$c_{\text{setup}} + c_{\text{BM development}} + 2n(c_{\text{per-patient}} + c_{\text{BM determination}}).$$

- Enrichment Design

$$c_{\text{setup}} + c_{\text{BM development}} + 2n\left(c_{\text{per-patient}} + \frac{c_{\text{BM determination}}}{\lambda_S}\right).$$

# Trial Costs $C(d)$

- Classical Design

$$c_{\text{setup}} + 2n c_{\text{per-patient}}.$$

- Stratified Design

$$c_{\text{setup}} + c_{\text{BM development}} + 2n(c_{\text{per-patient}} + c_{\text{BM determination}}).$$

- Enrichment Design

$$c_{\text{setup}} + c_{\text{BM development}} + 2n\left(c_{\text{per-patient}} + \frac{c_{\text{BM determination}}}{\lambda_S}\right).$$

Expected Utility:

$$E_{\pi} \{E_{\Delta}[U(d)]\}$$

The expectation is taken over

- the prior  $\pi$  on the effect sizes  $\Delta = (\delta_S, \delta_{S'})$  and
- the sampling distribution

Optimal design: Choose the design with maximal expected utility optimizing over

- Type of design (classical/stratified/enrichment)
- Sample size
- $\alpha$  allocation (for the stratified design)

Expected Utility:

$$E_{\pi} \{E_{\Delta}[U(d)]\}$$

The expectation is taken over

- the prior  $\pi$  on the effect sizes  $\Delta = (\delta_S, \delta_{S'})$  and
- the sampling distribution

**Optimal design:** Choose the design with maximal expected utility optimizing over

- Type of design (classical/stratified/enrichment)
- Sample size
- $\alpha$  allocation (for the stratified design)

# Prior Distributions $\pi$ On the Effects $\delta_S, \delta_{S'}$

$\delta_S$	0	$\theta$	$\theta$	$\theta$
$\delta_{S'}$	0	0	$\theta/2$	$\theta$
Weak Biomarker Prior	0.2	0.2	0.3	0.3
Strong Biomarker Prior	0.2	0.6	0.1	0.1

$\theta > 0 \dots$  effect size parameter.

- **Effect size parameter in the prior**  
 $\theta = 0.3$
- **Reward parameters**  
 $Nr_F = Nr_S = 1000\text{MUSD}$   
 $\mu_F = \mu_S = 0.1.$
- **Cost Parameters in (MUSD)**  
 $c_{\text{setup}} = 1$   
 $c_{\text{per-patient}} = 0.05$   
 $c_{\text{BM development}} = 1$   
 $c_{\text{BM determination}} = 0.005.$
- **Consistency parameters**  $\tau_S = \tau_{S'} = 0.3.$

- **Effect size parameter in the prior**  
 $\theta = 0.3$
- **Reward parameters**  
 $Nr_F = Nr_S = 1000\text{MUSD}$   
 $\mu_F = \mu_S = 0.1.$
- **Cost Parameters in (MUSD)**  
 $c_{\text{setup}} = 1$   
 $c_{\text{per-patient}} = 0.05$   
 $c_{\text{BM development}} = 1$   
 $c_{\text{BM determination}} = 0.005.$
- **Consistency parameters**  $\tau_S = \tau_{S'} = 0.3.$

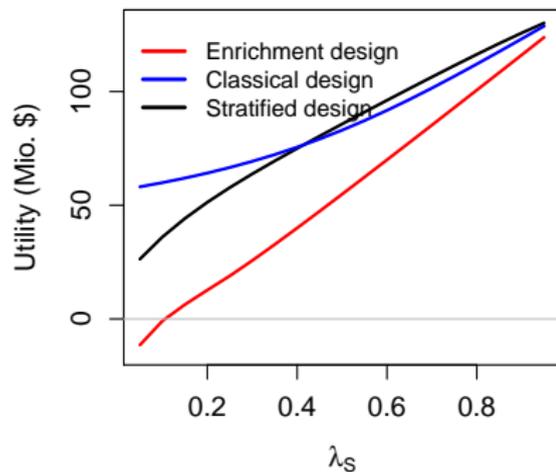
- **Effect size parameter in the prior**  
 $\theta = 0.3$
- **Reward parameters**  
 $Nr_F = Nr_S = 1000\text{MUSD}$   
 $\mu_F = \mu_S = 0.1.$
- **Cost Parameters in (MUSD)**  
 $c_{\text{setup}} = 1$   
 $c_{\text{per-patient}} = 0.05$   
 $c_{\text{BM development}} = 1$   
 $c_{\text{BM determination}} = 0.005.$
- **Consistency parameters**  $\tau_S = \tau_{S'} = 0.3.$

- **Effect size parameter in the prior**  
 $\theta = 0.3$
- **Reward parameters**  
 $Nr_F = Nr_S = 1000\text{MUSD}$   
 $\mu_F = \mu_S = 0.1.$
- **Cost Parameters in (MUSD)**  
 $c_{\text{setup}} = 1$   
 $c_{\text{per-patient}} = 0.05$   
 $c_{\text{BM development}} = 1$   
 $c_{\text{BM determination}} = 0.005.$
- **Consistency parameters**  $\tau_S = \tau_{S'} = 0.3.$

