Biostatistics (2018), **0**, 0, *pp.* 1–9
doi:10.1093/biostatistics/output

Power analysis in a SMART design: sample size estimation for determining the best embedded dynamic treatment regime supplementary materials

William J. Artman[1*], Inbal Nahum-Shani[2], Tianshuang Wu,

James R. McKay[3], Ashkan Ertefaie[1]

[1] *Department of Biostatistics and Computational Biology, University of Rochester Medical Center, Rochester, NY 14642, USA*

[2] *Institute for Social Research, University of Michigan, Ann Arbor, MI 48106, USA*

[3] *Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA.*

William_Artman@URMC.Rochester.edu

Appendix A.

1. Proof of Theorem 4.1

Note that $\text{Cov}\left(\frac{\sqrt{n}(\hat{\theta}_i - \hat{\theta}_N)}{\sigma_{iN}}, \frac{\sqrt{n}(\hat{\theta}_j - \hat{\theta}_N}{\sigma_{jN}}\right) = \frac{\Sigma_{ij} - \Sigma_{iN} - \Sigma_{jN} + \Sigma_{NN}}{\sigma_{iN}\sigma_{jN}}$.

Assume $\mathbf{\Sigma}$ is exchangeable, e.g., $\mathbf{\Sigma} = \sigma^2 \boldsymbol{I}_N + \rho\sigma^2 \left(\mathbb{1}_N \mathbb{1}_N' - \boldsymbol{I}_N\right)$ where $\mathbb{1}_N$ is a vector of N $1's$ and I_N is the N by N identity matrix.

Then, for all $i, j, i \neq j$, $\text{Cov}\left(W_i, W_j\right) = \text{Cov}\left(\frac{\sqrt{n}(\hat{\theta}_i - \hat{\theta}_N)}{\sigma_{iN}}, \frac{\sqrt{n}(\hat{\theta}_j - \hat{\theta}_N}{\sigma_{jN}}\right) = \frac{\rho\sigma^2 - 2\rho\sigma^2 + \sigma^2}{2\sigma^2(1-\rho)} =$

*To whom correspondence should be addressed.

$\dfrac{\sigma^2(1-\rho)}{2\sigma^2(1-\rho)} = \dfrac{1}{2}$ for all $\rho \in \left(-\dfrac{1}{N-1}, 1\right)$ and for all $\sigma^2 > 0$. Also, $c_{i,1-\alpha}$ is constant across all values of ρ and σ^2.

It follows from monotonicity of the probability measure that

$$\text{Power}_{\alpha,n}\left(\boldsymbol{\Sigma}, \boldsymbol{\Delta}, \Delta_{\min}\right) = \mathbb{P}\left(\bigcap_{i:\Delta_i \geqslant \Delta_{\min}} \left\{W_i < -c_{i,1-\alpha} + \frac{\Delta_i\sqrt{n}}{\sqrt{2\sigma^2(1-\rho)}}\right\}\right)$$

is monotone increasing in ρ and monotone decreasing in σ^2.

2. Simulation Study and EXTEND

In this section, we give additional details on how estimation was performed in the simulation studies and for EXTEND. The SMART designs in the simulation studies and the estimation procedures are based off those in Ertefaie *and others* (2015). We estimated $\boldsymbol{\theta}$ and $\boldsymbol{\Sigma}$ using AIPW.

For SMART design 1, the MSM is $m(\mathcal{T};\beta) = \beta_0 + \beta_1 A_1 + \beta_2 A_2^{\text{NR}}$. The conditional means are:

$$\mathbb{E}[Y \mid \bar{A}_2 = \text{EDTR}_k^V, \bar{O}_2] = \gamma_0 + \gamma_1 o_{11} + \gamma_2 o_{12} + \gamma_3 o_{21} + \gamma_4 o_{22} + a_1(\gamma_5 + \gamma_6 o_{11}) + \gamma_7 I(o_{21} < 0))a_2$$

$$\mathbb{E}[Y \mid A_1 = \text{EDTR}_{k,1}, O_1] = \gamma_8 + \gamma_9 o_{11} + \gamma_{10} o_{12} + \gamma_{11} a_1 + \gamma_{12} a_1 o_{11}$$

