Budgeting SMART: Sample Size and Repeated Measures with a Cost Constraint in a Longitudinal Sequential, Multiple-Assignment Randomized Trial

Nicholas J. Seewald

Department of Statistics
University of Michigan

Joint with D. Almirall

ENAR Spring Meeting 2021

Motivating Example: The ENGAGE Study

Patients with alcohol- and cocaine-related substance use disorders often disengage from treatment at high rates. How should clinicians best re-engage them?

. McKay, J. R., et al. (2015). *Journal of Consulting and Clinical Psychology*.

Motivating Example: The ENGAGE Study

Patients with alcohol- and cocaine-related substance use disorders often disengage from treatment at high rates. How should clinicians best re-engage them?

For these individuals, should we attempt to re-engage them in their original treatment, or offer them a choice of treatment modality?

. McKay, J. R., et al. (2015). *Journal of Consulting and Clinical Psychology*.

Motivating Example: The ENGAGE Study

Patients with alcohol- and cocaine-related substance use disorders often disengage from treatment at high rates. How should clinicians best re-engage them?

For these individuals, should we attempt to re-engage them in their original treatment, or offer them a choice of treatment modality?

What do we do if that doesn't work?

. McKay, J. R., et al. (2015). *Journal of Consulting and Clinical Psychology*.

Motivating Example: The ENGAGE Study

Patients with alcohol- and cocaine-related substance use disorders often disengage from treatment at high rates. How should clinicians best re-engage them?

For these individuals, should we attempt to re-engage them in their original treatment, or offer them a choice of treatment modality?

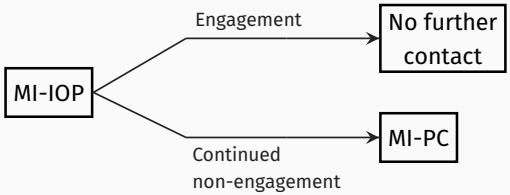
What do we do if that doesn't work?

This is a question about a *sequence* of treatments.

. McKay, J. R., et al. (2015). *Journal of Consulting and Clinical Psychology*.

Dynamic Treatment Regimens

Dynamic treatment regimens (DTRs) operationalize clinical decision-making by recommending particular treatments to certain subsets of patients at specific times.



- **MI-IOP**: 2 motivational interviews to re-engage patient in intensive outpatient program
- **MI-PC**: 2 motivational interviews to engage patient in treatment of their choice.

. Chakraborty, B., and E. E. M. Moodie (2013). *Statistical Methods for Dynamic Treatment Regimes*.

Sequential, Multiple-Assignment Randomized Trials

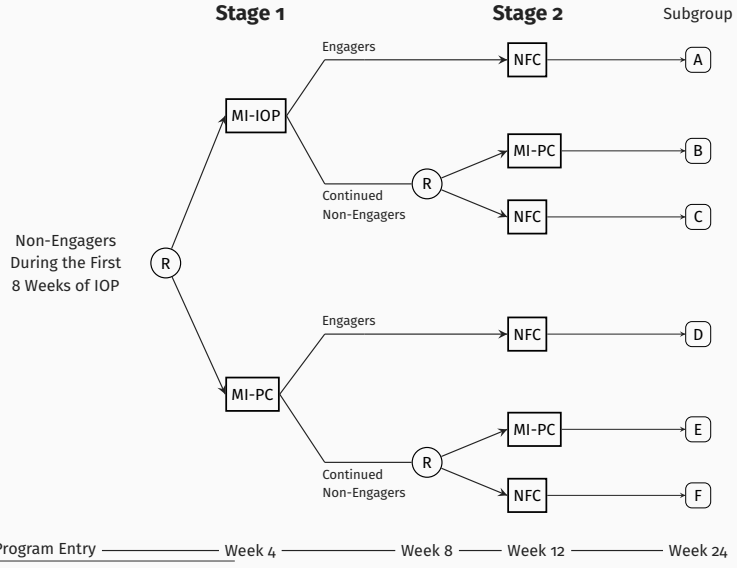
A **SMART** is one type of randomized trial design that can be used to answer questions at multiple stages of the development of a high-quality DTR.

Sequential, Multiple-Assignment Randomized Trials

A **SMART** is one type of randomized trial design that can be used to answer questions at multiple stages of the development of a high-quality DTR.

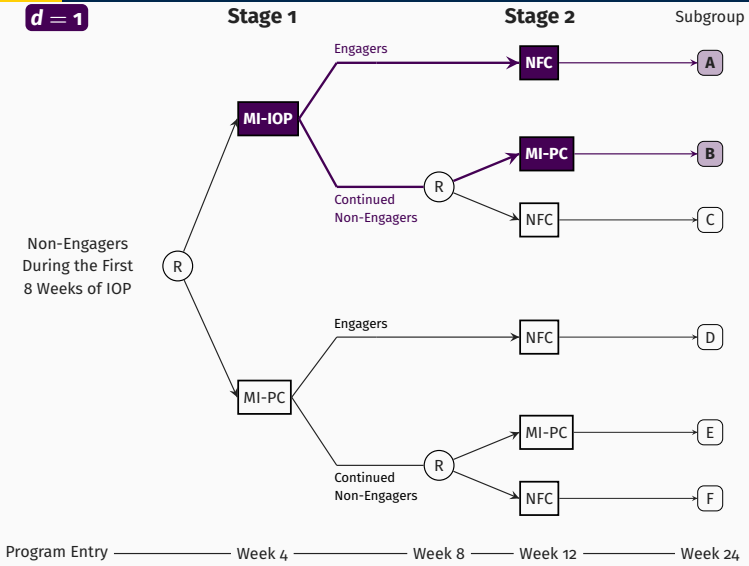
The key feature of a SMART is that some (or all) participants are randomized *more than once*.

