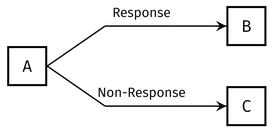
Sample size considerations for comparing dynamic treatment regimens in a sequential multiple-assignment randomized trial with a continuous longitudinal outcome

Nicholas J. Seewald[1] Kelley M. Kidwell[2] Inbal Nahum-Shani[3] Tianshuang Wu[4] James R. McKay[5] Daniel Almirall [1,3]

[1]Department of Statistics, University of Michigan [2]Department of Biostatistics, University of Michigan [3]Survey Research Center, Institute for Social Research, University of Michigan
[4]AbbVie [5]Department of Psychiatry, University of Pennsylvania

Dynamic Treatment Regimens

A **dynamic treatment regimen (DTR)** is a sequence of pre-specified decision rules which guides the delivery of an individualized sequence of treatments. This sequence is tailored based on ongoing information about the individual's progress in treatment.



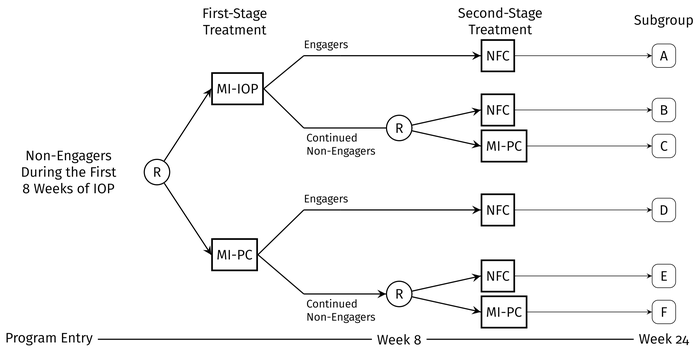
Sequential Multiple-Assignment Randomized Trials

A **sequential multiple-assignment randomized trial (SMART)** is an experimental design which can provide data that informs the construction of an effective DTR (Murphy, 2005). Some or all participants are randomized more than once. Each randomization corresponds to a critical question regarding the development of a DTR.
We consider two-stage SMARTs in which the primary outcome is continuous and repeatedly measured in participants over the course of the study.

The ENGAGE Trial

The ENGAGE study (J. McKay, PI; N = 500) is a SMART aimed at developing a DTR to increase motivation to engage in treatment among alcohol- and cocaine-dependent patients.

Figure 1: Diagram of the ENGAGE SMART. Circled R indicates randomization, boxes indicate treatments. MI-IOP refers to two phone-based sessions encouraging participation in an intensive outpatient program; MI-PC, two phone-based sessions offering patients a choice of treatment modalities; NFC is no further contact.



- The outcome of interest is **treatment readiness**, a measure of a patient's willingness and ability to commit to active participation in a substance abuse treatment program.
- Treatment readiness was assessed using an 8-item questionnaire scored from 0 to 40 and coded such that higher scores are better. We consider measurements taken at baseline and at weeks 8 and 24.
- There are 4 **embedded DTRs**, indexed by recommended first-stage treatment a_1 and recommended second-stage treatment for continued non-engagers, a_2.

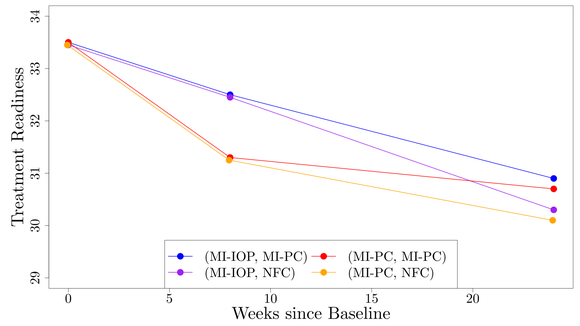
		Stage 2 Treatment		
(a_1, a_2)	Stage 1 Treatment	Engagers	Ctd. Non-Engagers	Subgroups
$(1,1)$	MI-IOP	NFC	MI-PC	A, C
$(1,-1)$	MI-IOP	NFC	NFC	A, B
$(-1,1)$	MI-PC	NFC	MI-PC	D, F
$(-1,-1)$	MI-PC	NFC	NFC	D, E

Marginal Mean Model

We are interested in $\mathrm{E}[Y_t^{(a_1,a_2)} \mid \boldsymbol{X}]$, the marginal mean of $\boldsymbol{Y}^{(a_1,a_2)}$ at time t under DTR (a_1, a_2) conditional on $\boldsymbol{X}$. This is the mean outcome at time t had all individuals with characteristics $\boldsymbol{X}$ been offered DTR (a_1, a_2).