The true $\boldsymbol{\beta}$ is approximately $(1.499, 0.251, 0.052)$. Then, $\hat{\boldsymbol{\theta}}_{\text{AIPW}} = \boldsymbol{D}\hat{\boldsymbol{\beta}}_{\text{AIPW}} = (1.802, 1.300, 1.699, 1.197)$ where

$$\boldsymbol{D} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & -1 \end{pmatrix}$$

The rows of $\boldsymbol{D}$ correspond to each of the four EDTRs listed in Table 2.

For SMART design 2, the MSM is:

$$m(\mathcal{T};\boldsymbol{\beta}) = \beta_0 + \beta_1 I(A_1 = -1) + I(A_1 = -1)[\beta_2 I(A_2^B = 1) + \beta_3 I(A_2^B = 2) + \beta_4 I(A_2^B = 3)].$$

The conditional means are:

$$\mathbb{E}[Y \mid \bar{a}_2 = \text{EDTR}_k^V, \bar{o}_2, \boldsymbol{\gamma}] = \gamma_0 + \gamma_1 o_{11} + \gamma_2 o_{12} + \gamma_3 o_{21} + \gamma_4 o_{22} + I(a_1 = -1)(\gamma_5 + \gamma_6 o_{11})$$
$$+ I(o_{22} < 0) I(a_1 = -1)[\gamma_7 I(a_2 = 1) + \gamma_8 I(a_2 = 2) + \gamma_9 I(a_2 = 3) + \gamma_{10} o_{21} I(a_2 = 2)]$$

$$\mathbb{E}[Y \mid a_1 \in \text{EDTR}_{k,1}, o_1, \boldsymbol{\gamma}] = \gamma_{11} + \gamma_{12} o_{11} + \gamma_{13} o_{12} + \gamma_{14} I(a_1 = -1) + \gamma_{15} I(a_1 = -1) o_{11}$$

The true $\boldsymbol{\beta}_{\text{AIPW}}$ value is approximately $(1.500, 2.001, -0.249, 0.750, 0.000)$. Then, $\hat{\boldsymbol{\theta}}_{\text{AIPW}} = \boldsymbol{D}\hat{\boldsymbol{\beta}}_{\text{AIPW}} = (1.500, 3.501, 3.251, 4.251, 3.501)$ where

$$\boldsymbol{D} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

The rows of $\boldsymbol{D}$ correspond to each of the five EDTRs listed in Table 3.

For EXTEND, we considered the MSM $m(\mathcal{T}; \boldsymbol{\beta}) = \beta_0 + \beta_1 A_1 + \beta_2 A_2^{\text{R}} + \beta_3 A_3^{\text{NR}}$. Let r denote the indicator of response to the initial NTX treatment. The conditional means are

$$\mathbb{E}[Y \mid \bar{a}_2 = \text{EDTR}_k^V, \bar{o}_2, \boldsymbol{\gamma}] = \gamma_0 + \gamma_1 o_{11} + \gamma_2 o_{12} + \gamma_3 o_{21} + \gamma_4 o_{22} + a_1(\gamma_5 + \gamma_6 o_{11}) + r a_2(\gamma_7 + \gamma_8 o_{21})$$
$$+ (1 - r) a_2 (\gamma_9 + \gamma_{10} o_{21})$$

$$\mathbb{E}[Y \mid a_1 \in \text{EDTR}_{k,1}, o_1, \boldsymbol{\gamma}] = \gamma_{11} + \gamma_{12} o_{11} + \gamma_{13} o_{12} + \gamma_{14} a_1 + \gamma_{15} a_1 o_{11} + \gamma_{16} a_1 o_{12}$$