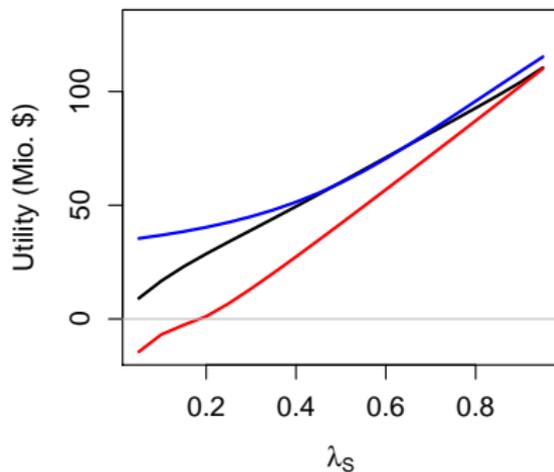
# Optimized Expected Utilities

Weak Biomarker Prior

## Sponsor



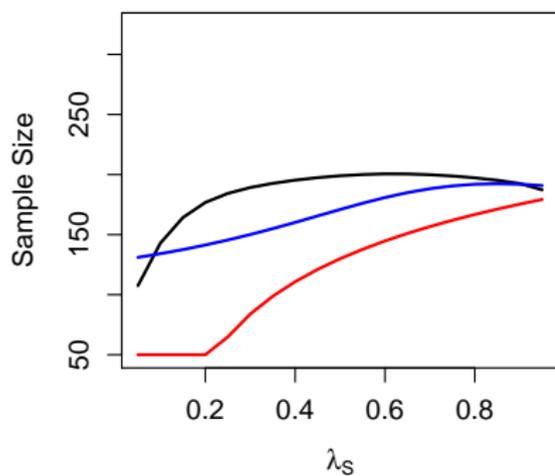
## Public Health



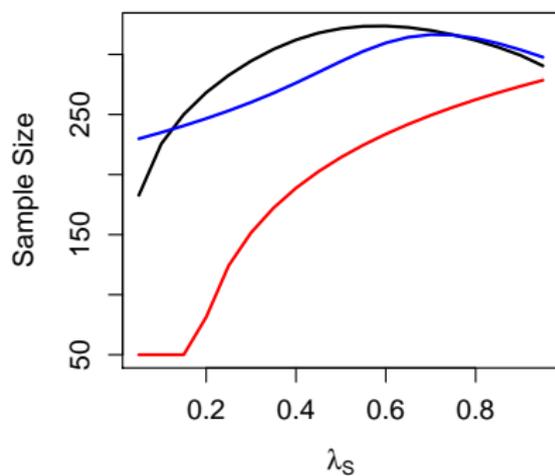
# Optimized Sample Size

Weak Biomarker Prior

**Sponsor**



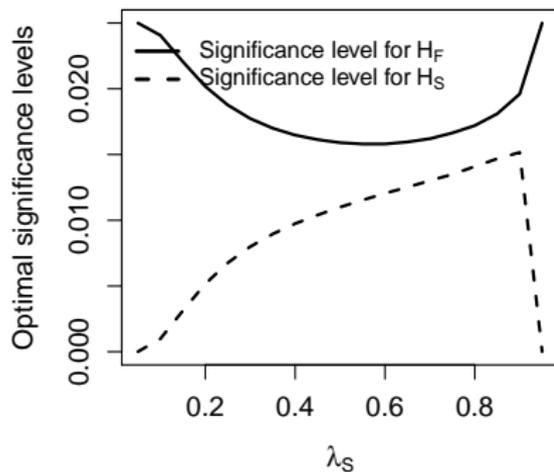
**Public Health**



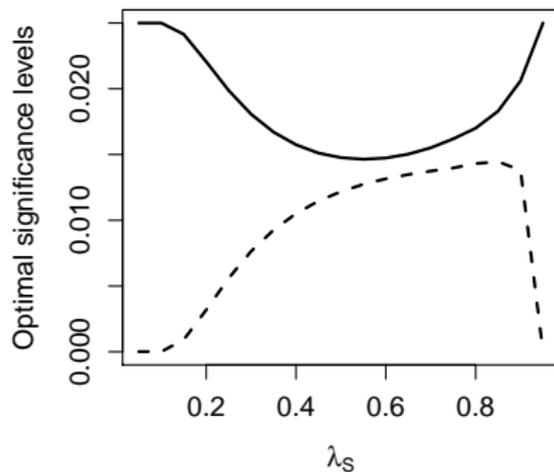
# Optimized Alpha Allocation

Weak Biomarker Prior

**Sponsor**

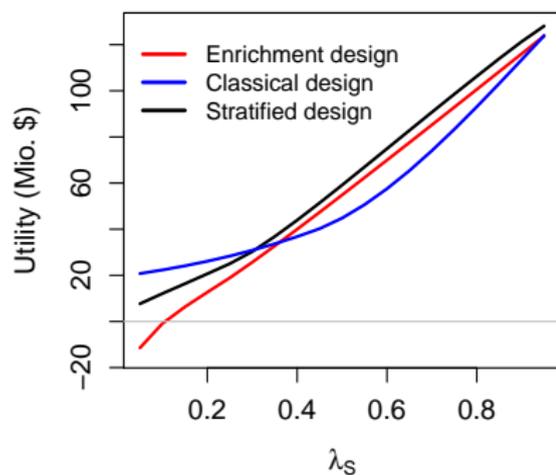


**Public Health**

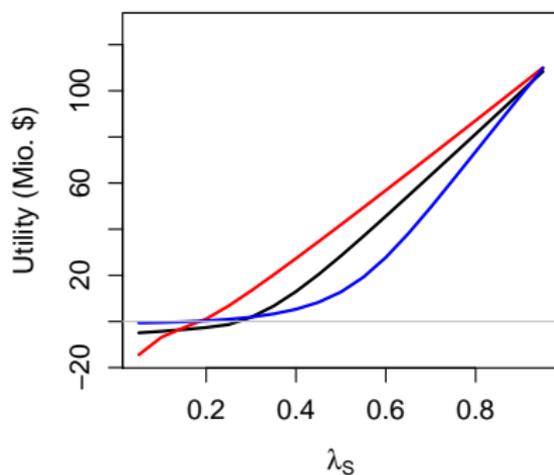


# Optimal Designs for the Strong Biomarker Prior

## Sponsor

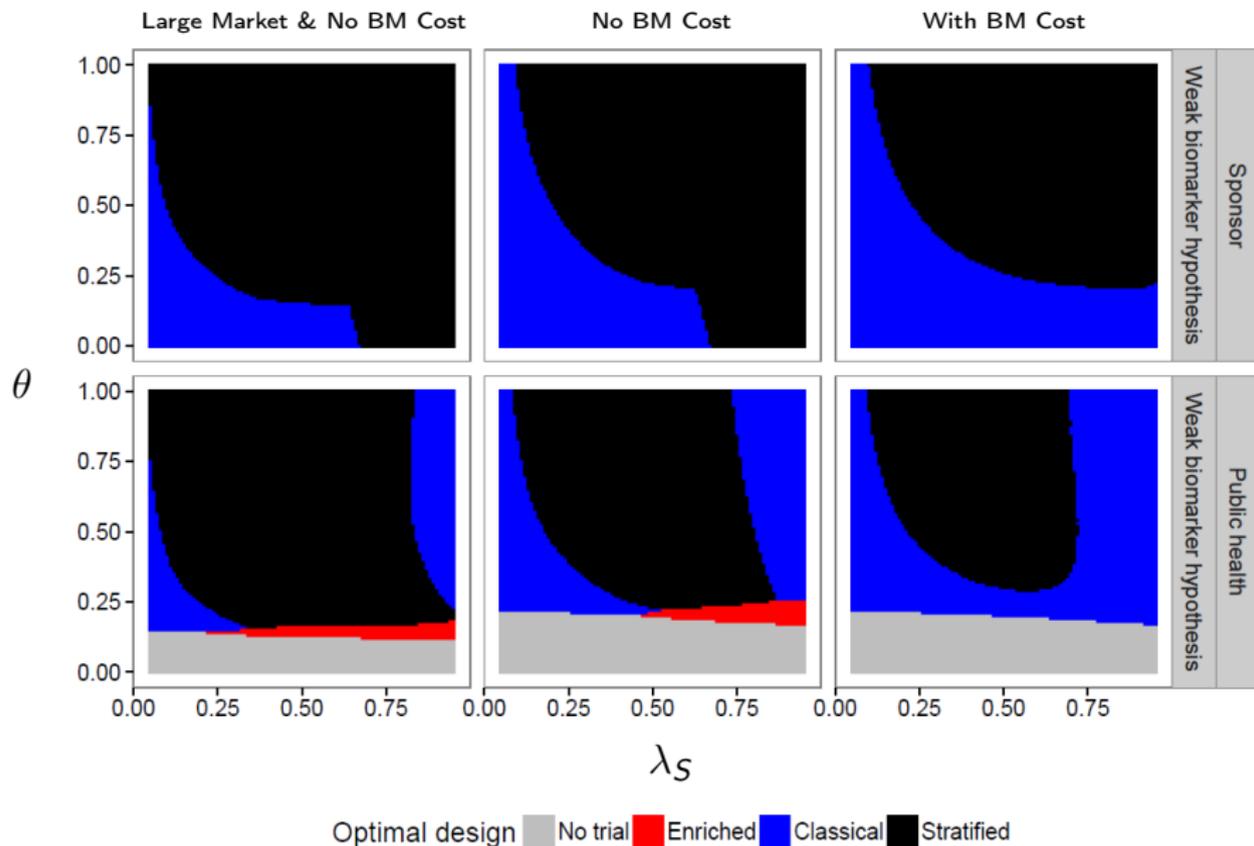


## Public Health



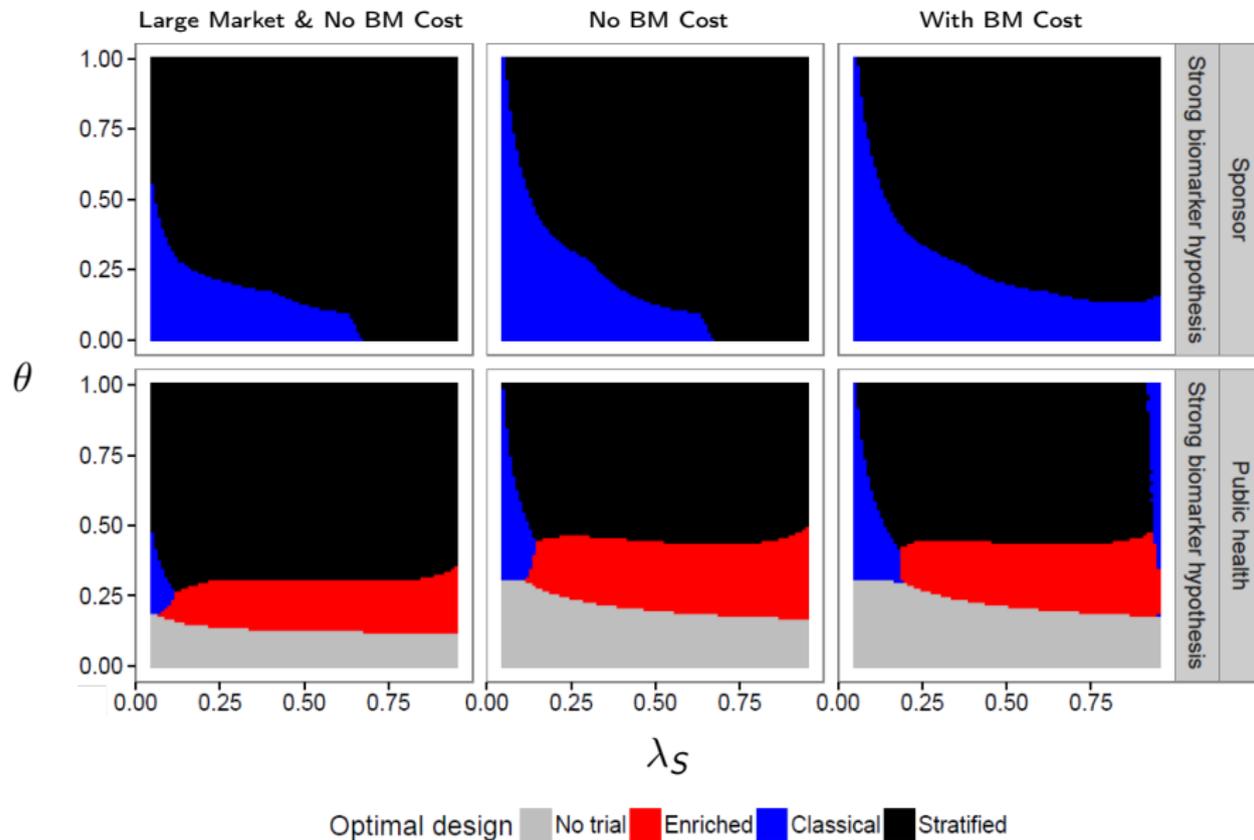