Motivating Example: The ENGAGE Study

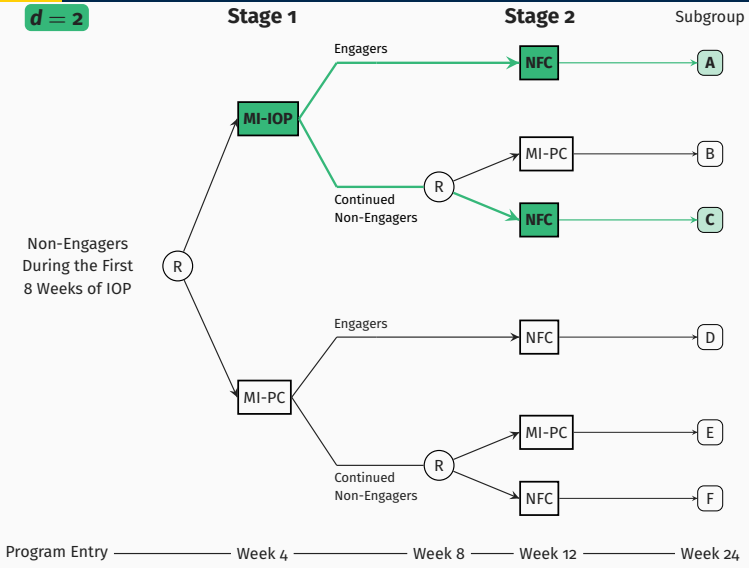


- McKay, J. R., et al. (2015). *Journal of Consulting and Clinical Psychology*.

