Designs and estimation for clinical trials with subpopulation selection

Yi-Da Chiu, Thomas Jaki
and
Franz Koenig, Martin Posch

Lancaster University, UK
and
Medical University of Vienna, Austria

Adaptive designs and multiple testing workshop, Padova

April 29, 2016

Subgroup effect: Variability in treatment responses between distinct, clearly identifiable groups within a patient population.

Common assumption underlying clinical trials:

- The trial population is homogeneous.

Potential errors: (Chen & Beckman, 2009)

- Missed treatment effect.
- Masked treatment effect.

Motivation: Rather limited literature addresses estimation problem for treatment effect in clinical trials with subgroup selection
(eg. Kimani *et al.*, 2015, Magnusson & Turnbull 2013)

Primary Aim: assess impacts on maximal likelihood estimate of subgroup treatment effects under various design features after the process of population selection.

Context: Cases with multiple distinct, pre-defined subgroups, attempt to select a subgroup and further confirm its efficacy of treatment effect.

The basic design set-up:

- λ_j: the known prevalence of subgroup S_j, $j = 1, \ldots, J$.
- subgroup sample sizes are fixedly decided by λ_j.
- total sample size keeps constant across K stages.
- randomly allocate patients to treatment and control group with ratio 1:1.
- Normally distributed observations are considered for all subpopulations with a common variance

Let $\mathcal{S}$ is the index set corresponding to any target population for selection. If $\mathcal{S} = \{1, f\}$, it indicates to select S_1 or F.

Given family-wise error rate α, the designs we consider is to control α in the strong sense and the testing hypotheses are

$$H_0 : \bigcap_s H_{0s}, \quad s \in \mathcal{S},$$

where the individual null hypothesis

$$H_{0s} : \theta_s \leq 0, \quad s \in \mathcal{S}.$$

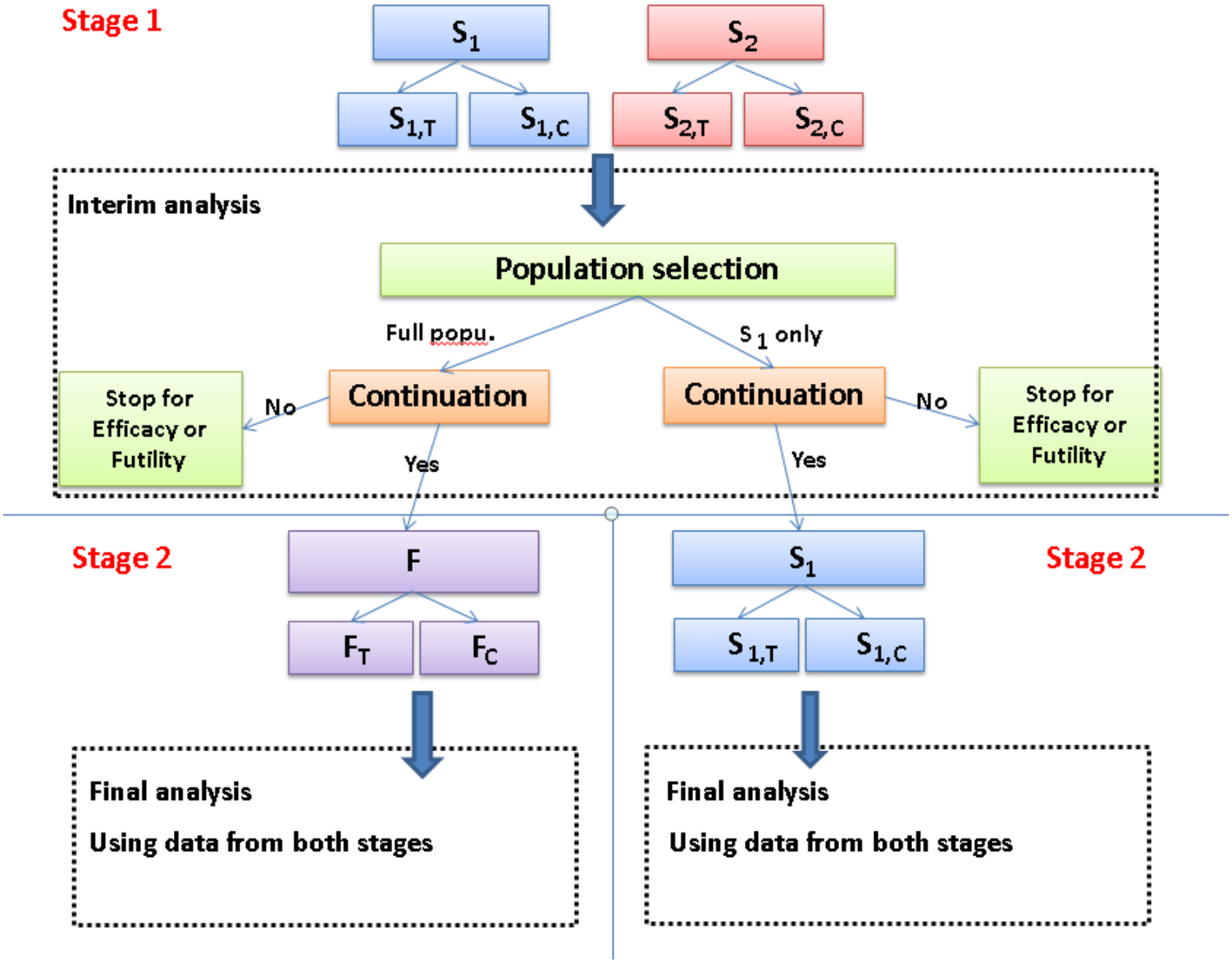
Here we consider the *best effect size* selection rule

$$Z_s^{(1)} = \max_{i \in \mathcal{S}} Z_i^{(1)}, \tag{1}$$

where

$$\begin{aligned} Z_i^{(1)} &= (\bar{Y}_{i,T}^{(1)} - \bar{Y}_{i,C}^{(1)}) \cdot I_i^{(1)}, \text{ a test statistic of } \theta_i \\ \bar{Y}_{i,T}^{(k)} &= \text{sample mean of treatment group at stage k for } S_i, \\ I_i^{(k)} &= \text{information level at stage k for testing } H_{0i}, \end{aligned}$$

Figure: Design Procedures for two-stage designs with two subgroups and considering $\mathcal{S} = \{1, f\}$



