A Decision Theoretic Approach to Subgroup Selection†

Thomas Ondra

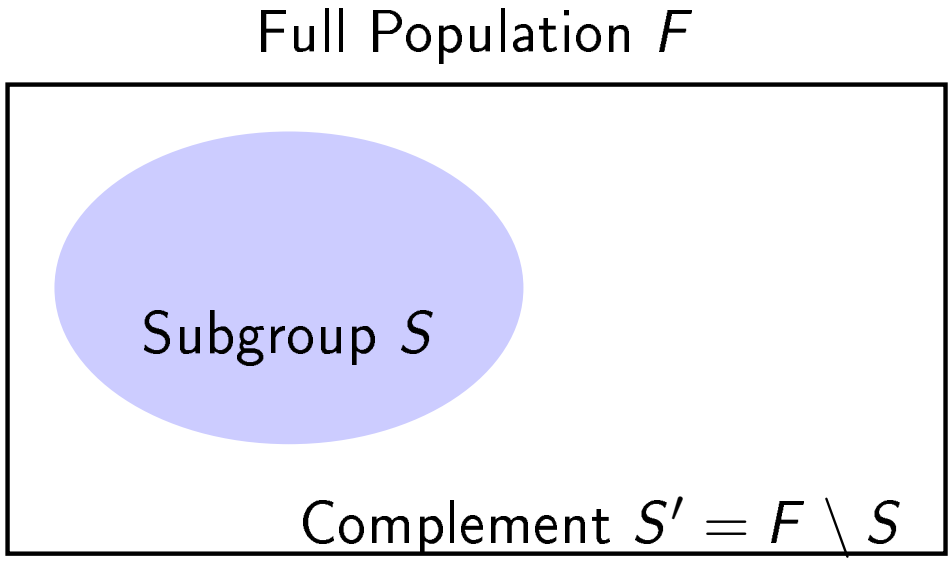
Section of Medical Statistics, Medical University of Vienna

Adaptive Designs and Multiple Testing Procedures
Cologne, 25 June 2015

† Thomas Ondra, Sebastian Jobjörnsson, Robert Beckman, Carl-Fredrik Burman, Franz König, Nigel Stallard and Martin Posch

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.

Framework



- Overall treatment effect

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

  where λ is the prevalence of subgroup S.
- We assume $\delta_{S'} \leq \delta_S$.
- Allows for investigating the hypotheses $H_F : \delta_F \leq 0$ and $H_S : \delta_S \leq 0$.

Classical, Enrichment and Stratification Design

Enrichment Design: Randomize only patients of subgroup S (say Biomarker +). Patients of the complement S' are excluded from the trial (Biomarker –).

Classical Design: Recruit from the full population F. No Biomarker is determined.

Stratification Design: Include Biomarker + and Biomarker – patients. Stratify randomization by biomarker status.

- With the enrichment design one can test H_S, i.e., for a treatment effect in the subpopulation.
- With the classical design one can test H_F.
- With the stratification design one can test H_S and H_F.

Testing Problem

Parallel group comparison of the means of normal distributions.

Enrichment Design:

- Test H_S with a z-test.

Classical Design:

- H_F with a z-test.

Stratification Design:

- Test H_S and H_F with a (stratified) z-test adjusting for multiplicity with a weighted test (Song and Chi, 2007; Spiessens and Debois, 2010)

The Spiessens-Debois test in the Stratified Design

For adjusted significance levels α_F, α_S the Spiessens and Debois test rejects

$$H_F \text{ if } p_F \leq \alpha_F \quad \text{and} \quad H_S \text{ if } p_S \leq \alpha_S,$$

where p_F, p_S are the p-values of the z-tests for H_F and H_S.

Some remarks:

- For fixed α_F and α, the level α_S is chosen such that

$$\mathbb{P}_{H_F \cap H_S}\left(p_F < \alpha_F \text{ or } p_S < \alpha_S\right) = \alpha.$$

- For fixed α_F the level α_S increases with the prevalence λ because the correlation of the test statistics increases.
- Formulas well known from group sequential tests.

Optimizing trial designs

- Traditionally power arguments can be the basis for determining the best trial design.
- An alternative is to apply a utility based approach (Graf et al., 2015; Beckman et al., 2011)
- We model the sponsors gain and costs of a particular trial design.
- Best trial design is determined by maximizing the sponsors profit.