# Optimal Trial Designs

## Weak Biomarker Prior



# Optimal Trial Designs

## Strong Biomarker Prior



- The optimal sample size under the public health view is typically larger than in the sponsor view.
- For the considered priors, the enrichment design is never optimal for the sponsor view
- The optimal design depends strongly on the particulars of the situation: Subgroup prevalence, trial costs and initial beliefs.

# Two Extensions of the Trial Designs

## Partial Enrichment Design

The prevalence of the subgroup in the trial  $\lambda_T$  is a design parameter and may differ from  $\lambda_S$ , the prevalence in the population.

*Special cases are the stratified design ( $\lambda_T = \lambda_S$ ) and the (full) enrichment design ( $\lambda_T = 1$ ).*

E.g., Zhao et al. (2010)

## Adaptive Enrichment Designs

Two stage design, where the second stage sample size and second stage trial prevalence may depend on the first stage outcome.

E.g., Brannath et al. (2009); Beckman et al. (2011); Jenkins et al. (2011); Friede et al. (2012); Simon and Simon (2013); Krisam and Kieser (2015); Graf et al. (2015); Fisher and Rosenblum (2016)

For simplicity we use as multiple testing procedure a single step unweighted Bonferroni test.

## Partial Enrichment Design: Hypothesis test of $H_F$

Because trial and population prevalence do not match, the standard z-test is not a valid test for  $H_F$ .