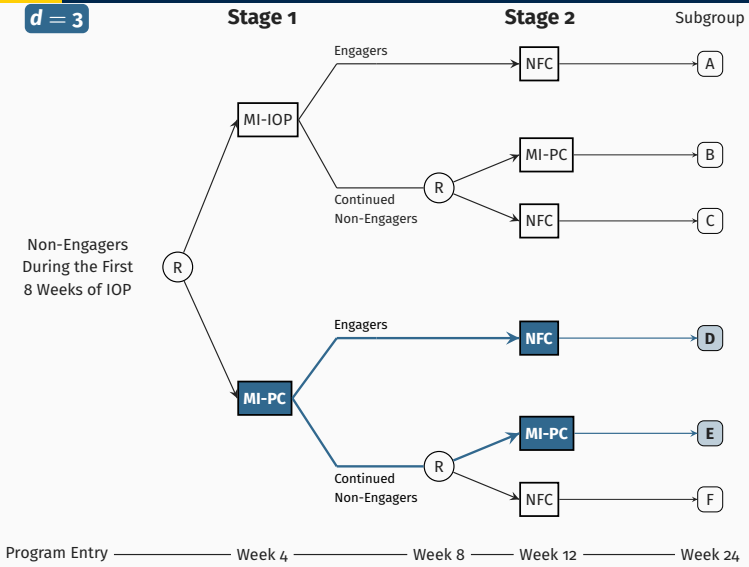
Four Embedded DTRs in ENGAGE



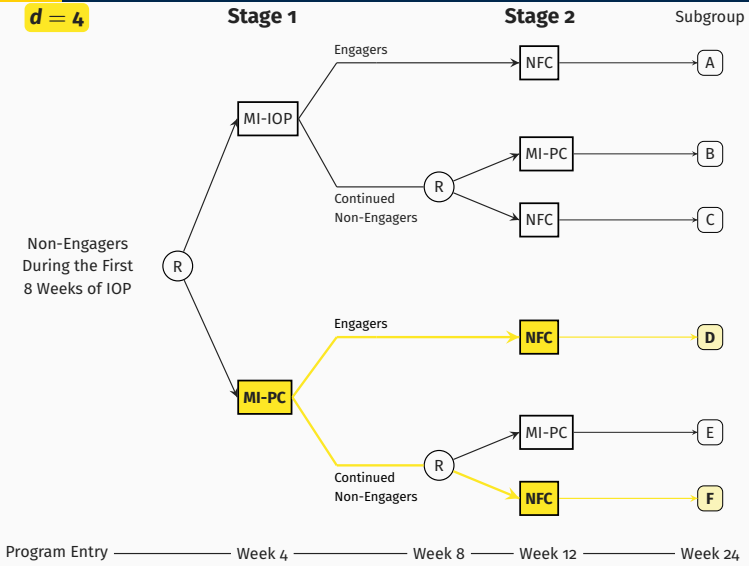
Four Embedded DTRs in ENGAGE



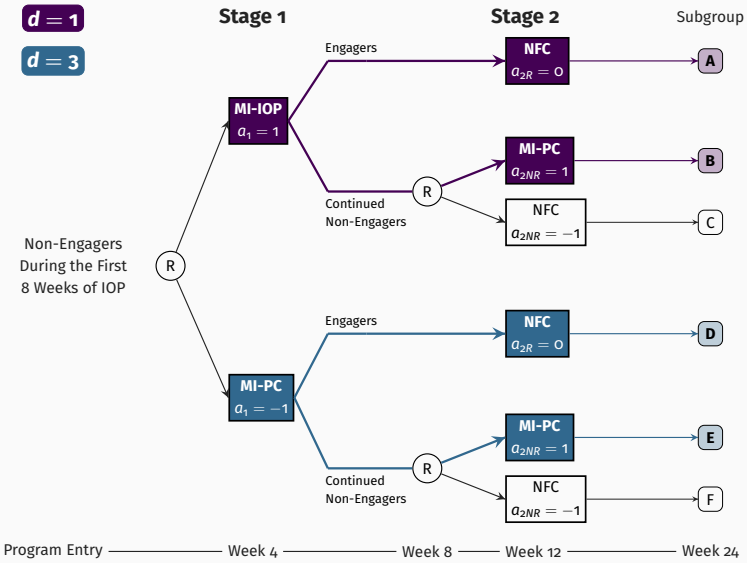
Four Embedded DTRs in ENGAGE



Four Embedded DTRs in ENGAGE

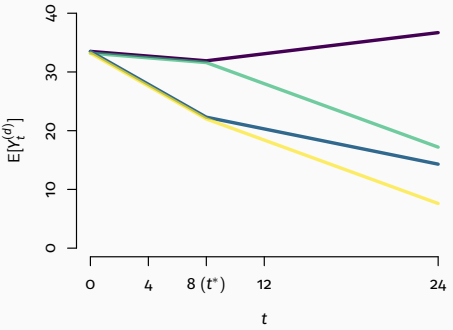


Common Primary Aim: Compare Embedded DTRs at End of Study



Our goal
is to develop a sample size formula for the comparison of two embedded DTRs at the end of the study using a longitudinal outcome collected at an arbitrary number of timepoints.

Example Model: Continuous Longitudinal Outcome in ENGAGE



	$d=1$	$d=2$	$d=3$	$d=4$
a_1	1	1	-1	-1
a_{2R}	0	0	0	0
a_{2NR}	1	-1	1	-1

$$
\begin{aligned}
\mathrm{E}\left[Y_t^{(d)} \mid \boldsymbol{X}\right] &:= \boldsymbol{\mu}^{(d)}(\beta) \\
&= \beta_0 \\
&\quad + \mathbb{1}\{t \le t^*\}\{\beta_1 t + \beta_2 a_1 t\} \\
&\quad + \mathbb{1}\{t > t^*\}\{t^*\beta_1 + t^*\beta_2 a_1 \\
&\qquad + \beta_3(t - t^*) + \beta_4(t - t^*)a_1 \\
&\qquad + \beta_5(t - t^*)a_{2NR} \\
&\qquad + \beta_6(t - t^*)a_1 a_{2NR}\}
\end{aligned}
$$

“GEE-Type” Estimating Equations for Model Parameters

$$0 = \sum_{i=1}^{N} \sum_{d} \left[\overbrace{\frac{I^{(d)}(A_{1,i}, R_i, A_{2,i})}{P(A_{1,i} = a_1) P(A_{2,i} = a_2 \mid A_{1,i} = a_1, R_i)}}^{W^{(d)}(A_{1,i}, R_i, A_{2,i})} \cdot \left(\boldsymbol{D}^{(d)} \right)^{\top} \cdot \boldsymbol{V}^{(d)}(\tau)^{-1} \cdot \left(\boldsymbol{Y}_i - \boldsymbol{\mu}^{(d)}(\beta) \right) \right],$$

- d specifies an embedded DTR,
- $I^{(d)}(A_{1,i}, R_i, A_{2,i}) = \mathbb{1}\{A_{1,i} = a_1\} \left(R_i + (1 - R_i)\, \mathbb{1}\{A_{2,i} = a_2\} \right)$
- $\boldsymbol{D}^{(d)} = \frac{\partial}{\partial \boldsymbol{\beta}^{\top}} \boldsymbol{\mu}^{(d)}(\beta)$
- $\boldsymbol{V}^{(d)}(\tau)$ is a working model for $\mathbf{Var}\left(\boldsymbol{Y}^{(d)} - \boldsymbol{\mu}^{(d)}(\beta) \right)$

. Lu, X., et al. (2016). *Statistics in Medicine*.

Goal

Goal: Develop a tractable sample size formula for the test

$$H_0 : \mathrm{E}\left[Y_T^{(d=1)} - Y_T^{(d=3)}\right] = 0 \quad \text{vs.} \quad H_1 : \mathrm{E}\left[Y_T^{(d=1)} - Y_T^{(d=3)}\right] = \Delta.$$

Goal

Goal: Develop a tractable sample size formula for the test

$$H_0 : \mathrm{E}\left[Y_T^{(d=1)} - Y_T^{(d=3)}\right] = 0 \quad \text{vs.} \quad H_1 : \mathrm{E}\left[Y_T^{(d=1)} - Y_T^{(d=3)}\right] = \Delta.$$

Under our example model,

$$\mathrm{E}\left[Y_T^{(d=1)} - Y_T^{(d=3)}\right] = \boldsymbol{c}^\top \boldsymbol{\beta}$$

Goal

Goal: Develop a tractable sample size formula for the test

$$H_0 : \mathrm{E}\left[Y_T^{(d=1)} - Y_T^{(d=3)}\right] = 0 \quad \text{vs.} \quad H_1 : \mathrm{E}\left[Y_T^{(d=1)} - Y_T^{(d=3)}\right] = \Delta.$$

Under our example model,

$$\mathrm{E}\left[Y_T^{(d=1)} - Y_T^{(d=3)}\right] = \boldsymbol{c}^\top \beta$$

We use a 1-degree of freedom (asymptotic) Wald test with test statistic

$$Z = \frac{\sqrt{n}\boldsymbol{c}^\top \hat{\beta}}{\sigma_c},$$

where $\sigma_c = \boldsymbol{c}^\top \mathbf{Var}\left(\hat{\beta}\right)\boldsymbol{c}$.