In particular we optimize the following aspects of a clinical trial:

- Which type of design (Enrichment Design/Classical Design/Stratified Design) to choose?
- Which sample size?
- Which significance levels α_F and α_S for H_F and H_S in the weighted multiple test for the stratified design are optimal?

The Sponsors Profit

- Price approval rule for the subgroup S. Measure of the revenue if the drug is licensed for the subgroup.

$$U = \lambda \cdot \varphi_{S,d} + (1-\lambda) \cdot \varphi_{S',d} - (c_{1,d} + c_{2,d} n)$$

The Sponsors Profit

- Price approval rule for the subgroup S. Measure of the revenue if the drug is licensed for the subgroup.

$$U = \lambda \cdot \varphi_{S,d} + (1-\lambda) \cdot \varphi_{S',d} - (c_{1,d} + c_{2,d} n)$$

- Prevalence

The Sponsors Profit

- Price approval rule for the subgroup S. Measure of the revenue if the drug is licensed for the subgroup.

$$U = \lambda \cdot \varphi_{S,d} + (1-\lambda) \cdot \varphi_{S',d} - (c_{1,d} + c_{2,d} n)$$

- Prevalence
- Price approval rule for S'

The Sponsors Profit

- Price approval rule for the subgroup S. Measure of the revenue if the drug is licensed for the subgroup.

$$U = \lambda \cdot \varphi_{S,d} + (1-\lambda) \cdot \varphi_{S',d} - (c_{1,d} + c_{2,d} n)$$

- Prevalence
- Price approval rule for S'
- Costs

The Sponsors Profit

- Price approval rule for the subgroup S. Measure of the revenue if the drug is licensed for the subgroup.

$$U = \lambda \cdot \varphi_{S,d} + (1-\lambda) \cdot \varphi_{S',d} - (c_{1,d} + c_{2,d} n)$$

- Prevalence
- Price approval rule for S'
- Costs

Costs split up into setup costs $c_{1,d}$ and costs per patient $c_{2,d}$.

Price Approval Rules of Payers and Regulators

We assume that the price approval rules $\varphi_S, \varphi_{S'}$ depend on the data via

- the observed effect sizes,
- a significant result in the respective statistical test.

If the drug is licensed in the full population, then $\varphi_S = \varphi_{S'}$.

Price Approval Rule for the Enrichment and the Classical Design

Enrichment Design:

$$\varphi_S = \psi_S \cdot g \cdot \left(\hat{\delta}_S - \mu \right)^+$$

- Test decision

Price Approval Rule for the Enrichment and the Classical Design

Enrichment Design:

$$\varphi_S = \psi_S \cdot g \cdot \left(\hat{\delta}_S - \mu \right)^+$$

- Test decision
- Revenue parameter

Price Approval Rule for the Enrichment and the Classical Design

Enrichment Design:

$$\varphi_S = \psi_S \cdot g \cdot \left(\hat{\delta}_S - \mu \right)^+$$

- Test decision
- Revenue parameter
- Efficacy estimate

Price Approval Rule for the Enrichment and the Classical Design

Enrichment Design:

$$\varphi_S = \psi_S \cdot g \cdot \left(\hat{\delta}_S - \mu \right)^+$$

- Test decision
- Revenue parameter
- Efficacy estimate
- Clinically relevant threshold

Price Approval Rule for the Enrichment and the Classical Design

Enrichment Design:

$$\varphi_S = \psi_S \cdot g \cdot \left(\hat{\delta}_S - \mu \right)^+$$

- Test decision
- Revenue parameter
- Efficacy estimate
- Clinically relevant threshold
- $\varphi_{S'} = 0$

Price Approval Rule for the Enrichment and the Classical Design

Enrichment Design:

$$\varphi_S = \psi_S \cdot g \cdot \left(\hat{\delta}_S - \mu \right)^+$$

- Test decision
- Revenue parameter
- Efficacy estimate
- Clinically relevant threshold
- $\varphi_{S'} = 0$

Classical Design:

- $\varphi_S = \varphi_{S'} = \psi_F \cdot g \cdot (\hat{\delta}_F - \mu)^+$.
- ψ_F denotes the test decision, based on the z-test (in the full population).