- $H_F$  is tested with a reweighted z-statistics

$$\tilde{Z}_F = \xi \left( \frac{\lambda_S}{\sqrt{\lambda_T}} Z_S + \frac{1 - \lambda_S}{\sqrt{1 - \lambda_T}} Z_{S'} \right),$$

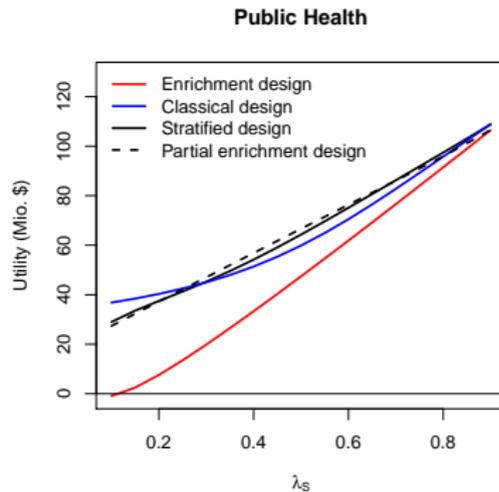
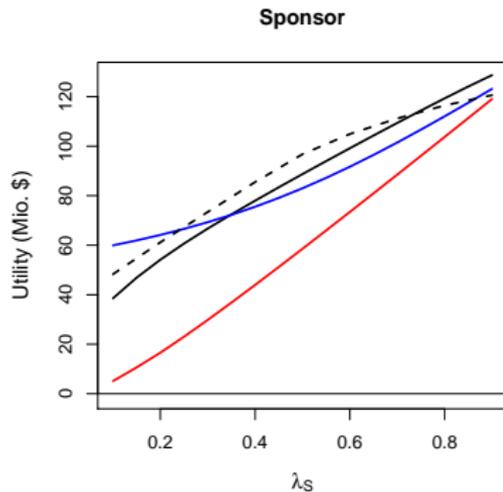
where  $Z_S, Z_{S'}$  denote the z-statistics for the subgroups  $S, S'$  and  $\xi$  is a normalizing constant.

e.g., Zhao et.al. 2010

- As above, to reject  $H_F$ , in addition the consistency condition needs to be fulfilled.

# Optimized Utilities

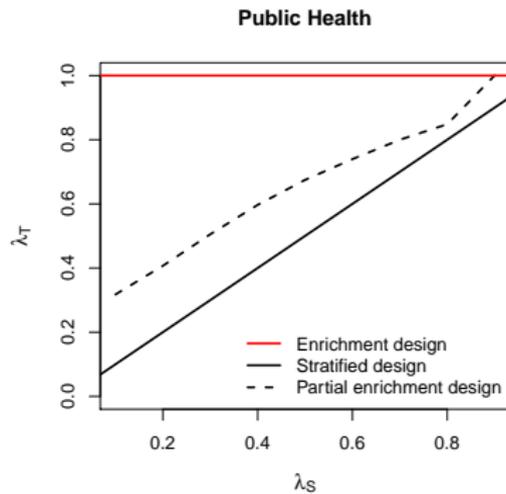
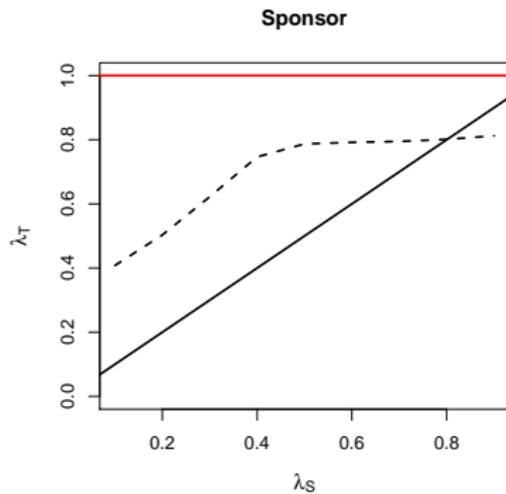
## Weak Biomarker Prior



(  $N_{rF} = N_{rS} = 1000$ ,  $c_{\text{setup}} = c_{\text{BM development}} = 1$ ,  $c_{\text{per-patient}} = 0.05$ ,  $c_{\text{BM determination}} = 0$ )

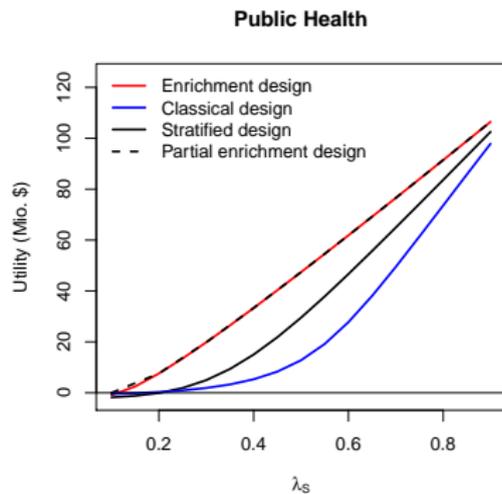
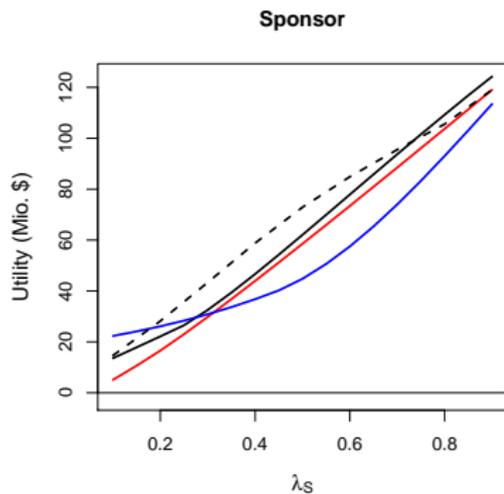
# Optimal Subgroup Prevalence $\lambda_T$

