Design, Analysis, and Sizing of Sequential Multiple Assignment Randomized Trials with Binary Outcomes

Graduate Student Statistical Topics Seminar Series

Nicholas J. Seewald[1], Daniel Almirall[2], Kelley M. Kidwell[1]

[1]Department of Biostatistics, University of Michigan
[2]Survey Research Center, University of Michigan

September 24, 2015

Overview

1. Dynamic Treatment Regimes
2. Sequential Multiple Assignment Randomized Trials
3. Analysis of Binary SMART Data
4. Sample Size for Binary-Outcome SMARTs

"Ignorance of whether or how to change psychotherapies is a major and persisting gap in psychiatric knowledge."

John Markowitz, Barbara Milrod (2015).
The Lancet Psychiatry, 2(2), 186-190.

Motivation

- Suppose you visit the doctor...

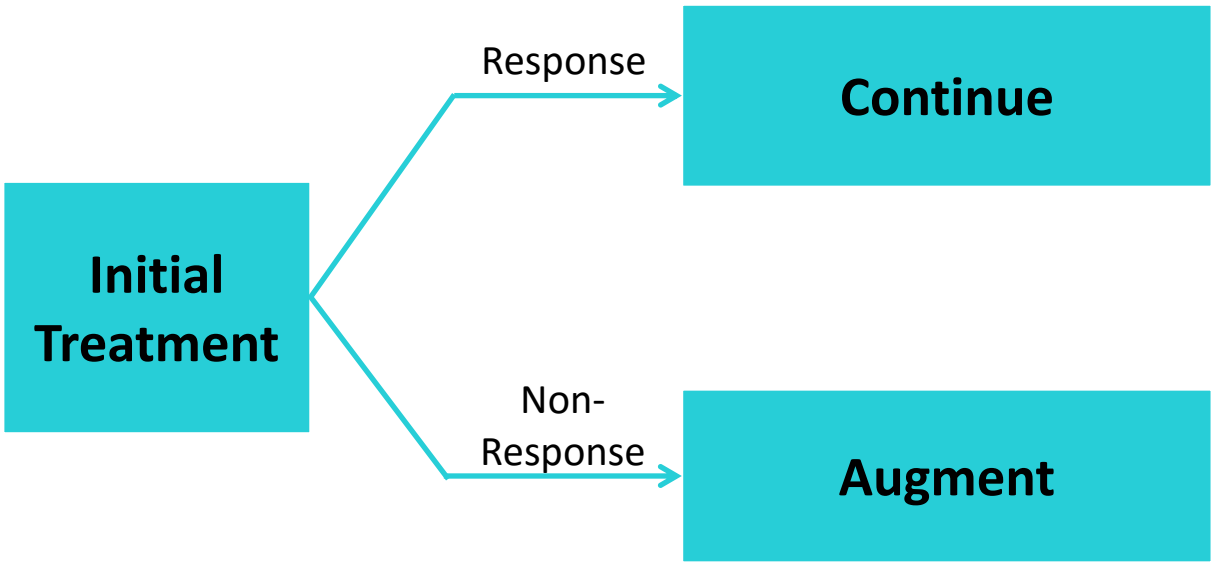
Motivation

- Suppose you visit the doctor...
- Your doctor *adapts* the treatment she provides to you based on your needs, which may change over time.

Motivation

- Suppose you visit the doctor...
- Your doctor *adapts* the treatment she provides to you based on your needs, which may change over time.
- Behind the scenes, she is choosing a treatment according to a *decision rule* that she has in her head.

The Doctor's Decision Rule



Can we mathematize this process?

Decision Rules

Suppose we want to make a decision about treatment at each of k timepoints.

- Denote the decision at time j by a_j.
- S_j, $j \leq k$, represents an *intermediate outcome*: information available after decision a_{j-1} and before decision a_j.

Decision Rules

Suppose we want to make a decision about treatment at each of k timepoints.

- Denote the decision at time j by a_j.
- S_j, $j \leq k$, represents an *intermediate outcome*: information available after decision a_{j-1} and before decision a_j.

Definition

A **decision rule** d_j is a function that takes intermediate outcomes prior to time j, $\bar{S}_j = \{S_1, \ldots, S_j\}$ and previous decisions $\{a_1, \ldots, a_{j-1}\}$ and outputs a treatment decision a_j.

Dynamic Treatment Regimes

The doctor makes a *sequence* of decisions according to a *sequence* of decision rules. We call this a **Dynamic Treatment Regime**.