Price Approval Rule for the Stratified Design

- If $\psi_F = 1$ then

$$\varphi_S = \varphi_{S'} = 1_{\{\hat{\delta}_S, \hat{\delta}_{S'} \geq \mu\}} \cdot \psi_F \cdot g \cdot (\hat{\delta}_F - \mu)^+.$$

- If the drug is not licensed in the full population but $\psi_S = 1$ then

$$\varphi_S = \psi_S \cdot g \cdot (\hat{\delta}_S - \mu)^+.$$
$$\varphi_{S'} = 0$$

Numerical Results - Parameters

Parameters	Classical	Enrichment	Stratified
Revenue parameter g	1 Billion\$	1 Billion\$	1 Billion\$
Efficacy threshold μ	0.1	0.1	0.1
Fixed costs	1Mio\$	16Mio\$	16Mio\$
Costs per Patient	100K\$	100K\$ + $\frac{10}{\lambda}$K\$	100K\$ + 10K\$

Numerical Results - Priors & Optimization

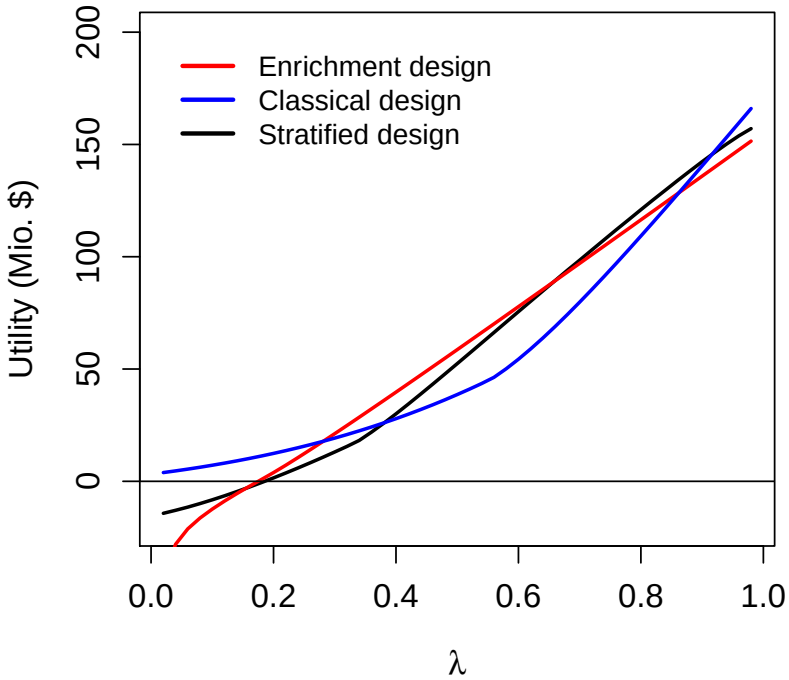
Optimal design is determined by maximizing the expected utility.
Single point prior on $\delta_S = 0.3$ (with weight 1).
Two scenarios for $\delta_{S'}$:

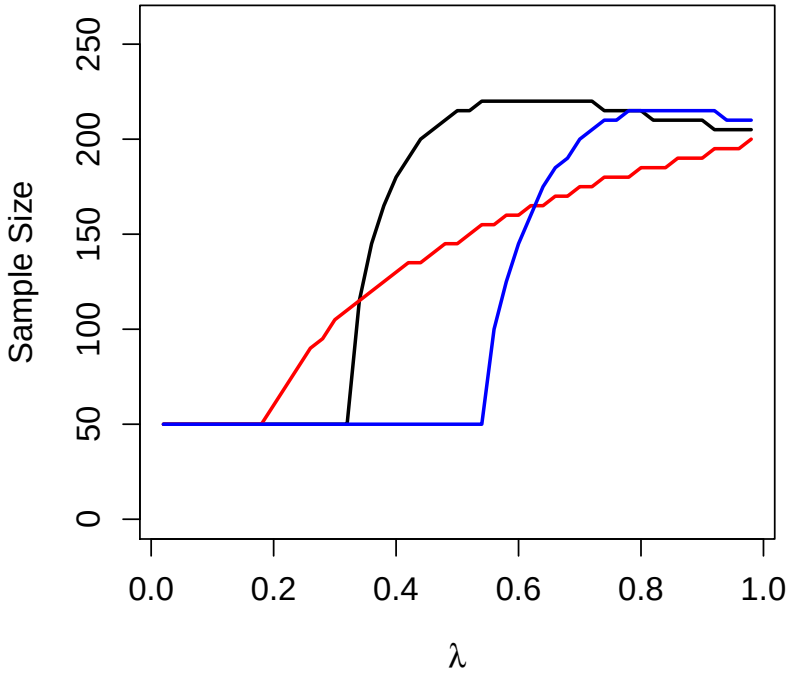
- Single point prior on $\delta_{S'} = 0$.
- Discrete prior on $\delta_{S'}$ on a grid in $[0, \delta_S]$.