The estimated $\hat{\boldsymbol{\beta}}_{\text{IPW}}$ is $(8.86, -0.99, -0.24, -0.07)$ and the estimated $\hat{\boldsymbol{\beta}}_{\text{AIPW}}$ is $(8.84, -0.90, -0.09, -0.21)$. Then, $\hat{\boldsymbol{\theta}}_{\text{IPW}} = \boldsymbol{D}\hat{\boldsymbol{\beta}}_{\text{IPW}} = (7.56, 9.53, 8.05, 10.02, 7.71, 9.68, 8.19, 10.17)$ and $\hat{\boldsymbol{\theta}}_{\text{AIPW}} = \boldsymbol{D}\hat{\boldsymbol{\beta}}_{\text{AIPW}} = (7.65, 9.44, 7.83, 9.62, 8.06, 9.85, 8.24, 10.03)$ where

$$\boldsymbol{D} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & -1 \end{pmatrix}$$

The rows of $\boldsymbol{D}$ correspond to each of the eight EDTRs listed in Table 4.

Appendix B.

3. MCB vs. Alternative Approaches

We now compare sizing a SMART using MCB with sizing for detecting a difference in pairwise comparisons. We also show that the required sample size for sizing SMARTs based off pairwise comparisons is sensitive to the choice of covariance matrix $\boldsymbol{\Sigma}$.

Ogbagaber *and others* (2016) proposed a method for sizing SMARTs to detect differences when performing pairwise comparisons while adjusting for multiple comparisons using the Bonferroni correction. For each pairwise comparison, the sample size is computed so that a difference can be detected with probability $1-\beta$ according to the following equation:

$$n = \frac{\sigma_{ij}^2 (z_{1-\alpha/2g} + z_{1-\beta})^2}{\Delta_{ij}^2} \tag{3.1}$$

where $\sigma_{ij}^2 = \mathrm{Var}(\hat{\theta}_i - \hat{\theta}_j)$, z_q is the q-quantile of the normal distribution, g is the number of pairwise comparisons, and $\Delta_{ij} = \theta_i - \theta_j$. The sample size is then chosen to be the maximum over all pairs i, j of Equation 3.1.

We consider two SMARTs: 1) simulation design 1 with four EDTRs; 2) simulation design 2 with five EDTRs. The true mean EDTR outcome vector for simulation design 1 is $(1.802, 1.300, 1.699, 1.197)$, and for design 2 is $(1.500, 3.501, 3.251, 4.251, 3.501)$.

For design 1 (4 EDTRs), we consider all pairwise comparisons in which there is a true difference of 0.1 or more. This is all pairwise comparisons. For design 2 we consider all pairwise comparisons such that they have true difference 0.7 or more: (1,2), (1,3), (1,4), (1,5), (2,4), (3,4), (4,5).

For MCB in design 1, the power is the probability of excluding from the set of best all EDTRs with outcome 0.1 or more away from the best. For MCB in design 2, the power is the probability of excluding from the set of best all EDTRs with outcome 0.7 or more away from the best. We show in Table 1 the sample size to achieve $1-\beta = 80\%$ power for each of the pairwise comparisons as well as the required sample size to achieve 80% power when using our MCB based methodology. We see that the required sample sizes to power a SMART are lower for MCB than for pairwise comparisons and that the pairwise comparisons based approach is also sensitive to the covariance matrix $\mathbf{\Sigma}$. Note the large sample sizes required by both methods is because we are powering to detect very small effect sizes.

4. Pilot SMARTs

In this section we discuss how to use data from a pilot SMART to estimate the correlations in the covariance matrix, $\mathbf{\Sigma}$, which can be used for sizing a full-scale SMART. We simulated pilot SMARTs with 50 individuals for the two simulation designs presented in the main manuscript.