Definition

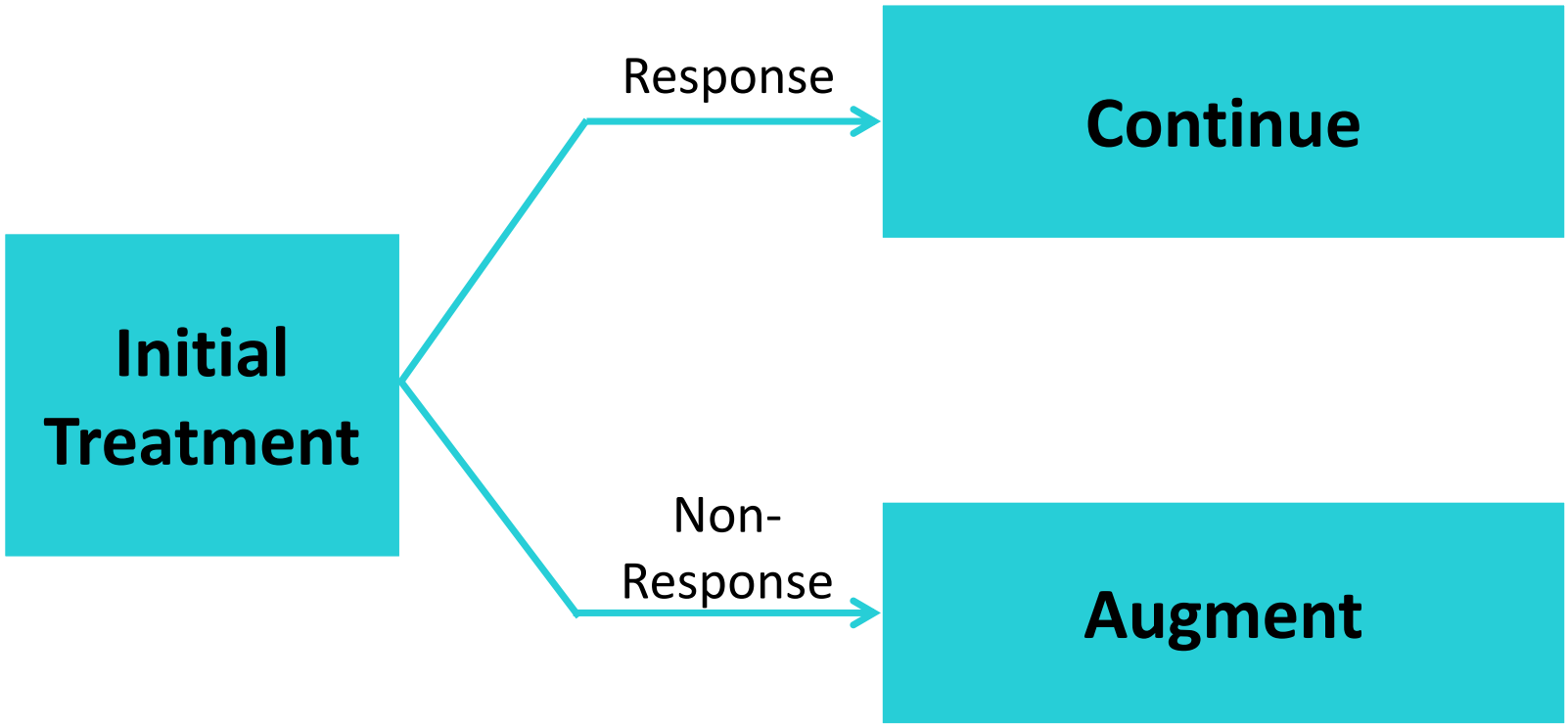
A **Dynamic Treatment Regime** (**DTR**) is a sequence of decision rules $\{d_1, \ldots, d_k\}$.

DTRs in Practice

A dynamic treatment regime (DTR) is...