Sample Size for an End-of-Study Comparison

Under mild working assumptions, exchangeable within-person correlation, and constant variance across time and DTRs:

$$N \geq \frac{4\left(z_{1-\alpha/2} + z_{1-\gamma}\right)^2}{\delta^2} \cdot \left(2 - P(R_i = 1)\right) \cdot \omega(\rho, \mathbf{t}, T_2)$$

- $\delta = \Delta/\sigma = \mathrm{E}[Y_T^{(d)} - Y_T^{(d')}] / \sqrt{\left(\mathrm{Var}(Y_T^{(d)}) + \mathrm{Var}(Y_T^{(d')})\right)/2}$ is the target standardized effect size
- α is the desired type-I error
- $1 - \gamma$ is the desired power
- $\rho = \mathrm{cor}(Y_t, Y_{t'})$ for $t \neq t'$
- $\mathbf{t}$ is a vector of measurement times
- T_2 is the number of measurements in stage 2

Sample Size for an End-of-Study Comparison

Under mild working assumptions, exchangeable within-person correlation, and constant variance across time and DTRs:

$$N \geq \underbrace{\frac{4\left(z_{1-\alpha/2} + z_{1-\gamma}\right)^2}{\delta^2}}_{\text{Standard sample size for a 2-arm trial}} \cdot \left(2 - P(R_i = 1)\right) \cdot \omega(\rho, \mathbf{t}, T_2)$$

- $\delta = \Delta/\sigma = \mathrm{E}[Y_T^{(d)} - Y_T^{(d')}] / \sqrt{\left(\mathrm{Var}(Y_T^{(d)}) + \mathrm{Var}(Y_T^{(d')})\right)/2}$ is the target standardized effect size
- α is the desired type-I error
- $1 - \gamma$ is the desired power
- $\rho = \mathrm{cor}(Y_t, Y_{t'})$ for $t \neq t'$
- $\mathbf{t}$ is a vector of measurement times
- T_2 is the number of measurements in stage 2

Sample Size for an End-of-Study Comparison

Under mild working assumptions, exchangeable within-person correlation, and constant variance across time and DTRs:

$$N \geq \frac{4\left(z_{1-\alpha/2} + z_{1-\gamma}\right)^2}{\delta^2} \cdot \underbrace{\left(\mathbf{2 - P(R_i = 1)}\right)}_{\text{Inflation: SMART design}} \cdot \omega(\rho, \mathbf{t}, T_2)$$

- $\delta = \Delta/\sigma = \mathrm{E}[Y_T^{(d)} - Y_T^{(d')}] / \sqrt{\left(\mathrm{Var}(Y_T^{(d)}) + \mathrm{Var}(Y_T^{(d')})\right)/2}$ is the target standardized effect size
- α is the desired type-I error
- $1 - \gamma$ is the desired power
- $\rho = \mathrm{cor}(Y_t, Y_{t'})$ for $t \neq t'$
- $\mathbf{t}$ is a vector of measurement times
- T_2 is the number of measurements in stage 2

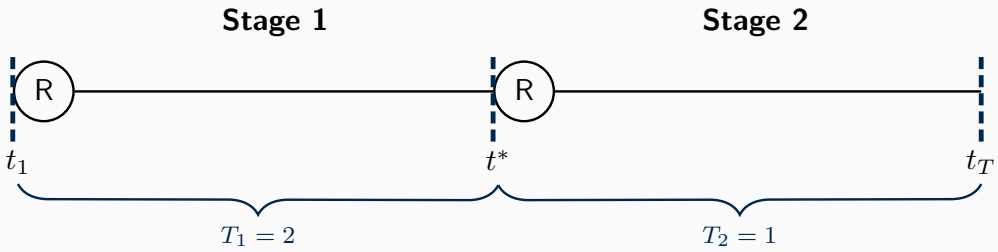
Sample Size for an End-of-Study Comparison

Under mild working assumptions, exchangeable within-person correlation, and constant variance across time and DTRs:

$$N \geq \frac{4\left(z_{1-\alpha/2} + z_{1-\gamma}\right)^2}{\delta^2} \cdot \left(2 - P(R_i = 1)\right) \cdot \underbrace{\omega(\rho, \mathbf{t}, T_2)}_{\text{Deflation: within-person outcome}}$$

- $\delta = \Delta/\sigma = \mathrm{E}[Y_T^{(d)} - Y_T^{(d')}] / \sqrt{\left(\mathrm{Var}(Y_T^{(d)}) + \mathrm{Var}(Y_T^{(d')})\right)/2}$ is the target standardized effect size
- α is the desired type-I error
- $1 - \gamma$ is the desired power
- $\rho = \mathrm{cor}(Y_t, Y_{t'})$ for $t \neq t'$
- $\mathbf{t}$ is a vector of measurement times
- T_2 is the number of measurements in stage 2

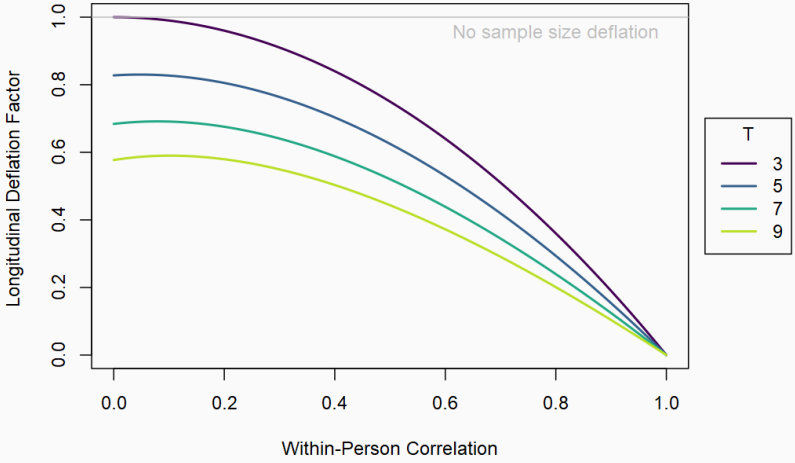
Special Case: 3 timepoints simplifies nicely



$$\omega(\rho, 1, 3) = (1 - \rho^2)$$

. Seewald, N. J., et al. (2020). *Statistical Methods in Medical Research*.