The iterative search of stopping boundaries is conducted based on

$$\alpha \geq \sum_{s \in \mathcal{S}} \left\{ \sum_{k=1}^{K} \left[\int \dots \int_{A_k} p_s^0 \left(\prod_{m=1}^{k} q_{sm}^0 dZ_s^{(1)} \dots dZ_s^{1:m} \right) \right] \right\}, \quad (2)$$

where

$$\begin{aligned} A_k &= \text{the integration region} [C_{l_1}, C_{u_1}] \times [C_{l_2}, C_{u_2}] \times \dots \times [C_{l_k}, \infty) \\ &\quad \text{on } (Z_s^{(1)}, Z_s^{1:2} \dots, Z_s^{1:k}), \\ p_s^0 &= \text{The joint density of } (Z_s^{(1)}, s) \text{ under } H_0, \\ q_{sm}^0 &= \text{The conditional density of } Z_s^{1:k} \text{ given its precursor} \\ &\quad Z_s^{1:(k-1)} \text{ at stage } k-1, \text{under } H_0 \end{aligned}$$

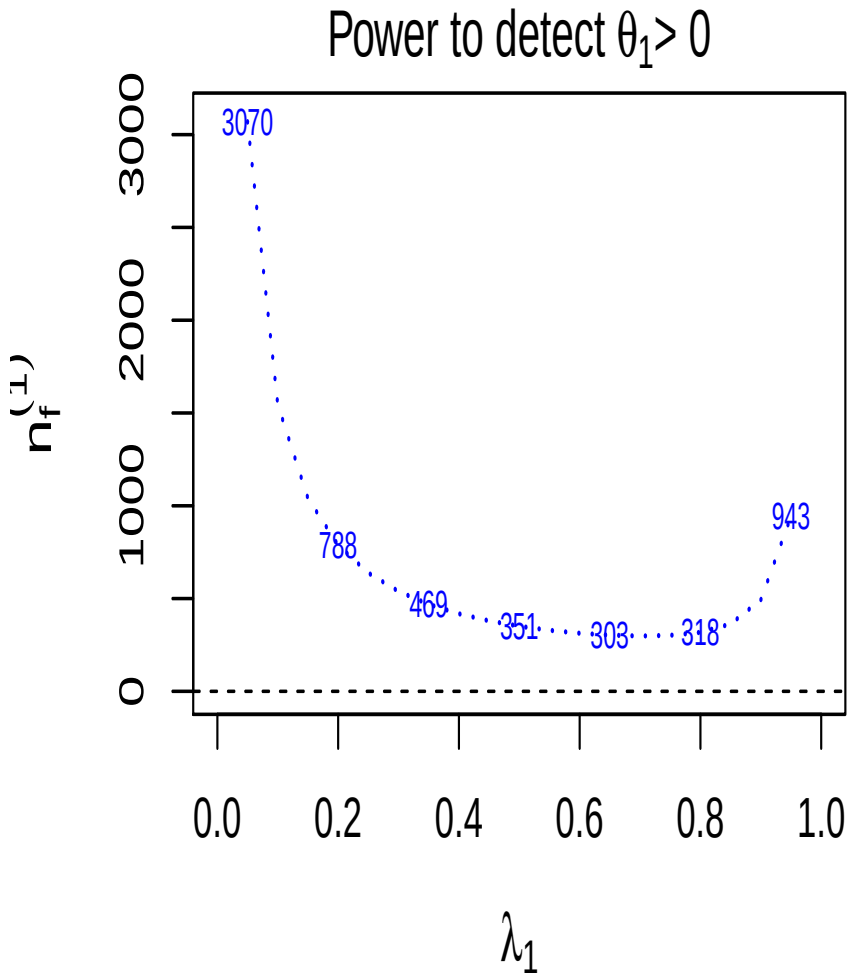
The stagewise total sample size for the full population $n_f^{(k)}$ can be searched iteratively upon

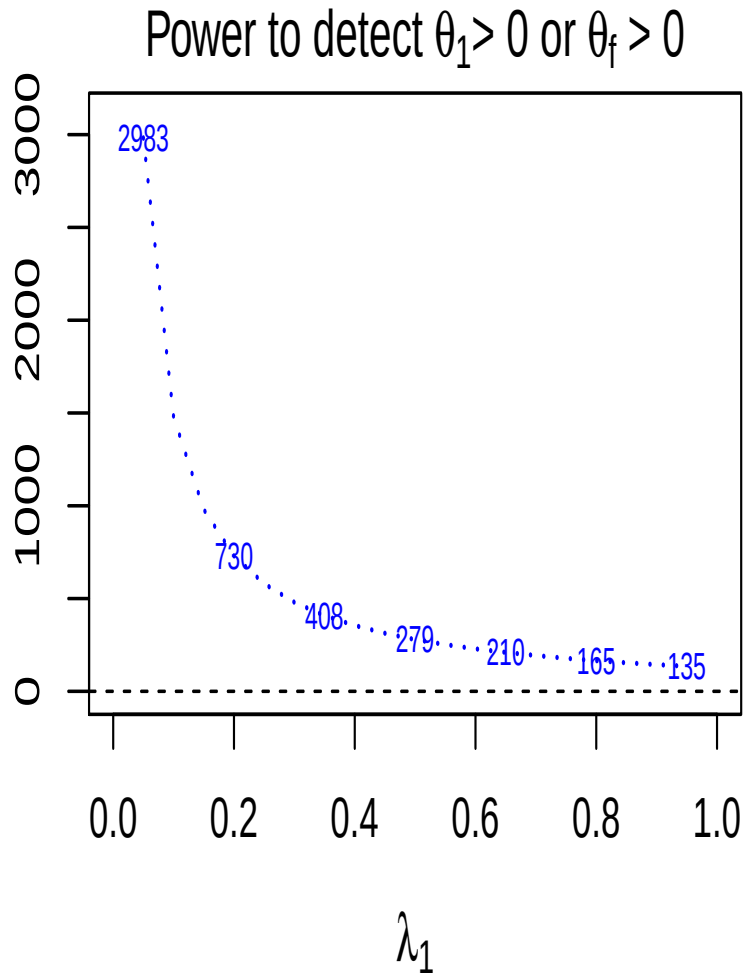
$$1-\beta \leq \sum_{k=1}^{K}\Big[\int\ldots\int_{A_k} p_s^a\Big(\prod_{m=1}^{k} q_{sm}^a dZ_s^{(1)}\ldots dZ_s^{1:m}\Big)\Big], \qquad (3)$$

where

$$\begin{aligned} A_k &= \text{the integral region found based on } (2), \\ p_s^a &= \text{The joint density of } (Z_s^{(1)}, s) \text{ under } H_a, \\ q_{sm}^a &= \text{The conditional density of } Z_s^{1:k} \text{ given its precursor} \\ &\quad Z_s^{1:(k-1)} \text{ at stage } k-1, \text{under } H_a \end{aligned}$$

Figure: The total sample sizes for $F\ (n_f^{(1)})$ across prevalence of $S_1\ (\lambda_1)$ for two different definitions of power. The single-stage design with two subpopulations where the underlying treatment effects $\theta_1 = 0.5$ and $\theta_2 = 0$ for S_1 and S_2, respectively. $\alpha = 0.025$ and $1 - \beta = 80\%$.