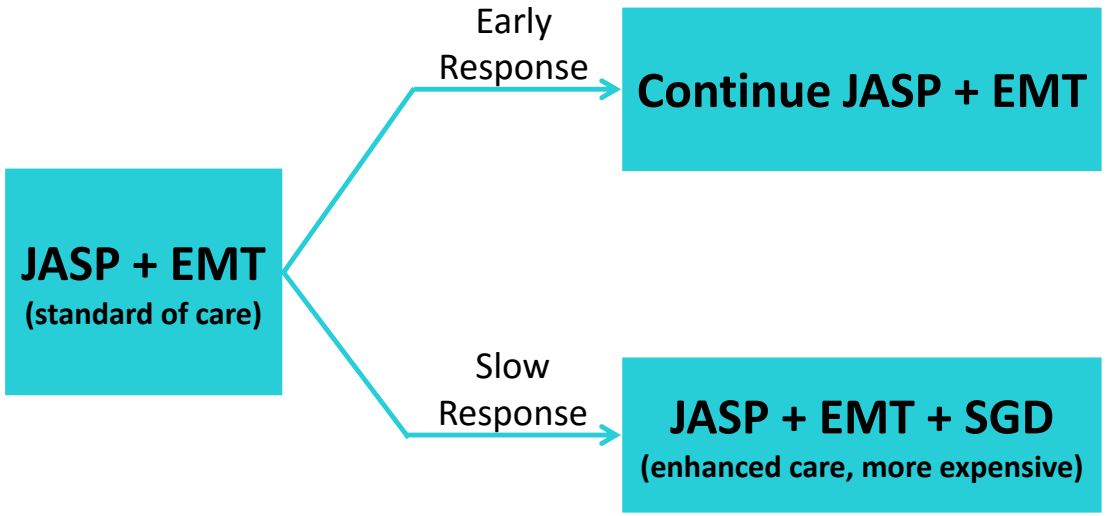
- an intervention guideline
- in which treatments are individualized
- according to the specific and changing needs of patients.

Visualizing the Doctor's DTR



A Real-World Example

Kasari, et al., 2014 [2]



"Start by giving standard care, JASP + EMT. If the child responds early, continue. If the child responds slowly, add a speech generating device (SGD) to JASP + EMT."