Understanding $\omega(\rho, \mathbf{t}, T_2)$: Increase T, fix $T_2 = \lfloor T/2 \rfloor$



Increasing T decreases sample size requirements (with diminishing returns).

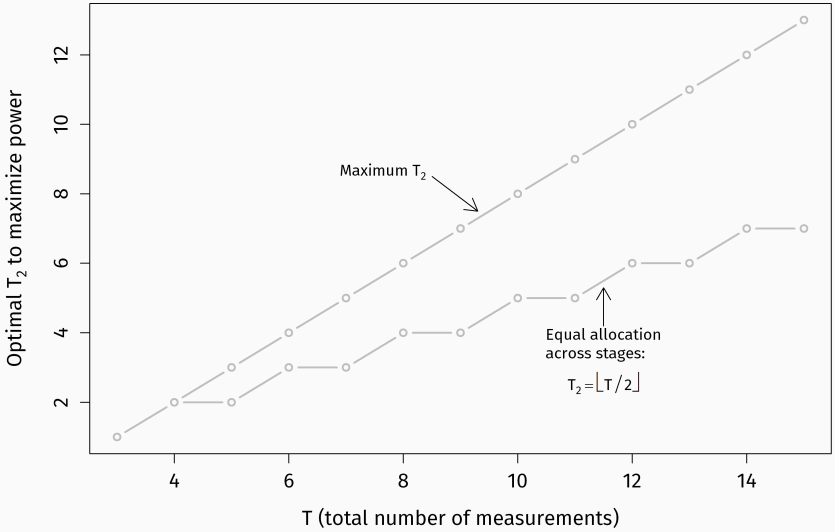
Big Question:
Given a fixed N, T, and ρ, how do we allocate measurements across stages of the SMART in order to maximize power?

Big Question:
Given a fixed N, T, and ρ, how do we allocate measurements across stages of the SMART in order to maximize power?

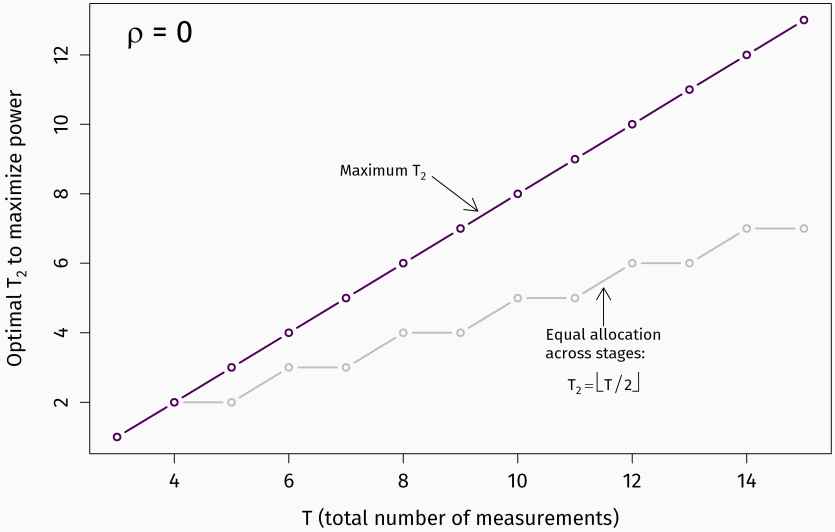
For simplicity, consider equally-spaced measurements throughout the trial.

Minimum of 2 measurements in stage 1 (baseline, before re-randomization)