Weak Biomarker Prior



# Optimized Utilities

## Strong Biomarker Prior



## First stage

Sample size  $n_1$ , subgroup trial prevalence  $\lambda_T^1$

## Second Stage

Second stage sample size  $n_2$  and subgroup trial prevalence  $\lambda_T^2$  are chosen based on first stage data.

## Testing procedure

- Overall test statistics computed with combination function:

$$\begin{aligned}Z_S &= \sqrt{\frac{1}{2}} Z_S^1 + \sqrt{\frac{1}{2}} Z_S^2 \\ \tilde{Z}_F &= \sqrt{\frac{1}{2}} \tilde{Z}_F^1 + \sqrt{\frac{1}{2}} \tilde{Z}_F^2\end{aligned}$$

where  $Z_S^1, Z_S^2$  and  $\tilde{Z}_F^1, \tilde{Z}_F^2$  are stage wise z-statistics.

- Unweighted Bonferroni test boundaries applied to  $Z_S$  and  $\tilde{Z}_F$  (if  $\lambda_T^2 < 1$ ).

## Optimal Adaptation Rule

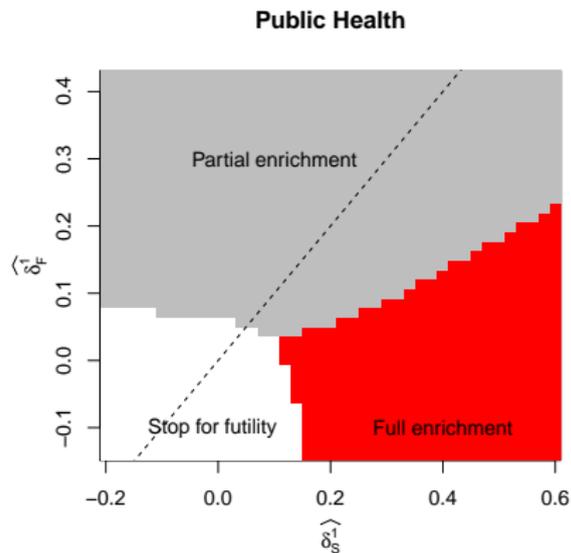
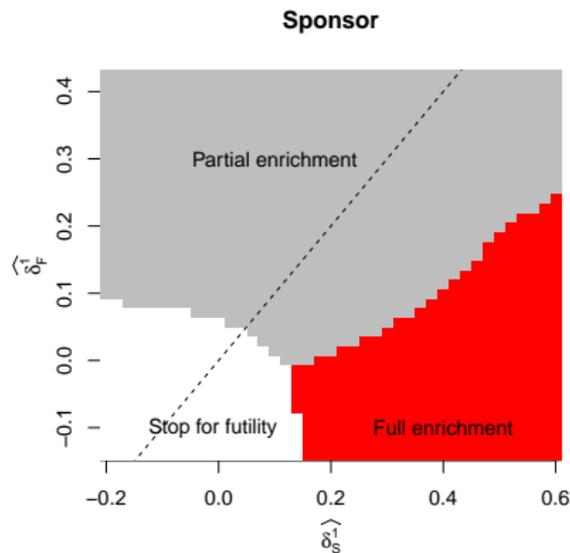
- In the interim analysis choose  $n_2, \lambda_T^2$  such that the expected utility conditional on the first stage data is maximized.
- Especially,  $n_2 = 0$  corresponds to stopping for futility,  $\lambda_T = 1$  to a second stage enrichment design.

## Optimizing first stage parameters

- Choose  $n_1, \lambda_T^1$  such that the expected utility (given the optimal adaptation rule is applied at interim) is maximized.