In each simulation study, baseline set-ups are that

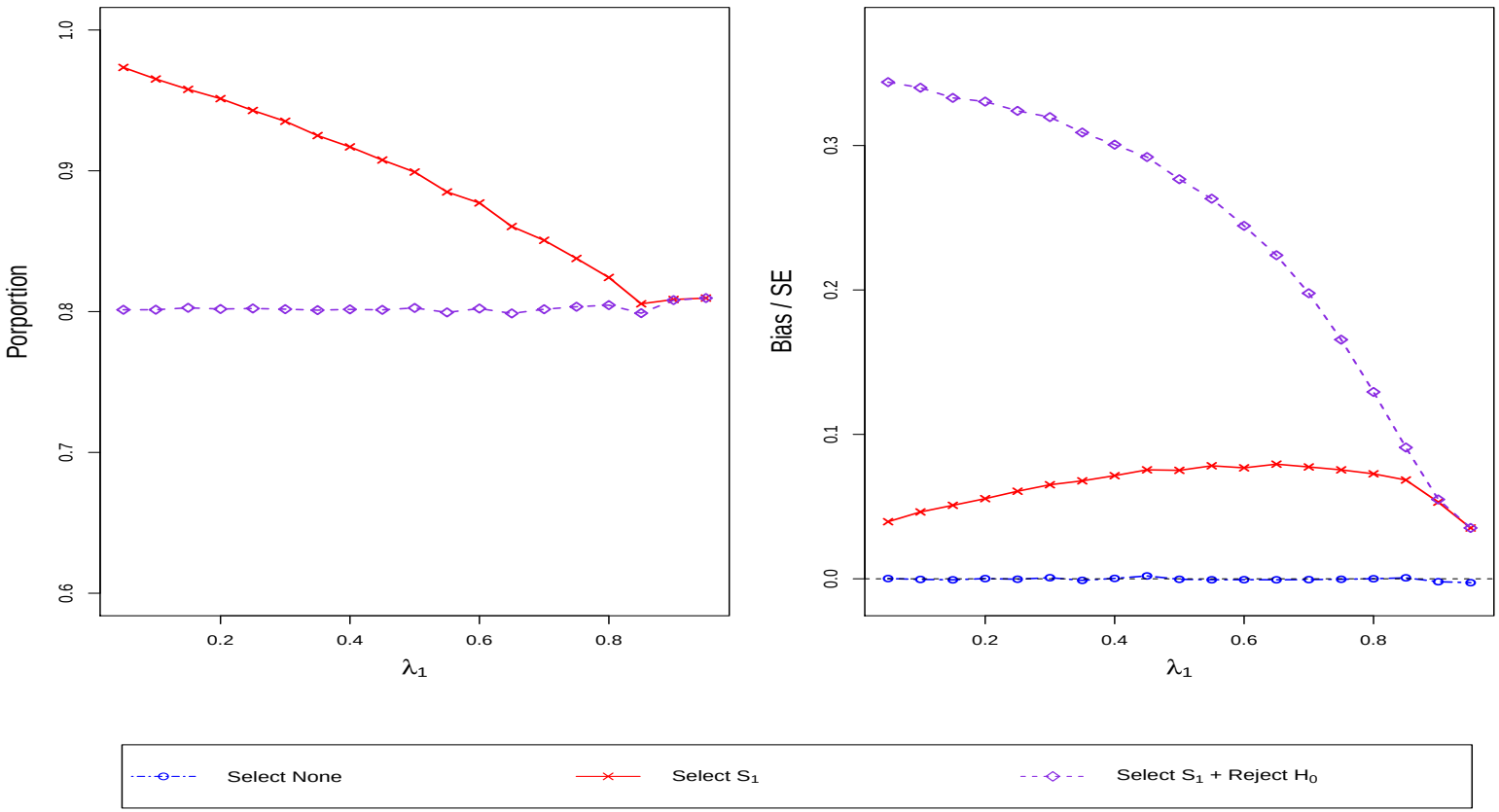
- sample sizes and stopping boundaries used are searched by (2) and (3) under $\alpha = 0.025$ and $1 - \beta = 0.80$,
- power is set to detect the treatment effect of S_1, $\theta_1 = 0.5$,
- The total simulation runs are 100,0000

For treatment effect θ_s, $s \in \mathcal{S}$, estimation assessment are based on,

$$\begin{aligned} \hat{\theta}_s &= Z_s^{1:M} / I_s^{1:M}, \\ Bias(\hat{\theta}_s) &= E(\hat{\theta}_s) - \theta_s, \\ MSE(\hat{\theta}_s) &= E((\hat{\theta}_s - \theta_s)^2). \end{aligned}$$