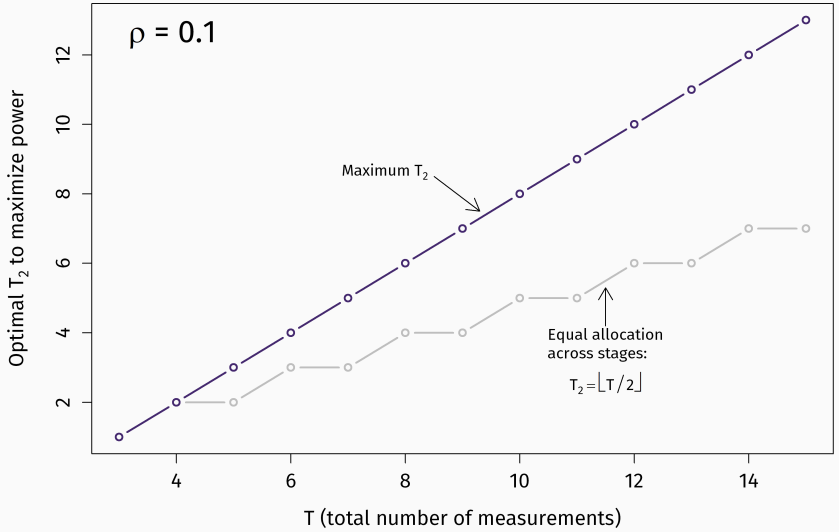
Choosing T_2 to maximize power



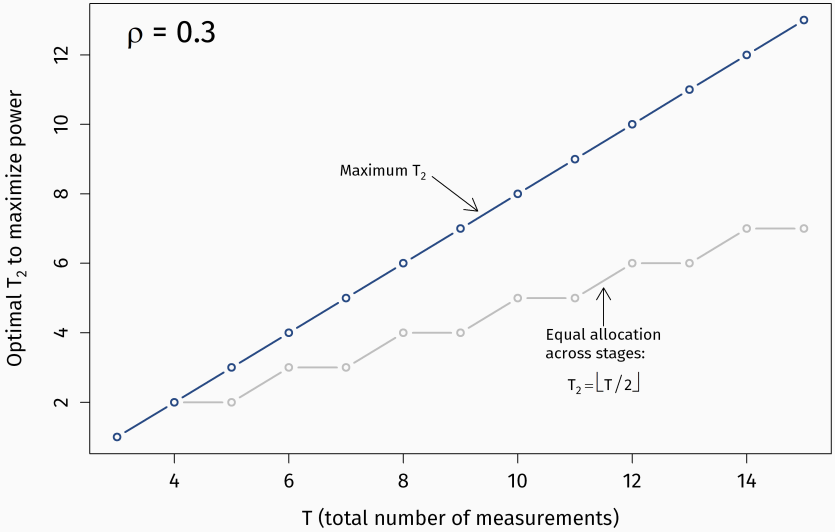
Choosing T_2 to maximize power



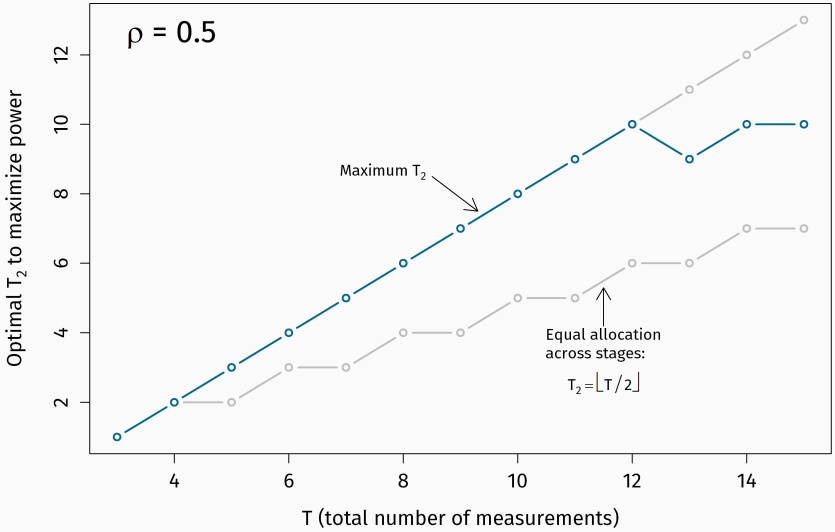
Choosing T_2 to maximize power



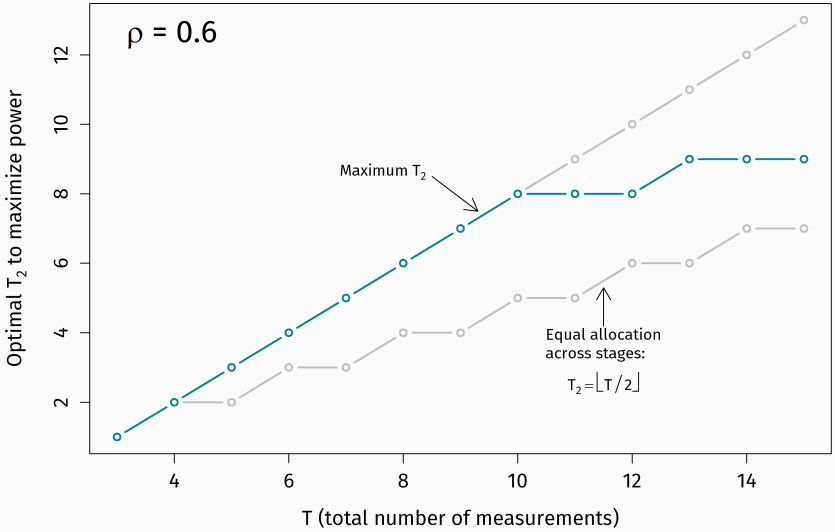
Choosing T_2 to maximize power



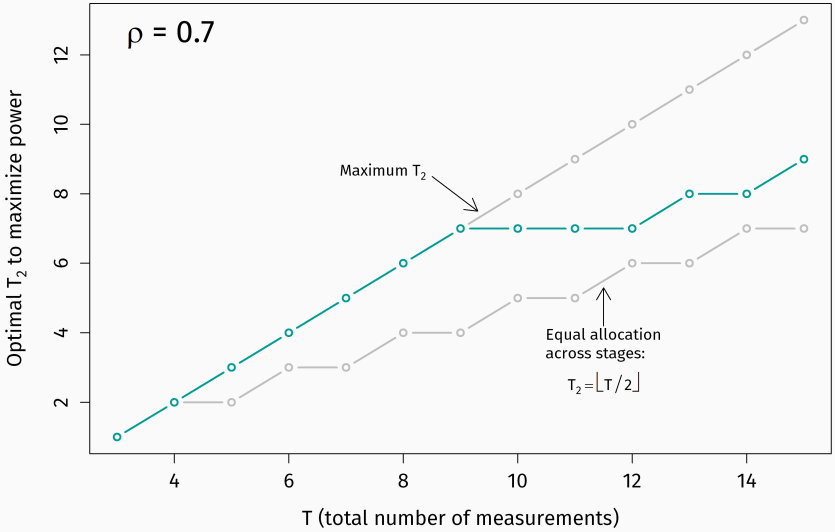
Choosing T_2 to maximize power



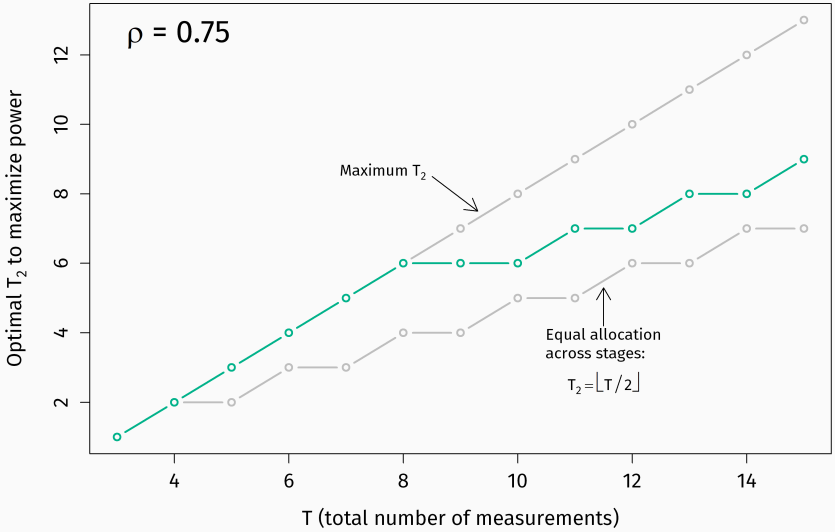
Choosing T_2 to maximize power



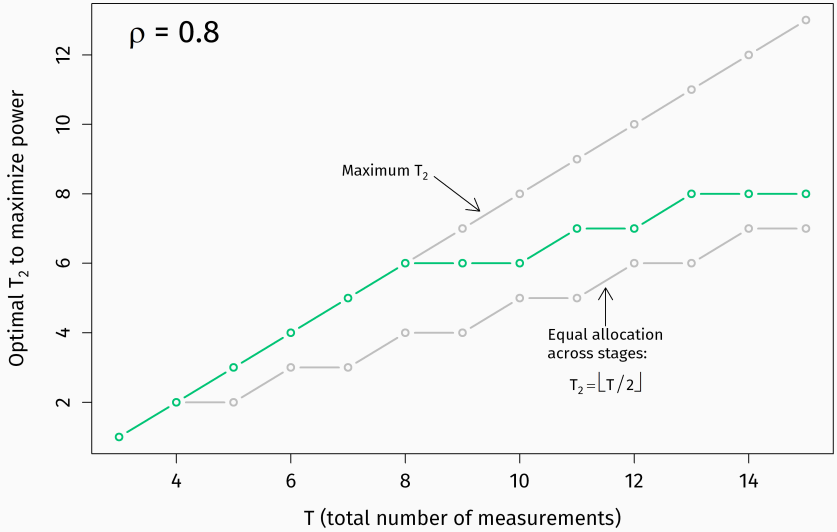
Choosing T_2 to maximize power



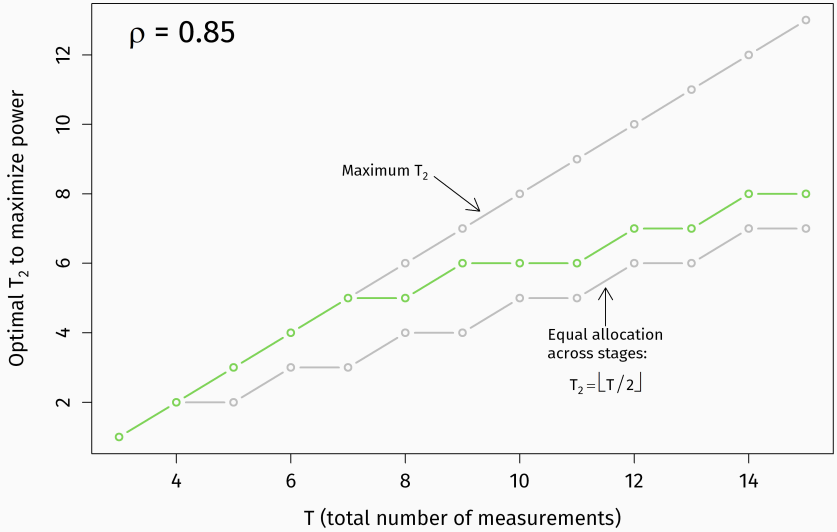
Choosing T_2 to maximize power



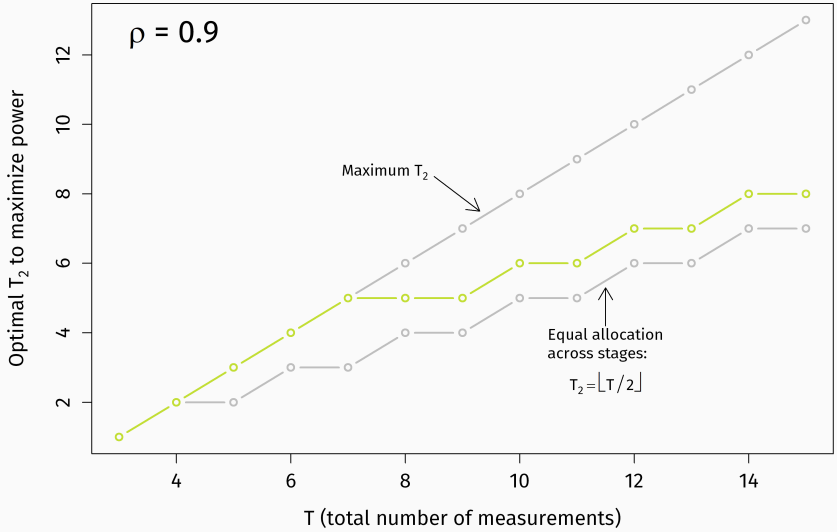
Choosing T_2 to maximize power



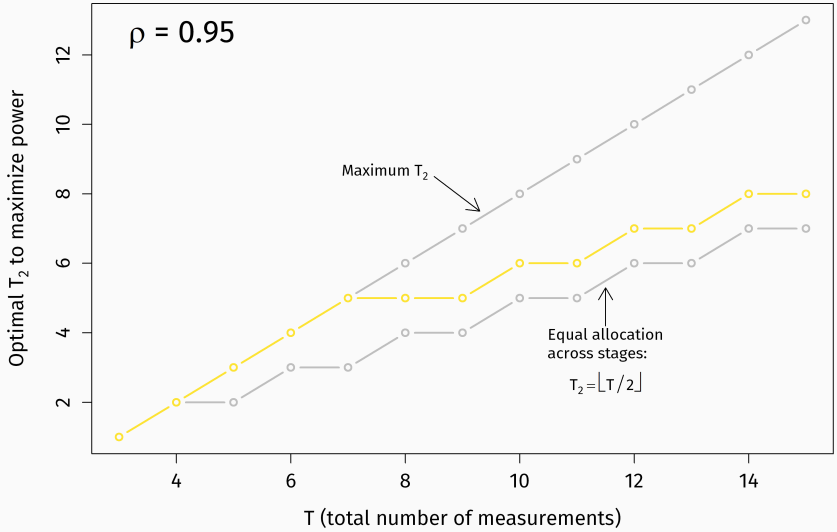
Choosing T_2 to maximize power



Choosing T_2 to maximize power



Choosing T_2 to maximize power



Choosing $T_2 = T - 2$ is sub-optimal for large ρ with large T

With equally-spaced measurements,

- For low ρ and/or low T, put as many measurements in stage 2 as possible.
  - At low ρ, power gains are likely from better modeling the linear trend in stage 2

Choosing $T_2 = T - 2$ is sub-optimal for large ρ with large T

With equally-spaced measurements,

- For low ρ and/or low T, put as many measurements in stage 2 as possible.
  - At low ρ, power gains are likely from better modeling the linear trend in stage 2
- For higher ρ and/or higher T, diminishing returns of more measurements in stage 2