Optimization of sample size and local significance levels over a grid (minimum sample size $n = 50$ per group).

- Sample size $n \in \{50, 55, \ldots, 400\}$
- $\alpha_F \in \{0, 0.001, \ldots, 0.025\}$

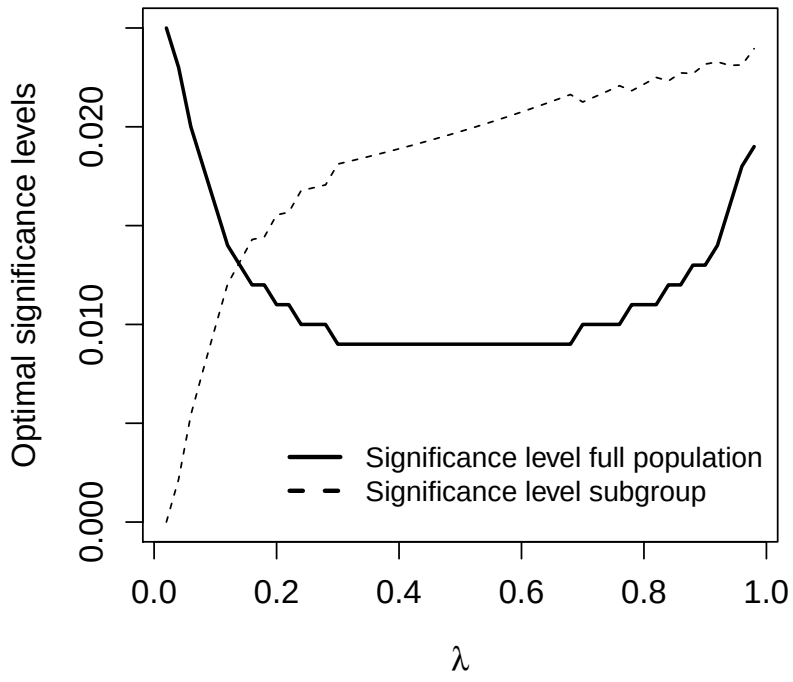
Point prior on $\delta_{S'} = 0$: Optimal sample size and corresponding utility





- For low prevalences all designs have negative or very low utilities.
- For intermediate prevalences enrichment dominates
- For larger prevalences the stratification design dominates.
- For very large prevalences the classical design dominates.