We constructed confidence intervals for the power by taking 1000 bootstrap samples of size 50 from the simulated pilot SMART data and re-estimated the covariance matrix for each bootstrapped pilot SMART using AIPW. For the known variance case (diagonal elements of $\mathbf{\Sigma}$ known) the covariance matrix is obtained by transforming the AIPW estimated covariance matrices to correlation matrices, and then transforming back to covariance matrices by left and right multiplying the correlation matrices by the diagonal matrix with entries consisting of the square root of the variances. For the unknown variance covariance matrix, the unstructured covariance matrix is computed for each bootstrapped pilot SMART (both the diagonal and off diagonal elements) using AIPW. Then, the power is computed across a grid of sample sizes for each of the bootstrapped covariance matrices to obtain the bootstrap power confidence intervals. The results are shown in Figure 3. When the variances are known, the proposed bootstrap algorithm yields accurate power predications. When the variances are unknown and estimated from the pilot SMART, the bootstrapped confidence intervals yield similar power to using a conservative estimate of the covariance matrix for design 1 (diagonal matrix with known variances). For design

2 when the variances are unknown, there are some gains in estimated power compared to using a conservative estimate of the covariance matrix.

References

Ertefaie, A., Wu, T., Lynch, K.G. and Nahum-Shani, I. (2015). Identifying a set that contains the best dynamic treatment regimes. *Biostatistics* **17**(1), 135–148.

Ogbagaber, S.B., Karp, J. and Wahed, A.S. (2016). Design of sequentially randomized trials for testing adaptive treatment strategies. *Statistics in medicine* **35**(6), 840–858.

Covariance Matrix	Pairwise		MCB	
	4 EDTRs	5 EDTRs	4 EDTRs	5 EDTRs
$\boldsymbol{\Sigma}_{\text{True}}$	1910	280	1274	247
$\boldsymbol{\Sigma}_{\text{Conservative}}$	24357	894	16730	786

Table 1. Sample size to achieve power $1-\beta = 80\%$ for pairwise comparisons and for MCB. $\boldsymbol{\Sigma}_{\text{True}}$ is the true covariance matrix. $\boldsymbol{\Sigma}_{\text{Conservative}}$ is a diagonal covariance matrix with all variances set to the true variances. Note that the required sample sizes for pairwise comparisons are higher than those for our MCB based methodology.

EDTR	Decision Rule
$\text{EDTR}_1 : (+1, +1)$	start with $A_1 = +1$. If the patient is non-responsive, take $A_2 = +1$; if the patient is responsive, continue with $A_1 = +1$
$\text{EDTR}_2 : (-1, +1)$	start with $A_1 = -1$. If the patient is non-responsive, take $A_2 = +1$; if the patient is responsive, continue with $A_1 = -1$
$\text{EDTR}_3 : (+1, -1)$	start with $A_1 = +1$. If the patient is non-responsive, take $A_2 = -1$; if the patient is responsive, continue with $A_1 = +1$
$\text{EDTR}_4 : (-1, -1)$	start with $A_1 = -1$. If the patient is non-responsive, take $A_2 = -1$; if the patient is responsive, continue with $A_1 = -1$

Table 2. SMART design 1: EDTR

EDTR	Decision Rule
$\text{EDTR}_1 : (1, 0)$	start with $A_1 = +1$. If the patient is non-responsive, intensify A_1; if the patient is responsive, continue with $A_1 = +1$
$\text{EDTR}_2 : (-1, 4)$	start with $A_1 = -1$. If the patient is non-responsive, take $A_2 = 4$; if the patient is responsive, continue with $A_1 = -1$
$\text{EDTR}_3 : (-1, 1)$	start with $A_1 = -1$. If the patient is non-responsive, take $A_2 = 1$; if the patient is responsive, continue with $A_1 = -1$
$\text{EDTR}_4 : (-1, 2)$	start with $A_1 = -1$. If the patient is non-responsive, take $A_2 = 2$; if the patient is responsive, continue with $A_1 = -1$
$\text{EDTR}_5 : (-1, 3)$	start with $A_1 = -1$. If the patient is non-responsive, take $A_2 = 3$; if the patient is responsive, continue with $A_1 = -1$