  - At high ρ, more information per measurement; share the love with stage 1

Choosing $T_2 = T - 2$ is sub-optimal for large ρ with large T

With equally-spaced measurements,

- For low ρ and/or low T, put as many measurements in stage 2 as possible.
  - At low ρ, power gains are likely from better modeling the linear trend in stage 2
- For higher ρ and/or higher T, diminishing returns of more measurements in stage 2
  - At high ρ, more information per measurement; share the love with stage 1
- Difficult to identify exactly what "low ρ" and "high T" mean, since $\omega(\rho, \mathbf{t}, T_2)$ is complicated.

A brief discussion of budget constraints

A work in progress! Inspired by Zhang and Ahn (2011)

Setup

- Total budget B
- Cost C_R of recruiting one participant
- Cost C_M of measuring outcome per participant
- Assume equally-spaced measurements across stages

. Zhang, S., and C. Ahn (2011). *Statistics in Biopharmaceutical Research*.

A brief discussion of budget constraints

A work in progress! Inspired by Zhang and Ahn (2011)

Setup

- Total budget B
- Cost C_R of recruiting one participant
- Cost C_M of measuring outcome per participant
- Assume equally-spaced measurements across stages

Budget Constraint
Choose N, T, T_2 to maximize power such that

$$NC_R + NTC_M \leq B.$$

. Zhang, S., and C. Ahn (2011). *Statistics in Biopharmaceutical Research*.

Extremely preliminary numerical results