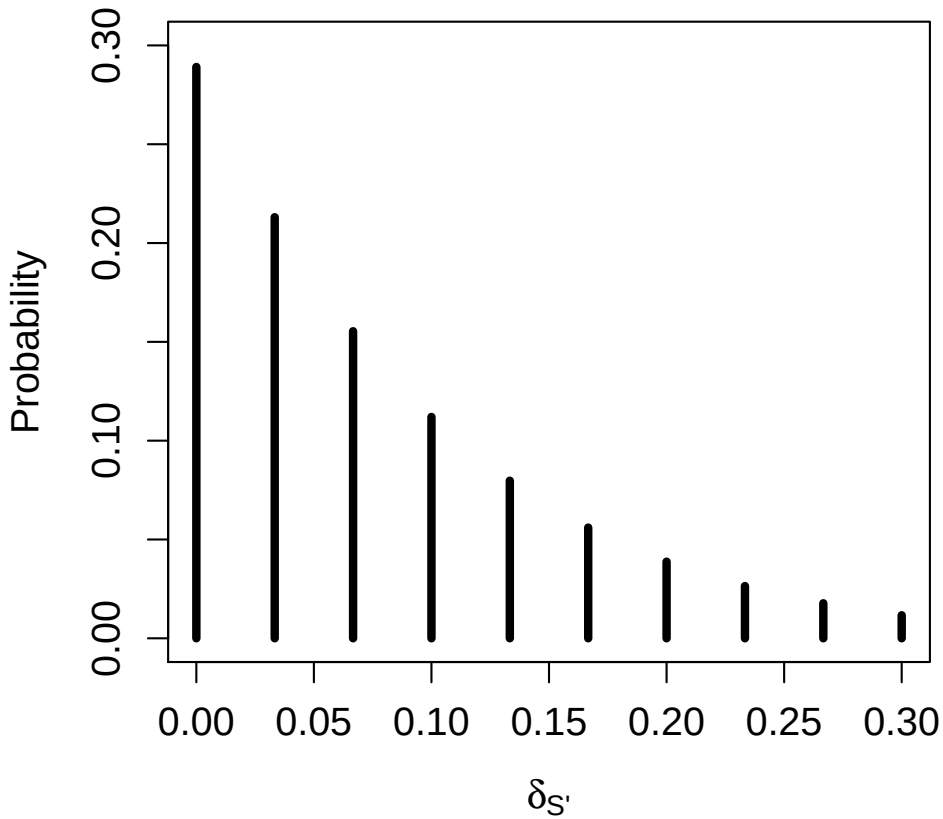
Point prior on $\delta_{S'} = 0$: Optimal Significance Levels for the Stratified Design



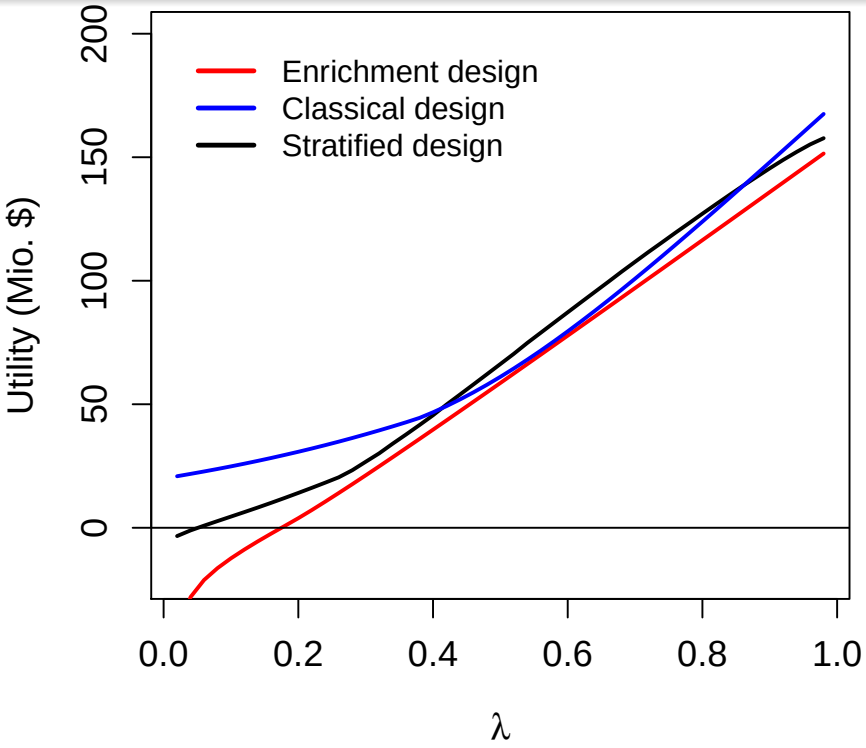
- Low prevalences: optimal test puts most of the weight on H_F.
- Intermediate prevalences: weight is more evenly distributed on H_F and H_S.
- Large prevalences: test statistics highly correlated. Both hypotheses can be tested at nearly unadjusted levels.