Table 3. SMART design 2: EDTR

EDTR	Decision Rule
$\text{EDTR}_1 : (+1, +1, +1)$	start with $A_1 = +1$. If the patient is non-responsive, take $A_2 = +1$; if the patient is responsive, take $A_2 = +1$
$\text{EDTR}_2 : (-1, +1, +1)$	start with $A_1 = -1$. If the patient is non-responsive, take $A_2 = +1$; if the patient is responsive, take $A_2 = +1$
$\text{EDTR}_3 : (+1, -1, +1)$	start with $A_1 = +1$. If the patient is non-responsive, take $A_2 = +1$; if the patient is responsive, take $A_2 = -1$
$\text{EDTR}_4 : (-1, -1, +1)$	start with $A_1 = -1$. If the patient is non-responsive, take $A_2 = +1$; if the patient is responsive, take $A_2 = -1$
$\text{EDTR}_5 : (+1, +1, -1)$	start with $A_1 = +1$. If the patient is non-responsive, take $A_2 = -1$; if the patient is responsive, take $A_2 = +1$
$\text{EDTR}_6 : (-1, +1, -1)$	start with $A_1 = -1$. If the patient is non-responsive, take $A_2 = -1$; if the patient is responsive, take $A_2 = +1$
$\text{EDTR}_7 : (+1, -1, -1)$	start with $A_1 = +1$. If the patient is non-responsive, take $A_2 = -1$; if the patient is responsive, take $A_2 = -1$
$\text{EDTR}_8 : (-1, -1, -1)$	start with $A_1 = -1$. If the patient is non-responsive, take $A_2 = -1$; if the patient is responsive, take $A_2 = -1$