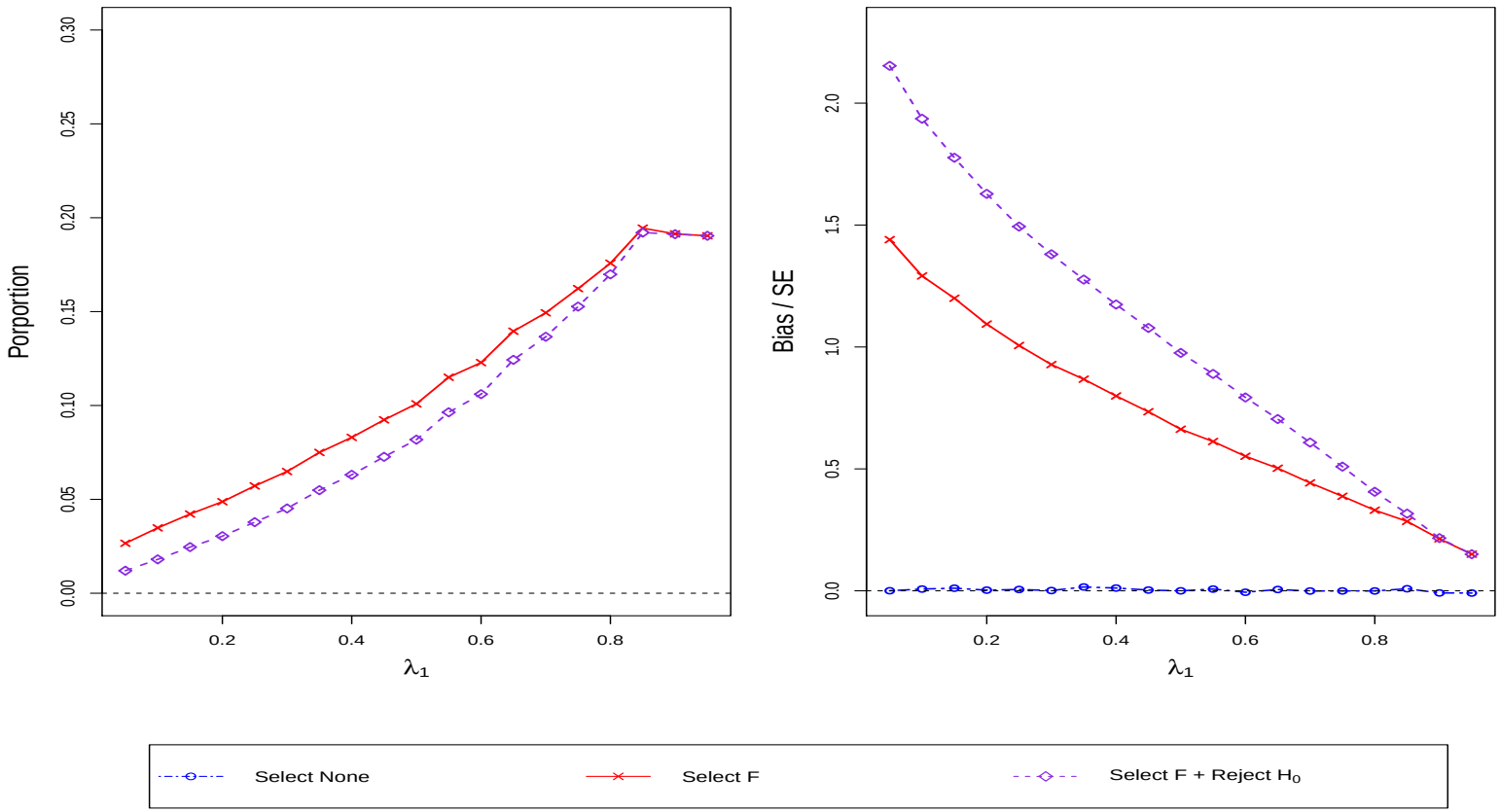
Consider $\mathcal{S} = \{1, f\}$

Figure: ($\theta_1 = 0.5$ and $\theta_2 = 0$) the standardized bias of MLE $\hat{\theta}_1$ and the simulation proportions for different circumstances against the prevalence of S_1, λ_1.



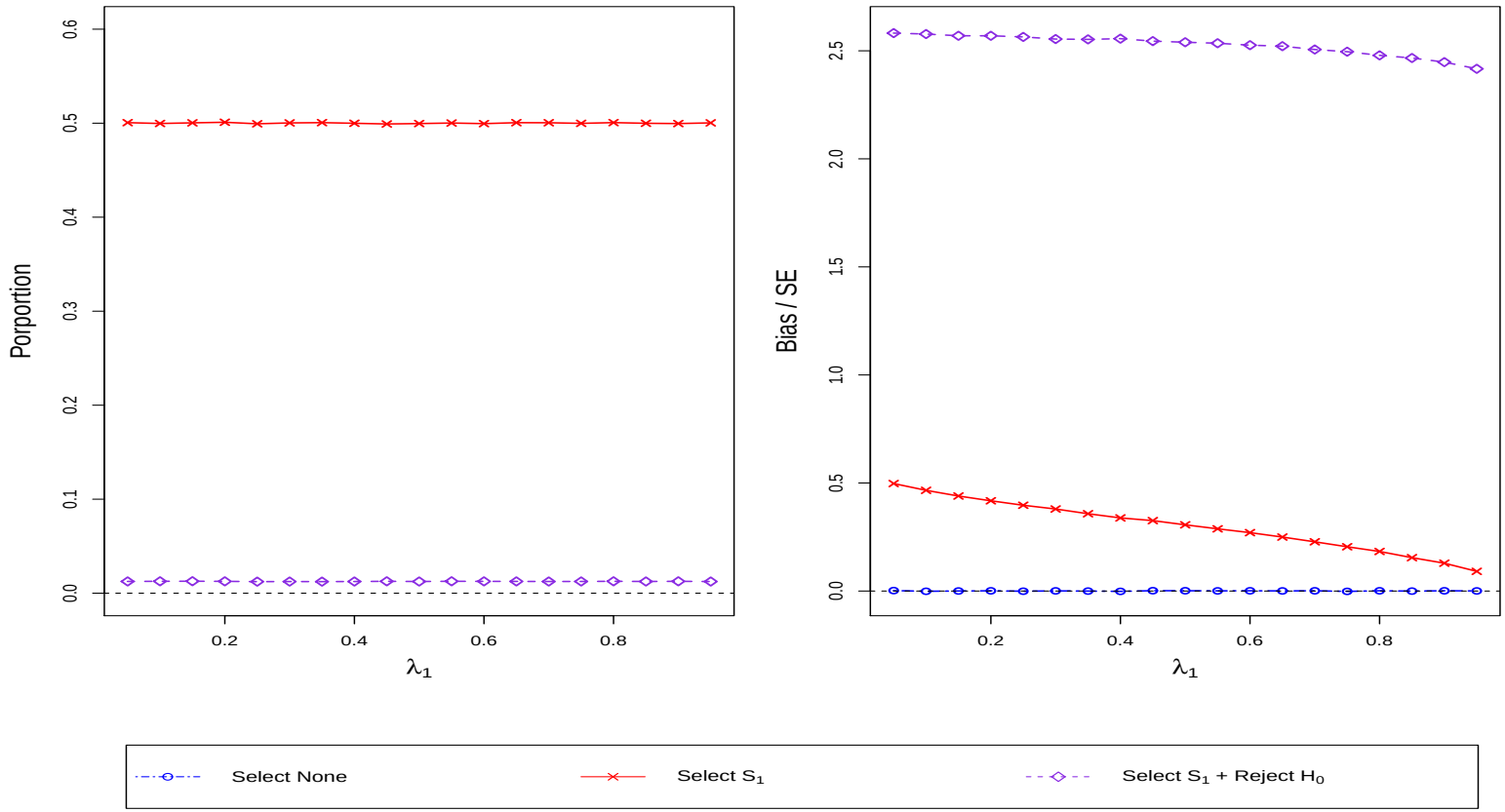
Consider $\mathcal{S} = \{1, f\}$

Figure: ($\theta_1 = 0.5$ and $\theta_2 = 0$) the standardized bias of MLE $\hat{\theta}_f$ and the simulation proportions for different circumstances against the prevalence of S_1, λ_1.