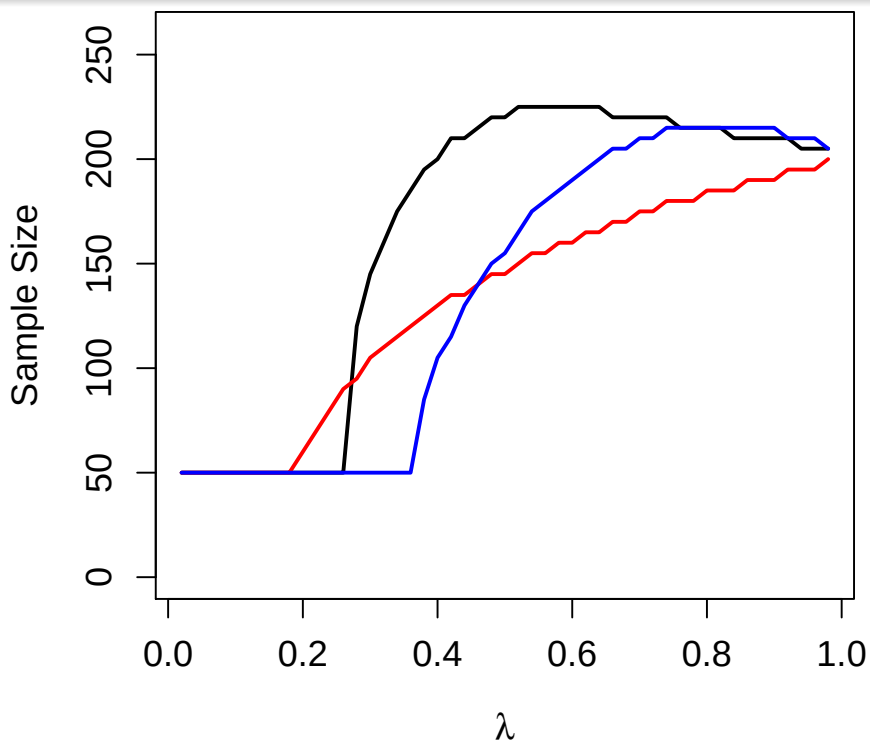
A sceptic prior for $\delta_{S'}$

- Point prior with weight 1 on $\delta_S = 0.3$.
- Sceptic prior on $\delta_{S'}$ corresponding to the belief that the efficacy in S' is some fraction of that in S.
- Larger probabilities attached to smaller effects.



Sceptic prior for $\delta_{S'}$: Optimal sample size and corresponding utility





- For low prevalences the qualitative results are as for the case of $\delta_{S'} = 0$.
- For intermediate prevalences the stratification design is optimal (The possibility that $\delta_{S'} > 0$ increases its utility).
- As above, for very larger prevalences the classical design is optimal.

Conclusions

- In general, the decision theoretic approach gives guidance regarding the choice of the trial design (including the type of the trial, the choice of the sample size and the weights in the multiple test)
- The analysis of the example shows that the optimal sponsor decision depends strongly on the particulars of the situation. Subgroup prevalence, trial costs and initial beliefs are all important to consider when making the design choice.
- The model can be extended in several directions.
  - Include the public health perspective
  - Allow for partial enrichment.
  - Include adaptive enrichment designs.

References

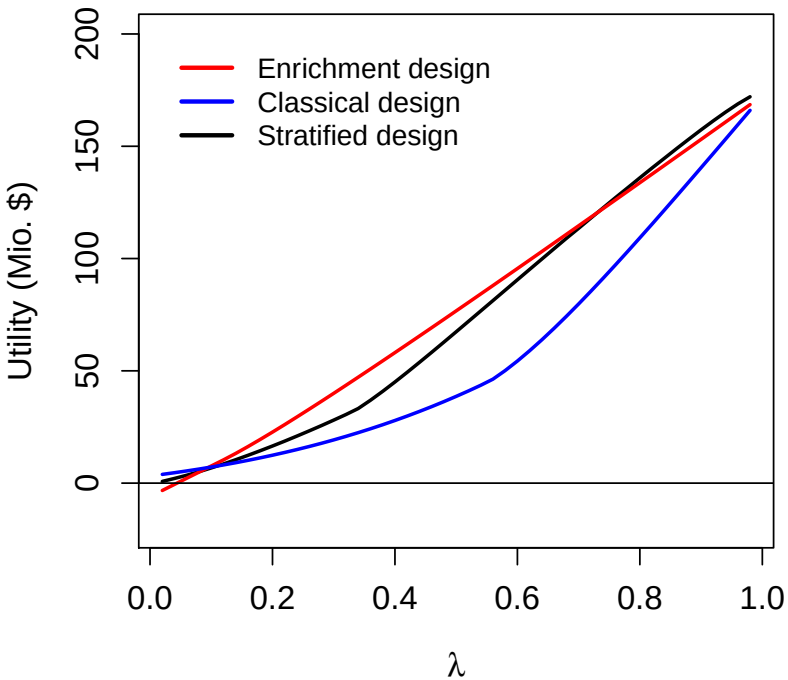
Beckman, R. A., J. Clark, and C. Chen (2011). Integrating predictive biomarkers and classifiers into oncology clinical development programmes. *Nature Reviews Drug Discovery 10*(10), 735–748.