	Exch. Correlation			
C_R/C_M	$\rho = 0.2$	$\rho = 0.4$	$\rho = 0.6$	$\rho = 0.8$
1	3	3	3	3
10	13	9	7	6
25	15	15	12	10
50	15	15	15	15

Notes:

- Set $T_2 = T - 2$
- Considering $T \leq 15$ for all scenarios
- For chosen T, use maximum-affordable N

Summary and Recap

- Interpretable sample size formula for end-of-study comparisons of embedded DTRs using a continuous longitudinal outcome
  - Depends on ρ and measurement times
- Optimal allocation of measurements favors stage 2
- Budget constraint seems to have little middle ground

Coming Soon

- A work in progress!
- Still to Come:
  - User-friendly sample size tool: $\omega(\rho, \mathbf{t}, T_2)$ is complicated
  - Software for helping clinicians optimize N, T, T_2 within a budget

Sample Size Considerations for 3 Measurements

Article

SMMR
STATISTICAL METHODS IN MEDICAL RESEARCH

Statistical Methods in Medical Research
2020, Vol. 29(7) 1891–1912

DOI: 10.1177/0962280219877520
journals.sagepub.com/home/smm
SAGE

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] and Daniel Almirall[1,3]

arXiv:1810.13094 [stat.ME]

Funding

This work was supported by the National Institutes of Health (R01 DA039901, R01 HD095973, R01 DA047279) and the Institute of Education Sciences (R32 4B180003).The content of this presentation is solely the responsibility of the authors and does not necessarily represent the official views of funding agencies.

https:\\nickseewald.com