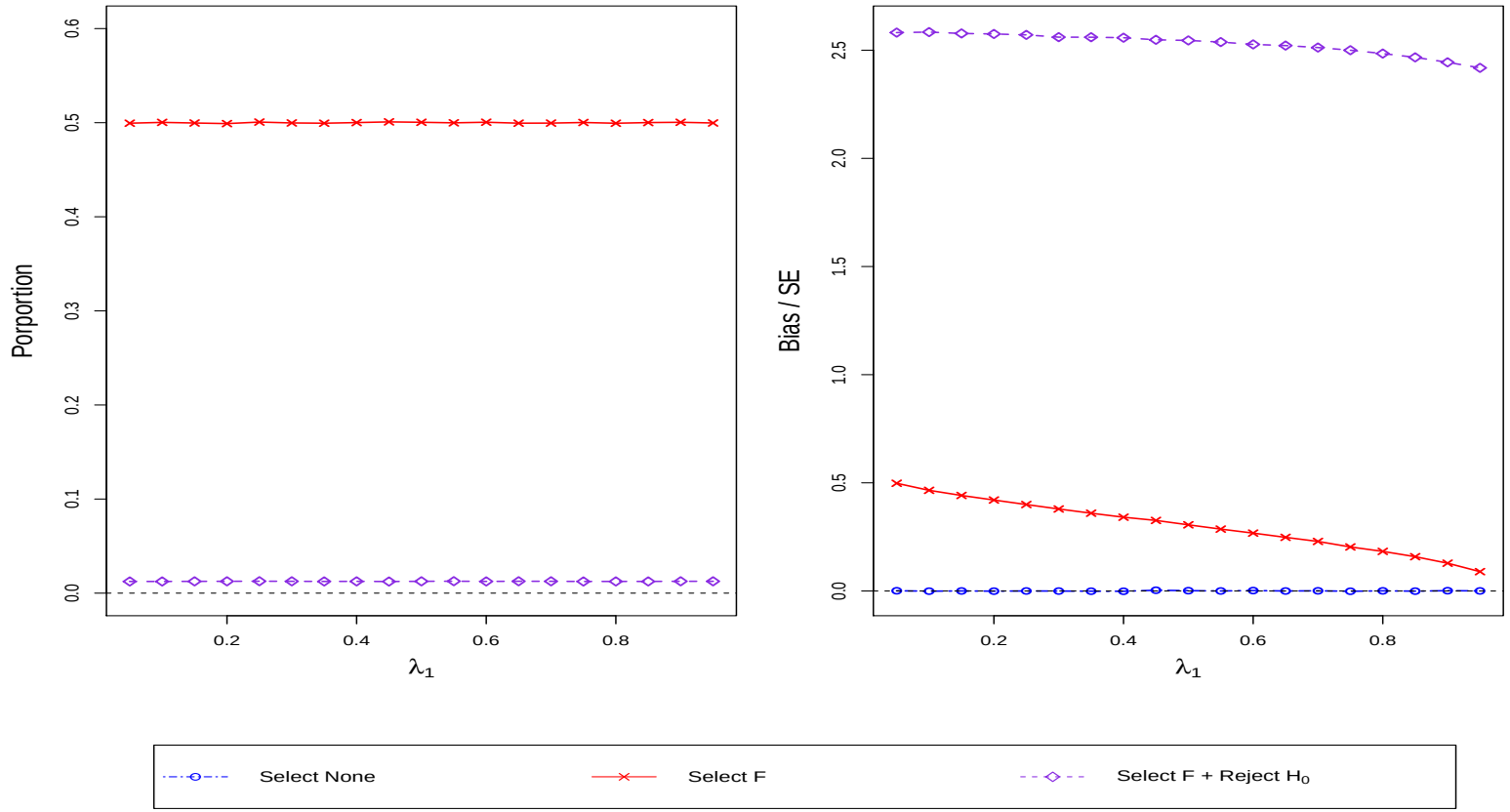
Consider $\mathcal{S} = \{1, f\}$

Figure: ($\theta_1 = 0$ and $\theta_2 = 0$) the standardized biases of MLE $\hat{\theta}_1$ and the simulation proportions for different circumstances against the prevalence of S_1, λ_1.