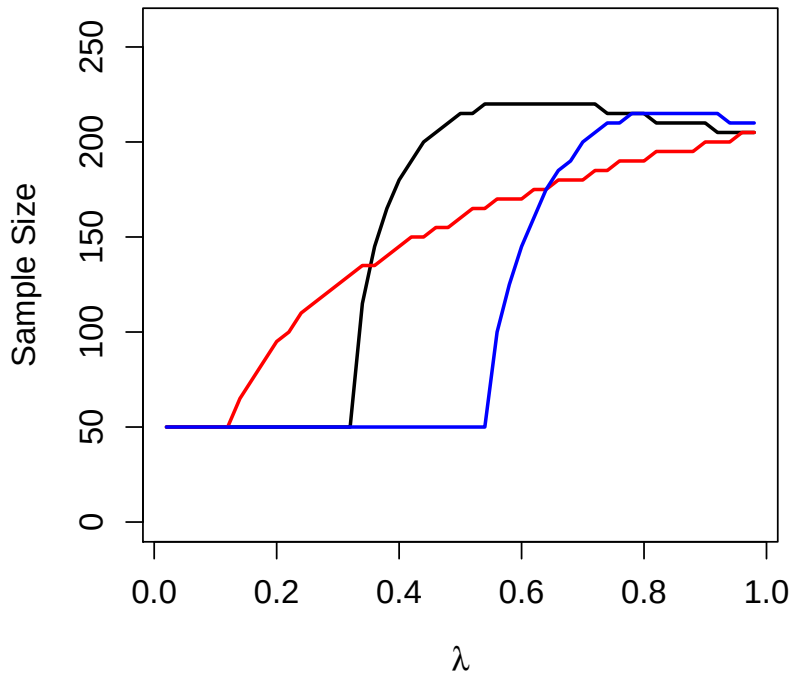
Graf, A. C., M. Posch, and F. Koenig (2015, January). Adaptive designs for subpopulation analysis optimizing utility functions. *Biometrical Journal 57*, 76–89.

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.

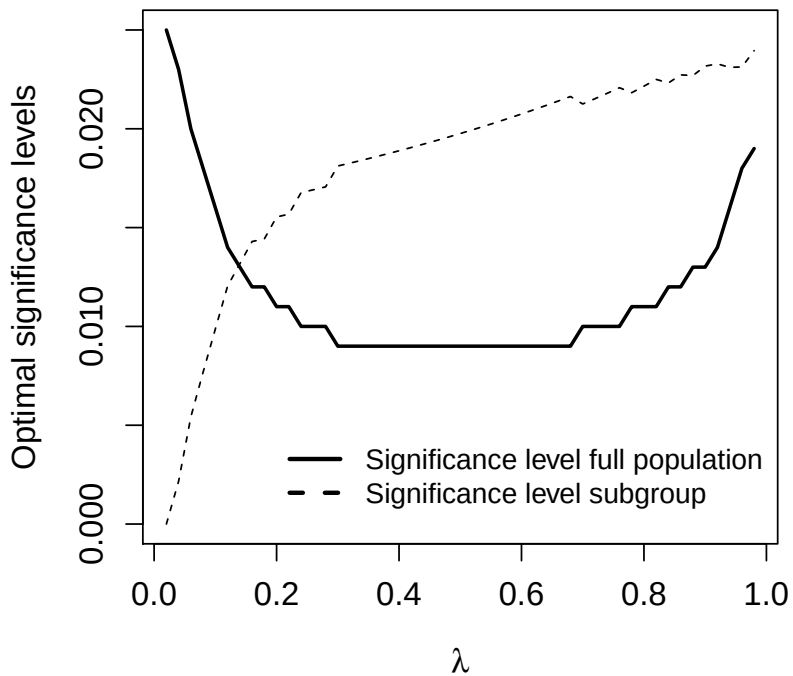
Spiessens, B. and M. Debois (2010, November). Adjusted significance levels for subgroup analyses in clinical trials. *Contemporary Clinical Trials 31*(6), 647–656.