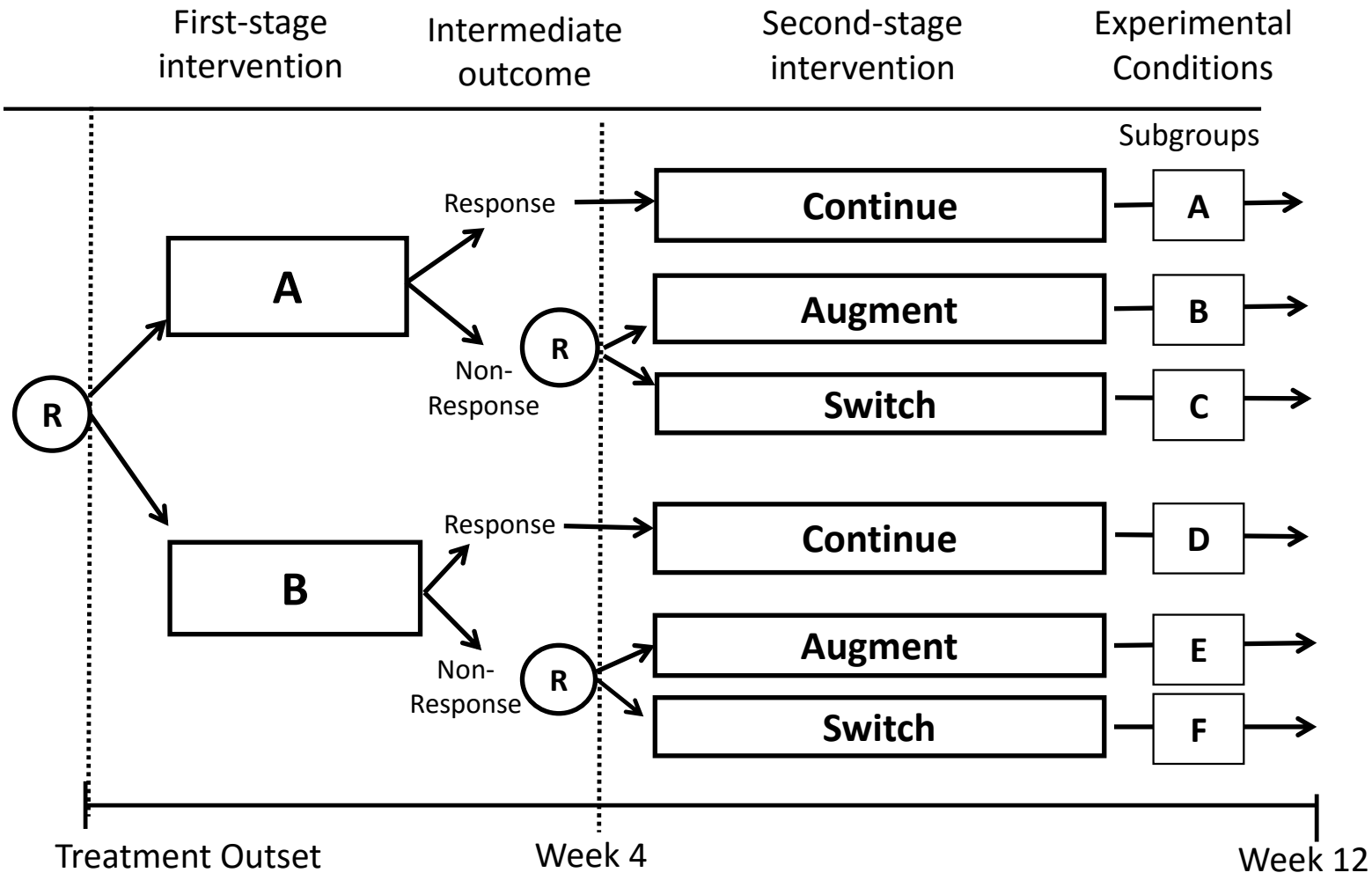
A (hypothetical) Motivating Example

Suppose I want to develop a high-quality DTR to treat Netflix addiction.
However, there's not enough evidence to determine...

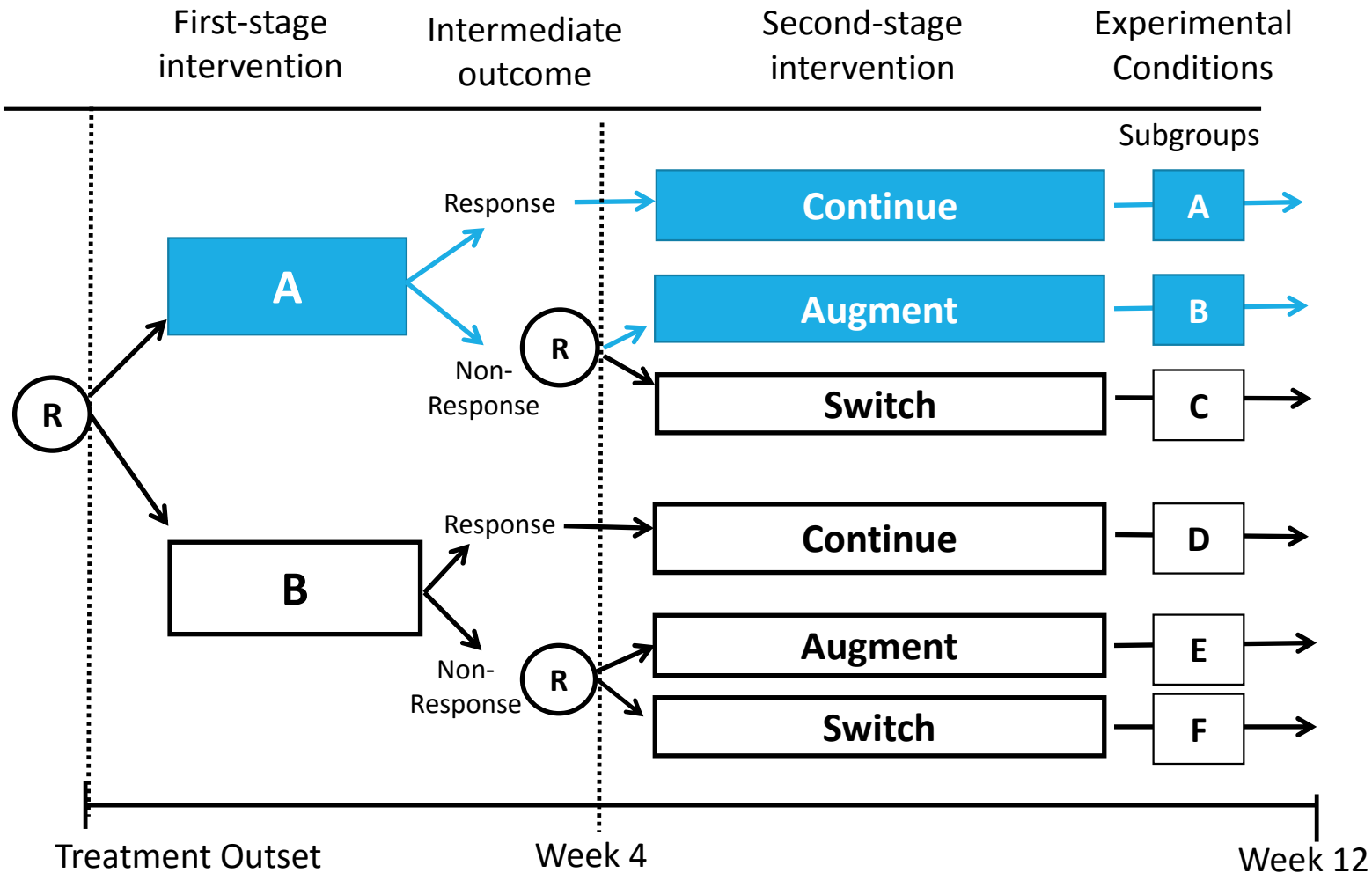
- ...how to **initiate treatment** (A or B?)
- ...how to **modify treatment** for early non-responders (switch or augment?)