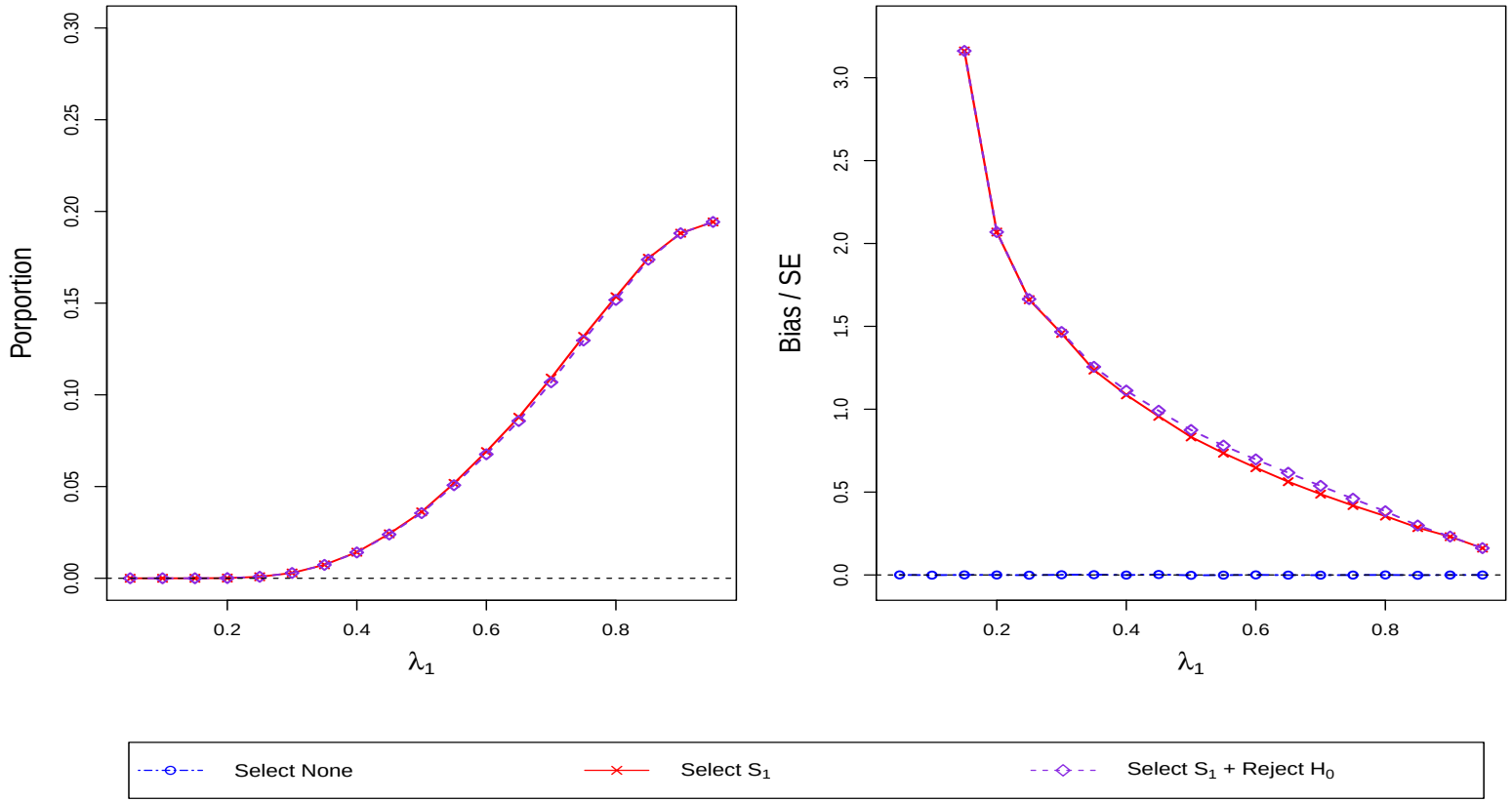
Consider $\mathcal{S} = \{1, f\}$

Figure: ($\theta_1 = 0$ and $\theta_2 = 0$) the standardized biases of MLE $\hat{\theta}_f$ and the simulation proportions for different circumstances against the prevalence of S_1, λ_1.



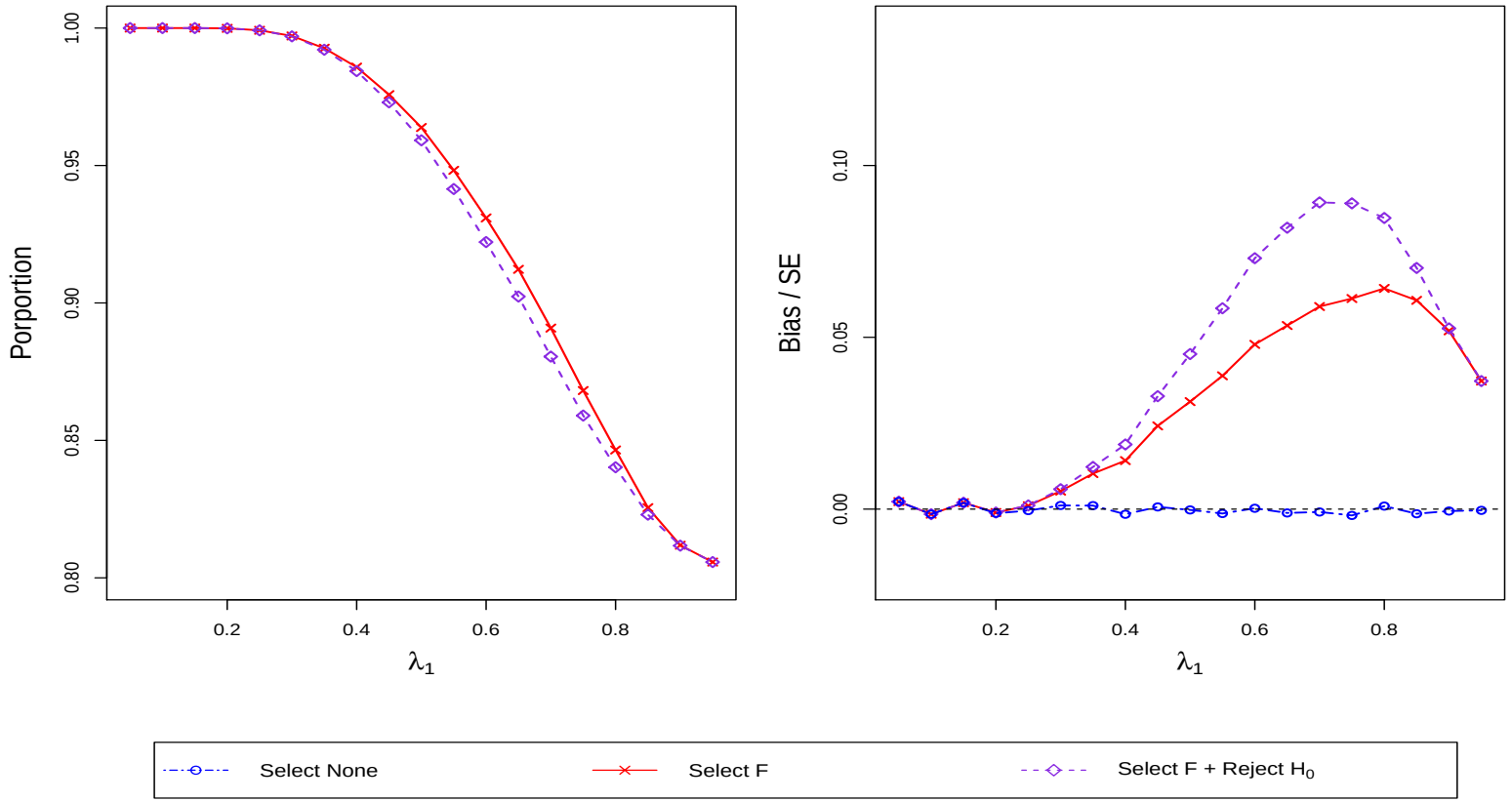
Consider $\mathcal{S} = \{1, f\}$

Figure: ($\theta_1 = 0.5$ and $\theta_2 = 0.5$) the standardized biases of MLE $\hat{\theta}_1$ and the simulation proportions for different circumstances against the prevalence of S_1, λ_1.