Table 4. EXTEND trial: EDTR

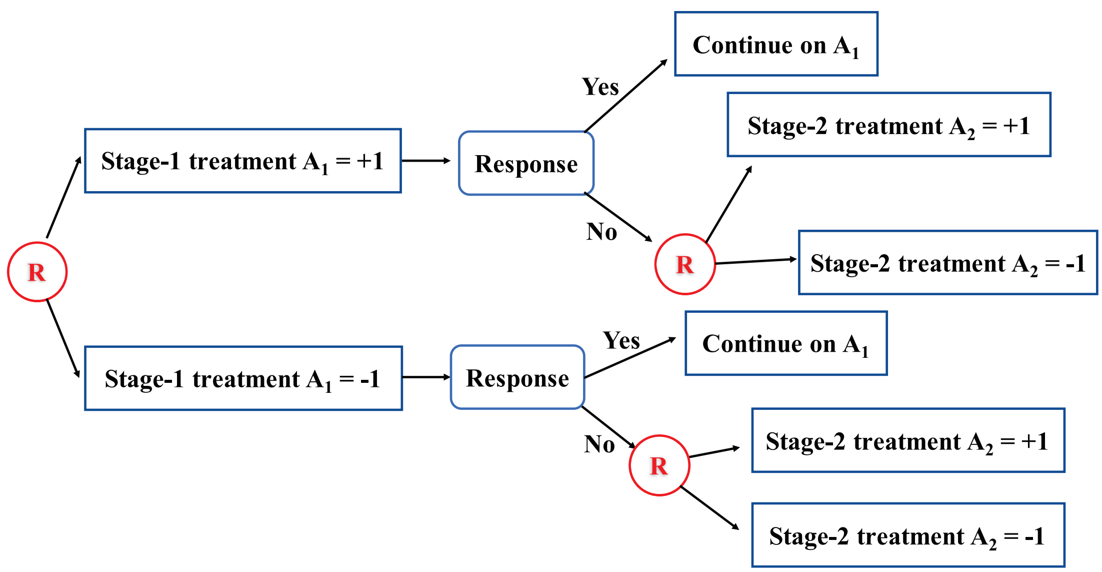

Fig. 1. SMART simulation design 1 with four EDTRs

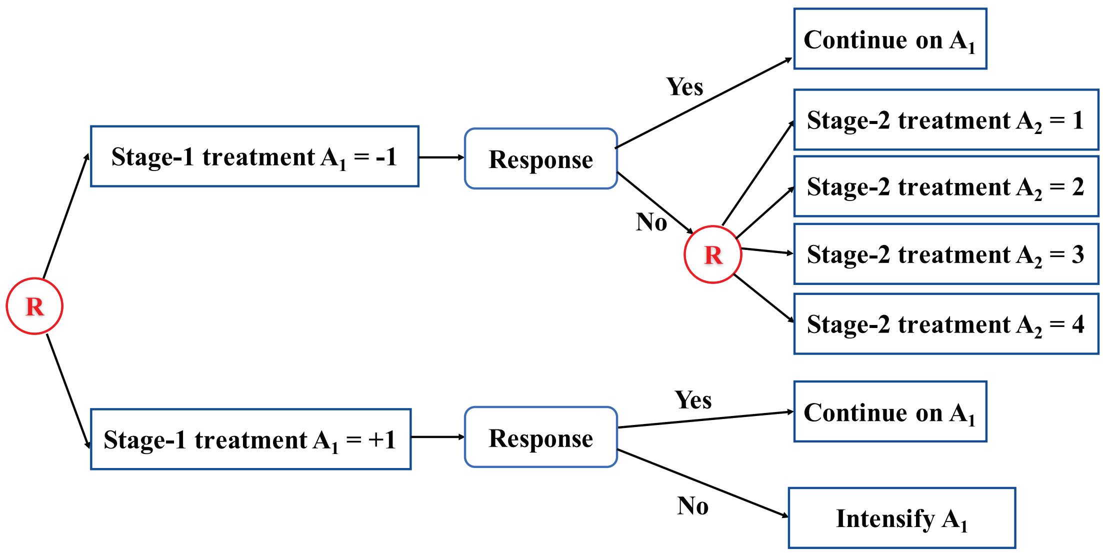

Fig. 2. SMART simulation design 2 with five EDTRs

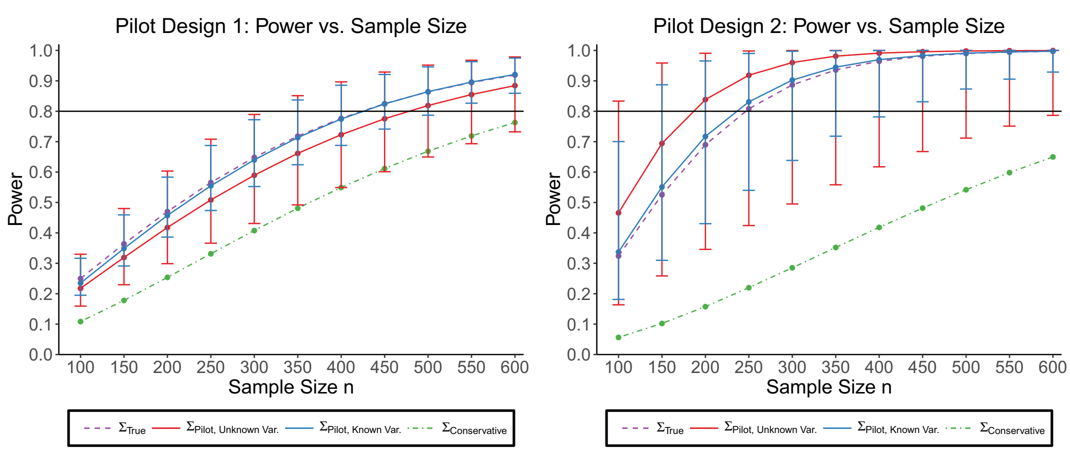

Fig. 3. Power by sample size computed using estimates of $\mathbf{\Sigma}$ obtained from pilot SMARTs with 50 individuals. $\mathbf{\Sigma}_{\text{Pilot, Unknown Var.}}$ is the covariance matrix when estimating both diagonal elements and correlations in $\mathbf{\Sigma}$ from the pilot SMART, $\mathbf{\Sigma}_{\text{Pilot, Known Var.}}$ is the covariance matrix when estimating only the correlations and the variances are known. $\mathbf{\Sigma}_{\text{Conservative}}$ is the diagonal covariance matrix with known variances, and $\mathbf{\Sigma}_{\text{True}}$ is the true covariance matrix. The error bars represent 95% bootstrap confidence intervals.