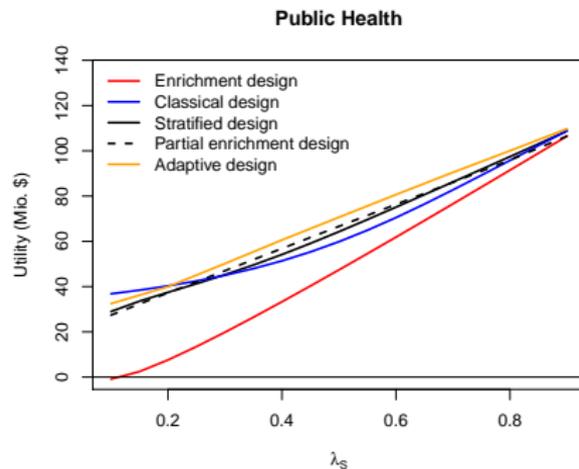
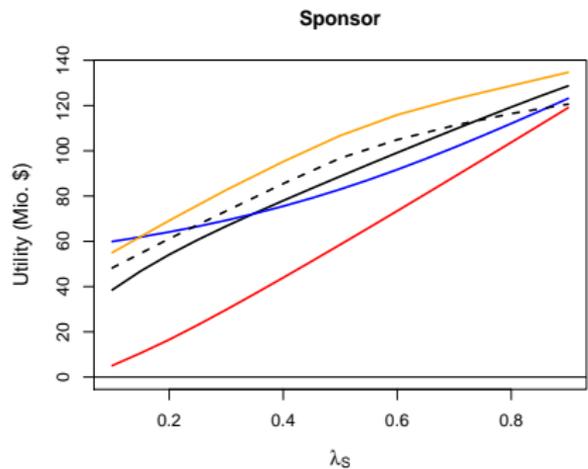
# Example for the Optimal Decision Rule

Weak Biomarker Prior ( $n_1 = 100, \lambda_S = \lambda_T^1 = 0.5$ )



# Optimized Utilities

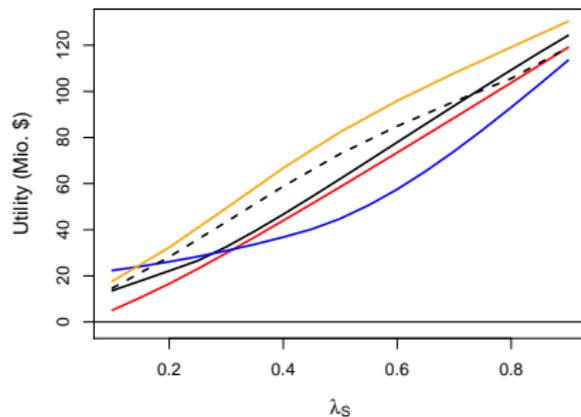
## Weak Biomarker Prior



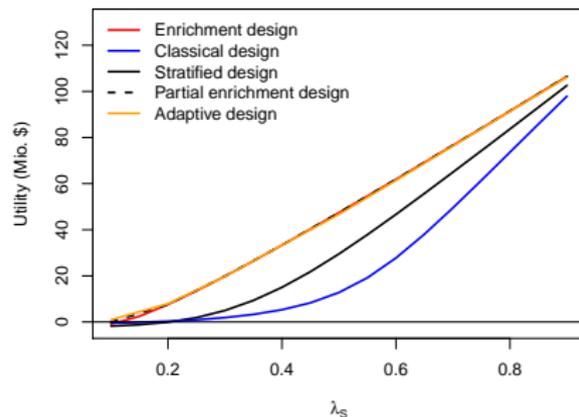
# Optimized Utilities

## Strong Biomarker Prior

### Sponsor



### Public Health



- Partial Enrichment Designs can increase the utility mainly for the sponsor utility function.
- Adaptive Enrichment Designs further increase the expected utility, also for the public health utility function.
- Extensions: weighted, stepwise Spiessens Debois test for the partial enrichment design, optimized weights in combination function...

Ondra, T., Jobjörnsson, S., Beckman, R. A., Burman, C. F., König, F., Stallard, N., Posch, M. (2016). Optimizing Trial Designs for Targeted Therapies. arXiv preprint arXiv:1606.03987.

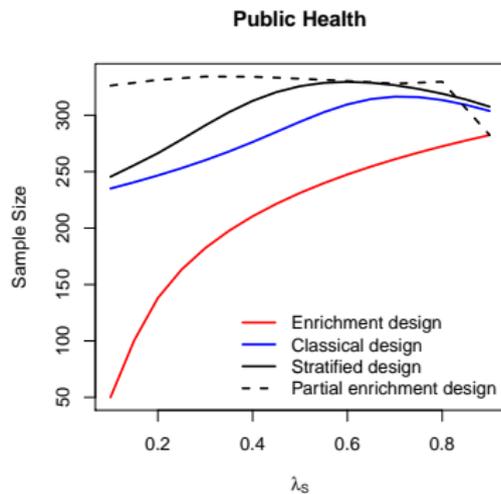
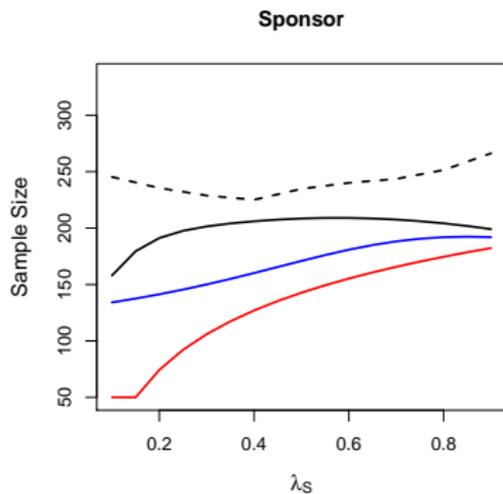
- Beckman, R. A., J. Clark, and C. Chen (2011). Integrating predictive biomarkers and classifiers into oncology clinical development programmes. *Nature Reviews Drug Discovery* 10(10), 735–748.
- Brannath, W., E. Zuber, M. Branson, F. Bretz, P. Gallo, M. Posch, and A. Racine-Poon (2009, May). Confirmatory adaptive designs with Bayesian decision tools for a targeted therapy in oncology. *Statistics in Medicine* 28(10), 1445–1463.
- Fisher, A. and M. Rosenblum (2016). Stochastic optimization of adaptive enrichment designs for two subpopulations. *Johns Hopkins University Bloomberg School of Public Health*.
- Friede, T., N. Parsons, and N. Stallard (2012, December). A conditional error function approach for subgroup selection in adaptive clinical trials. *Statistics in Medicine* 31(30), 4309–4320.
- Graf, A. C., M. Posch, and F. Koenig (2015, January). Adaptive designs for subpopulation analysis optimizing utility functions. *Biometrical Journal* 57, 76–89.
- Jenkins, M., A. Stone, and C. Jennison (2011, July). An adaptive seamless phase II/III design for oncology trials with subpopulation selection using correlated survival endpoints†. *Pharmaceutical Statistics* 10(4), 347–356.