- We impose a modeling assumption:
$$E[Y_t^{(a_1,a_2)} \mid \boldsymbol{X}] = \mu_t^{(a_1,a_2)}(\boldsymbol{X}; \boldsymbol{\theta}),$$
where $\mu_t^{(a_1,a_2)}(\boldsymbol{X}; \boldsymbol{\theta})$ is a **marginal structural mean model** with unknown parameters $\boldsymbol{\theta} = (\boldsymbol{\eta}, \boldsymbol{\gamma})$.
- $\mu_t^{(a_1,a_2)}(\boldsymbol{X}; \boldsymbol{\theta})$ should account for the design of the SMART.
- An example model for ENGAGE is
$$\begin{aligned}\mu_t^{(a_1,a_2)}(\boldsymbol{X}; \boldsymbol{\theta}) = \boldsymbol{\eta}^\top \boldsymbol{X} &+ \gamma_0 + \mathbb{1}_{\{t \le 8\}}\left(\gamma_1 t + \gamma_2 a_1 t\right) \\ &+ \mathbb{1}_{\{t>8\}}\left(8\gamma_1 + 8\gamma_2 a_1 + \gamma_3(t-8) + \gamma_4(t-8)a_1 \right. \\ &\left. + \gamma_5(t-8)a_2 + \gamma_6(t-8)a_1 a_2\right),\end{aligned}$$
where $\mathbb{1}_{\{E\}}$ is the indicator function for the event E.

Figure 2: Plot of treatment readiness vs. time using data from ENGAGE.



Estimation of Model Parameters

The estimate $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ is the solution to the following the estimating equations:

Estimating Equations

$$\mathbf{0} = \frac{1}{n}\sum_{i=1}^{n}\sum_{(a_1,a_2)}\left[W^{(a_1,a_2)}\left(A_{1,i}, R_i, A_{2,i}\right) \cdot \boldsymbol{D}^{(a_1,a_2)}(\boldsymbol{X}_i)^\top \boldsymbol{V}^{(a_1,a_2)}(\boldsymbol{X}_i)^{-1}\left(\boldsymbol{Y}_i - \boldsymbol{\mu}^{(a_1,a_2)}(\boldsymbol{X}_i; \boldsymbol{\theta})\right)\right]$$

where

- $\left(a_1, a_2\right)$ specifies an embedded DTR,
- $W^{(a_1,a_2)}\left(A_{1,i}, R_i, A_{2,i}\right) = 2 \cdot \mathbb{1}\left\{A_{1,i} = a_1\right\}\left(R_i + 2(1-R_i)\mathbb{1}\left\{A_{2,i} = a_2\right\}\right)$
- $\boldsymbol{D}^{(a_1,a_2)}(\boldsymbol{X}_i) = \frac{\partial}{\partial \boldsymbol{\theta}}\boldsymbol{\mu}^{(a_1,a_2)}(\boldsymbol{X}_i; \boldsymbol{\theta})$
- $\boldsymbol{V}^{(a_1,a_2)}\left(\boldsymbol{X}_i\right)$ is a working model for $\mathrm{Var}\left(\boldsymbol{Y}^{(a_1,a_2)} - \boldsymbol{\mu}^{(a_1,a_2)}(\boldsymbol{X}_i; \boldsymbol{\theta}) \mid \boldsymbol{X}_i\right)$

Assuming that $\boldsymbol{\mu}^{(a_1,a_2)}(\boldsymbol{X}_i; \boldsymbol{\theta})$ is correctly specified, $\hat{\boldsymbol{\theta}}$ is consistent for the true parameter value, regardless of the choice of $\boldsymbol{V}^{(a_1,a_2)}\left(\boldsymbol{X}_i\right)$ (Lu et al., 2016). Under usual regularity conditions for M-estimators and given data from a SMART, $\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ has an asymptotic multivariate normal distribution: $\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \Rightarrow \mathcal{N}(\mathbf{0}, \boldsymbol{B}^{-1}\boldsymbol{M}\boldsymbol{B}^{-1})$, where

- $\boldsymbol{B} := \mathrm{E}\left[\sum_{(a_1,a_2)} W^{(a_1,a_2)}\left(A_{1,i}, R_i, A_{2,i}\right)\boldsymbol{D}^{(a_1,a_2)}(\boldsymbol{X}_i)^\top \boldsymbol{V}^{(a_1,a_2)}(\boldsymbol{X}_i)^{-1}\boldsymbol{D}^{(a_1,a_2)}(\boldsymbol{X}_i)\right]$
- $\boldsymbol{M} := \mathrm{E}\left[\left(\sum_{(a_1,a_2)} W^{(a_1,a_2)}\left(A_{1,i}, R_i, A_{2,i}\right)\boldsymbol{D}^{(a_1,a_2)}(\boldsymbol{X}_i)^\top \boldsymbol{V}^{(a_1,a_2)}(\boldsymbol{X}_i)^{-1}\left(\boldsymbol{Y}_i - \boldsymbol{\mu}^{(a_1,a_2)}(\boldsymbol{X}_i; \boldsymbol{\theta})\right)\right)^{\otimes 2}\right]$