No Biomarker Costs, Point Prior on $\delta_{S'} = 0$, Optimal Sample Size and corresponding Utility

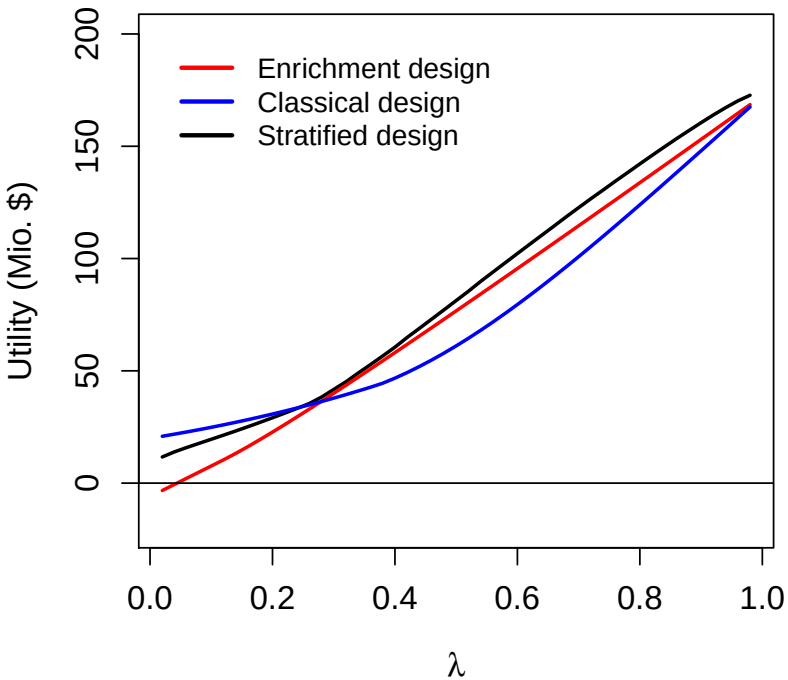


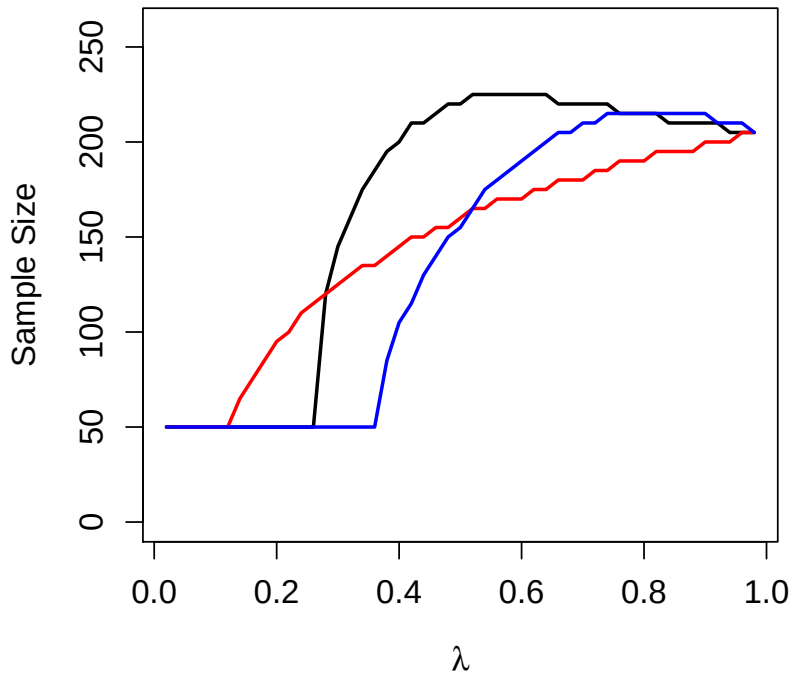


No Biomarker Costs, Point Prior on $\delta_{S'} = 0$, Optimal Sample Size and corresponding Utility



No Biomarker Costs, Sceptic Prior on $\delta_{S'} = 0$, Optimal Sample Size and corresponding Utility





No Biomarker Costs, Sceptic Prior on $\delta_{S'} = 0$, Optimal Sample Size and corresponding Utility