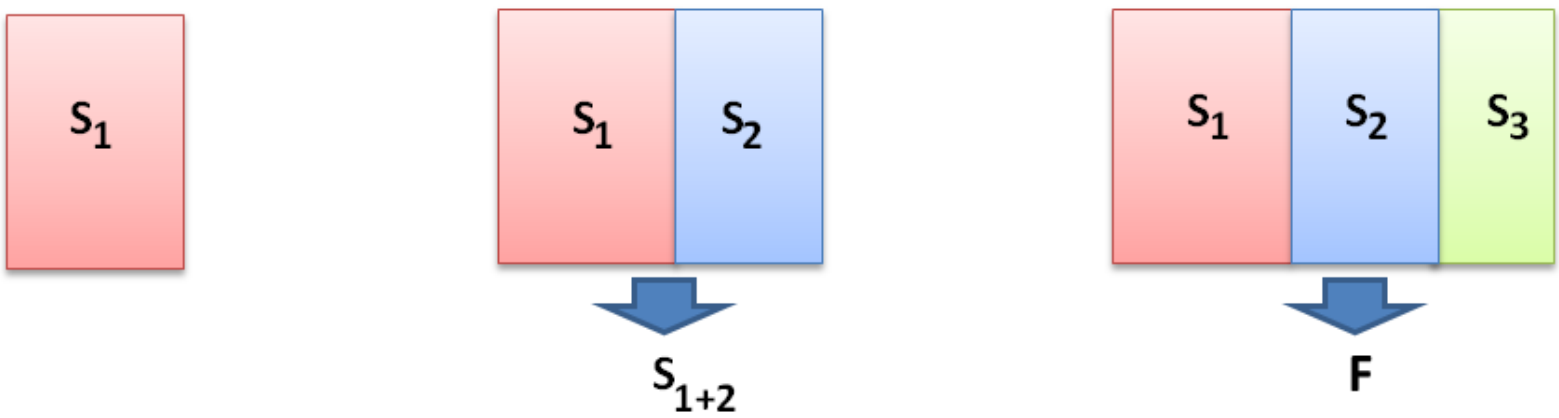


Consider $\mathcal{S} = \{1, f\}$

Figure: ($\theta_1 = 0.5$ and $\theta_2 = 0.5$) the standardized biases of MLE $\hat{\theta}_f$ and the simulation proportions for different circumstances against the prevalence of S_1, λ_1.



Consider $\mathcal{S} = \{1, 1+2, f\}$, that means to select



The selection rule is:

$$\begin{cases} \text{select } S_1 & \text{if } Z_1^{(1)} > \max(Z_f^{(1)}, Z_{1+2}^{(1)}) \\ \text{select } S_{1+2} & \text{if } Z_1^{(1)} \not> \max(Z_f^{(1)}, Z_{1+2}^{(1)}), \text{and } Z_{1+2}^{(1)} > Z_f^{(1)} \\ \text{select } F & \text{if } Z_1^{(k)} \not> \max(Z_f^{(1)}, Z_{1+2}^{(1)}), \text{and } Z_{1+2}^{(1)} < Z_f^{(1)}, \end{cases} \tag{4}$$

Estimation: simulation studies 2

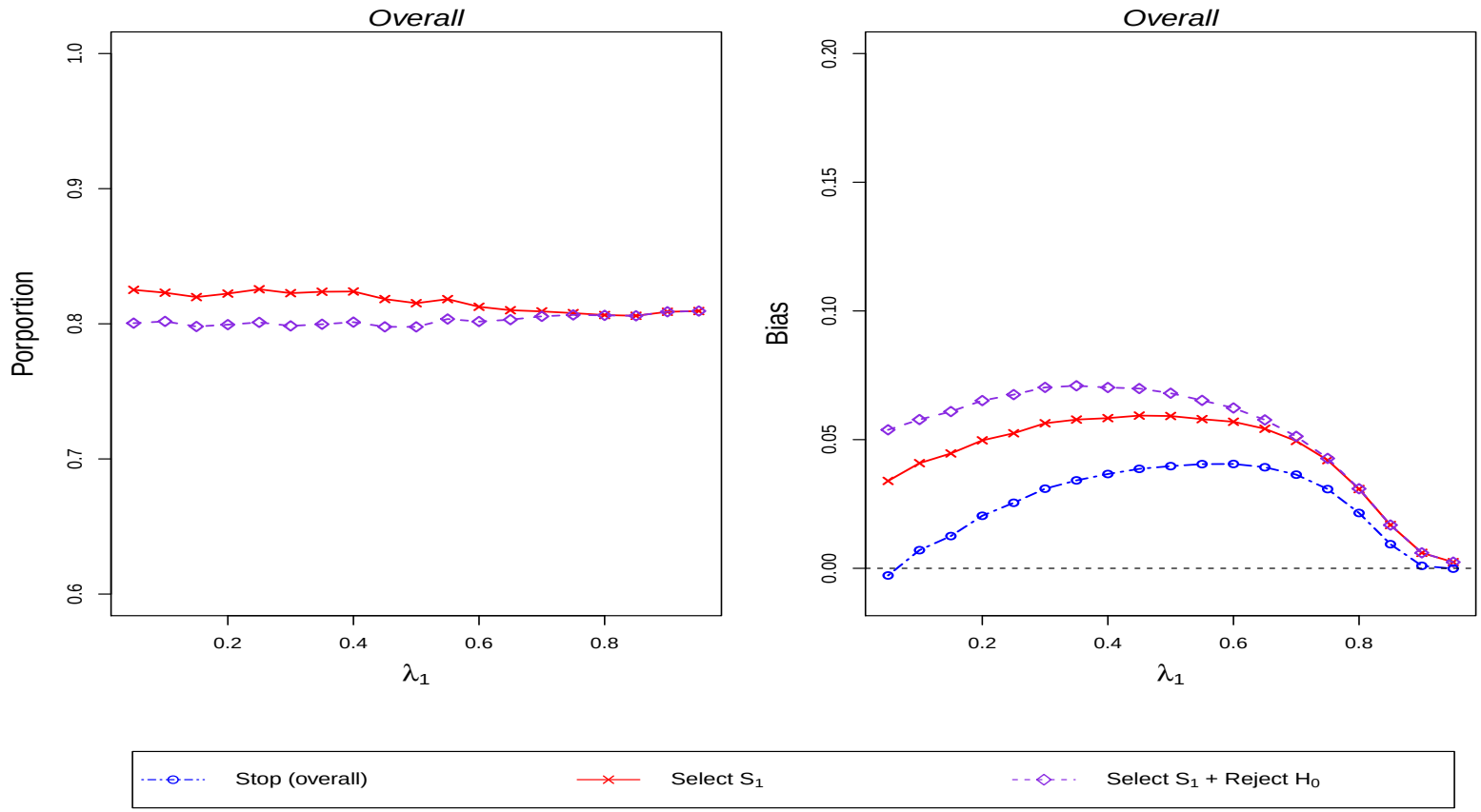
Single-Stage designs with three subpopulations: results 1

Table: ($\theta_1 = 0.5, \theta_2 = 0$ and $\theta_3 = 0$) the standardized biases of the MLEs ($\hat{\theta}_1$, $\hat{\theta}_{1+2}$ and $\hat{\theta}_f$) for different circumstances, where the prevalences of three subgroups are 1/3; proportion (Prop.) stands for how often the corresponding circumstance occurs.