Sample Size

We developed a sample size formula for an ENGAGE-style SMART with a continuous longitudinal outcome in which the primary aim is to compare two embedded DTRs which recommend different first-stage treatments on the end-of-study measurement. We ignore baseline covariates and consider three timepoints, $t = 0, 1, 2$.
To compare DTRs $(1,1)$ and $(-1,-1)$, we size the trial based on a Wald test of

$$H_0 : \boldsymbol{c}^\top \boldsymbol{\theta} = 0 \quad \text{vs.} \quad H_1 : \boldsymbol{c}^\top \boldsymbol{\theta} = \Delta,$$

where $\boldsymbol{c}$ is a contrast vector such that $\boldsymbol{c}^\top \boldsymbol{\theta} = \mathrm{E}[Y_2^{(1,1)} - Y_2^{(-1,-1)}]$. The test statistic is

$$Z = \frac{\sqrt{n}\boldsymbol{c}^\top \hat{\boldsymbol{\theta}}}{\sigma_c}.$$

We make three working assumptions to simplify the form of $\sigma_c = \sqrt{\boldsymbol{c}^\top \boldsymbol{B}^{-1}\boldsymbol{M}\boldsymbol{B}^{-1}\boldsymbol{c}}$:

1. The variance in the outcome among non-responders after the second randomization is not too much larger than the corresponding variances in responders,
2. $\mathrm{Cov}(Y_t^{(a_1,a_2)}, Y_2^{(a_1,a_2)} \mid R^{(a_1)} = 1) \le \mathrm{Cov}(Y_t^{(a_1,a_2)}, Y_2^{(a_1,a_2)} \mid R^{(a_1)} = 0)$ for $t = 0, 1$,
3. The marginal variance of $\boldsymbol{Y}^{(a_1,a_2)}$ is constant across time and DTR, and has an exchangeable correlation structure with correlation ρ, i.e., $\mathrm{Var}\left(\boldsymbol{Y}^{(a_1,a_2)}\right) = \sigma^2 \mathbf{Exch}_3(\rho)$.

Suppose we want to detect a **standardized effect size** $\delta = \Delta/\sigma$. Define r_{a_1} to be the probability of response to first-stage treatment a_1.

Sample Size Formula for an ENGAGE-Type SMART

$$n \ge \frac{4\left(z_{1-\alpha/2} + z_{1-\beta}\right)^2}{\delta^2} \cdot \left(1-\rho^2\right) \cdot \left(2 - \frac{1}{2}(r_1 + r_{-1})\right)$$

The sample size formula is the product of three components: **(1)** the formula for the minimum sample size for the comparison of two means in a standard two-arm trial, **(2)** a deflation factor of $1-\rho^2$ that accounts for the use of a longitudinal outcome, and **(3)** a SMART-specific "design effect", an inflation factor that accounts for the SMART design.

Table 1: Example sample sizes for comparing the end-of-study outcomes of two embedded DTRs in an ENGAGE-type SMART which start with different treatments. $r_1 = r_{-1} = 0.4$, $\alpha = 0.05$ (two-sided), and $\beta = 0.2$.

	Within-Person Correlation		
Std. Effect Size	$\rho = 0$	$\rho = 0.3$	$\rho = 0.6$
$\delta = 0.3$	559	508	358
$\delta = 0.5$	201	183	129
$\delta = 0.8$	79	72	51

Acknowledgements

This work was supported by the following awards from the National Institutes of Health: R01DA039901, P50DA039838, R01HD073975, R03MH097954, P01AA016821, RC1AA019092, U54EB020404. The content of this poster is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

Lu, X., I. Nahum-Shani, C. Kasari, K. G. Lynch, D. W. Oslin, W. E. Pelham, G. Fabiano, and D. Almirall (2016). Comparing dynamic treatment regimes using repeated-measures outcomes: Modeling considerations in SMART studies. *Stat Med 35*(10), 1595--1615.

Murphy, S. A. (2005). An experimental design for the development of adaptive treatment strategies. *Stat Med 24*(10), 1455--1481.

Contact Information

- Email: nseewald@umich.edu
- Web: http://nickseewald.com

LSA STATISTICS UNIVERSITY OF MICHIGAN