- Krisam, J. and M. Kieser (2015). Optimal decision rules for biomarker-based subgroup selection for a targeted therapy in oncology. *International journal of molecular sciences* 16(5), 10354–10375.
- Ondra, T., S. Jobjörnsson, R. A. Beckman, C. Burman, F. König, S. Stallard, and P. M (2016). Optimizing trial designs for targeted therapies.
- Rosenblum, M., H. Liu, and E.-H. Yen (2014). Optimal tests of treatment effects for the overall population and two subpopulations in randomized trials, using sparse linear programming. *Journal of the American Statistical Association* 109(507), 1216–1228.
- Simon, N. and R. Simon (2013). Adaptive enrichment designs for clinical trials. *Biostatistics* 14(4), 613–625.
- Song, Y. and G. Y. H. Chi (2007, August). A method for testing a prespecified subgroup in clinical trials. *Statistics in Medicine* 26(19), 3535–3549.
- Spiessens, B. and M. Debois (2010, November). Adjusted significance levels for subgroup analyses in clinical trials. *Contemporary Clinical Trials* 31(6), 647–656.
- Zhao, Y. D., A. Dmitrienko, and R. Tamura (2010). Design and analysis considerations in clinical trials with a sensitive subpopulation. *Statistics in Biopharmaceutical Research* 2(1), 72–83.

## Backup Slides

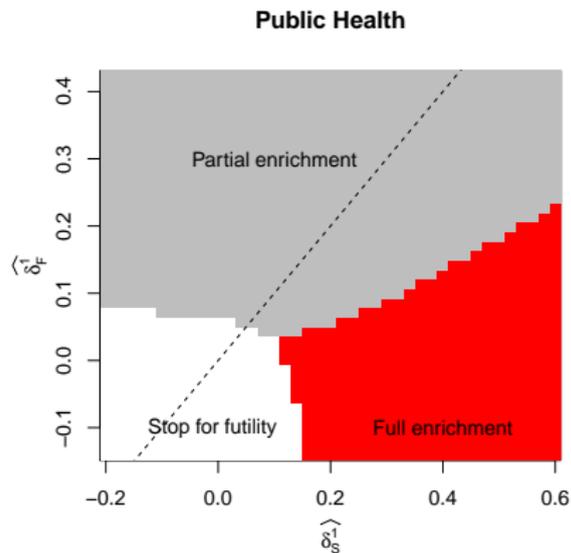
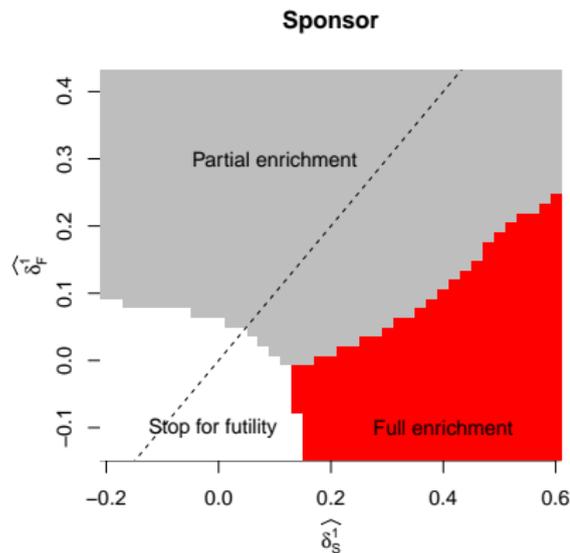
# Optimal Sample Size of the Partial Enrichment Design

## Weak Biomarker Prior



# Example for the Optimal Decision Rule

Weak Biomarker Prior ( $n_1 = 100, \lambda_S = \lambda_T^1 = 0.5$ )



# Example for the Optimal Decision Rule

Strong Biomarker Prior ( $n_1 = 100, \lambda_5 = \lambda_7 = 0.5$ )