	Prop.(%)	Bias/SE
$\hat{\theta}_f$ (Select None)		-0.00186
$\hat{\theta}_f$ (Select F)	3.74	0.96546
$\hat{\theta}_f$ (Select F + Reject H_0)	2.91	1.31217
$\hat{\theta}_1$ (Select None)		-0.00151
$\hat{\theta}_1$ (Select S_1)	88.58	0.09094
$\hat{\theta}_1$ (Select S_1 + Reject H_0)	80.20	0.27068
$\hat{\theta}_{1+2}$ (Select None)		-0.00118
$\hat{\theta}_{1+2}$ (Select S_{1+2})	7.68	0.76128
$\hat{\theta}_{1+2}$ (Select S_{1+2} + Reject H_0)	6.47	1.02579

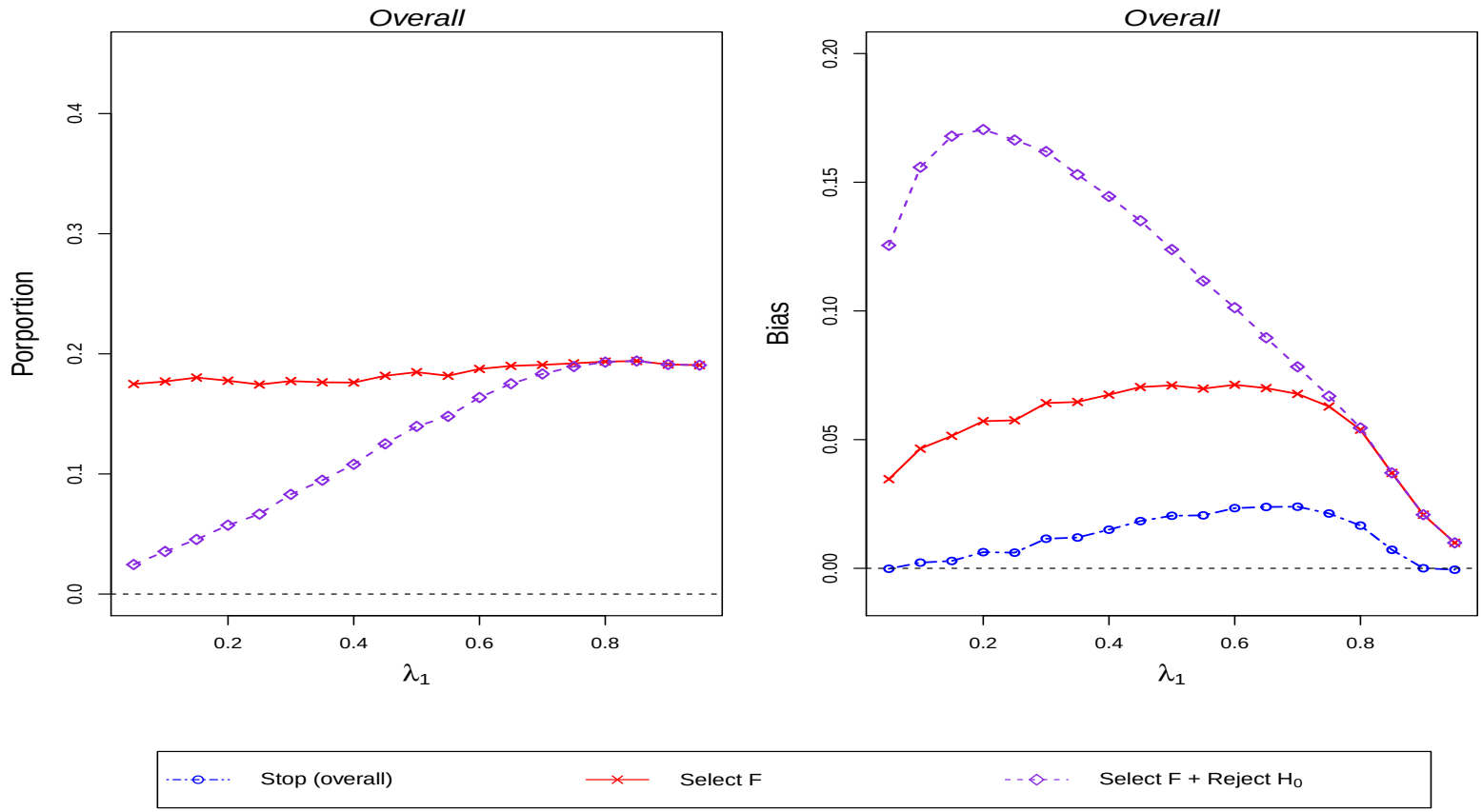
Consider $\mathcal{S} = \{1, f\}$

Figure: ($\theta_1 = 0.5$ and $\theta_2 = 0$) the overall biases of $\hat{\theta}_1$ and the overall simulation proportions for different circumstances while the design stops, against the prevalence of S_1, λ_1.



Consider $\mathcal{S} = \{1, f\}$

Figure: ($\theta_1 = 0.5$ and $\theta_2 = 0$) the overall biases of $\hat{\theta}_f$ and the overall simulation proportions for different circumstances while the design stops, against the prevalence of S_1, λ_1.



Summary and Future work

Summary: bias is always observed in the ML estimate of treatment effect under various design features with population selection.

Future work: develop a unbiased estimator with less MSE.

References

Kimani P et al. (2015) Estimation after subpopulation selection in adaptive seamless trials. *Statistics in Medicine*; 34(18): 2581–2601.

Magnusson B, Turnbull B (2013) Group sequential enrichment design incorporating subgroup selection. *Statistics in Medicine*; 32(16):2695–2754.

Chen & Beckman (2009) Hypothesis testing in a confirmatory phase III trial with a possible subset effect. *Statistics in Biopharmaceutical Research*; 1(4):431–440.

Thank you! :)