For responders to A or B, continue with initial treatment.

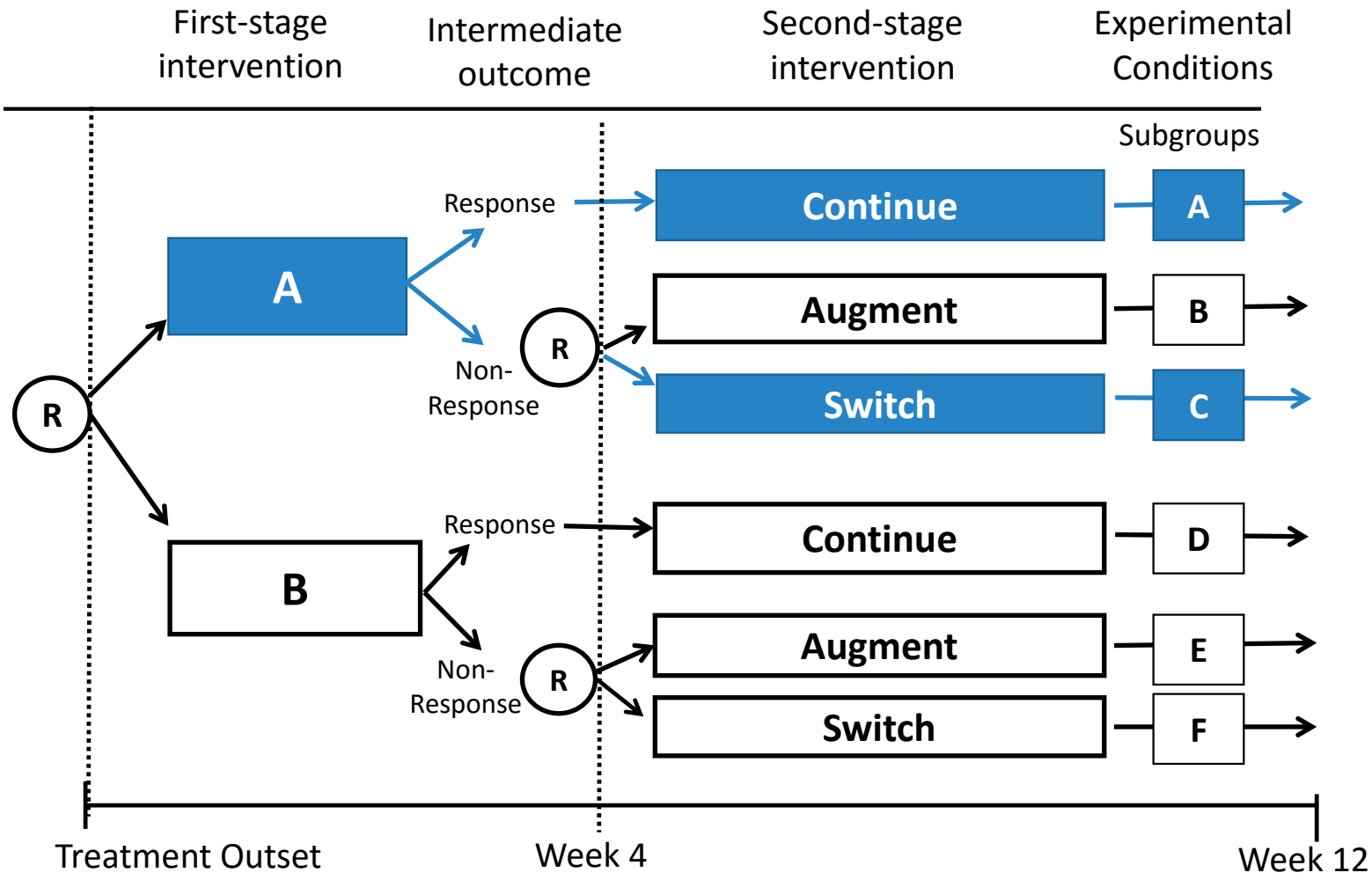
A (hypothetical) Motivating Example



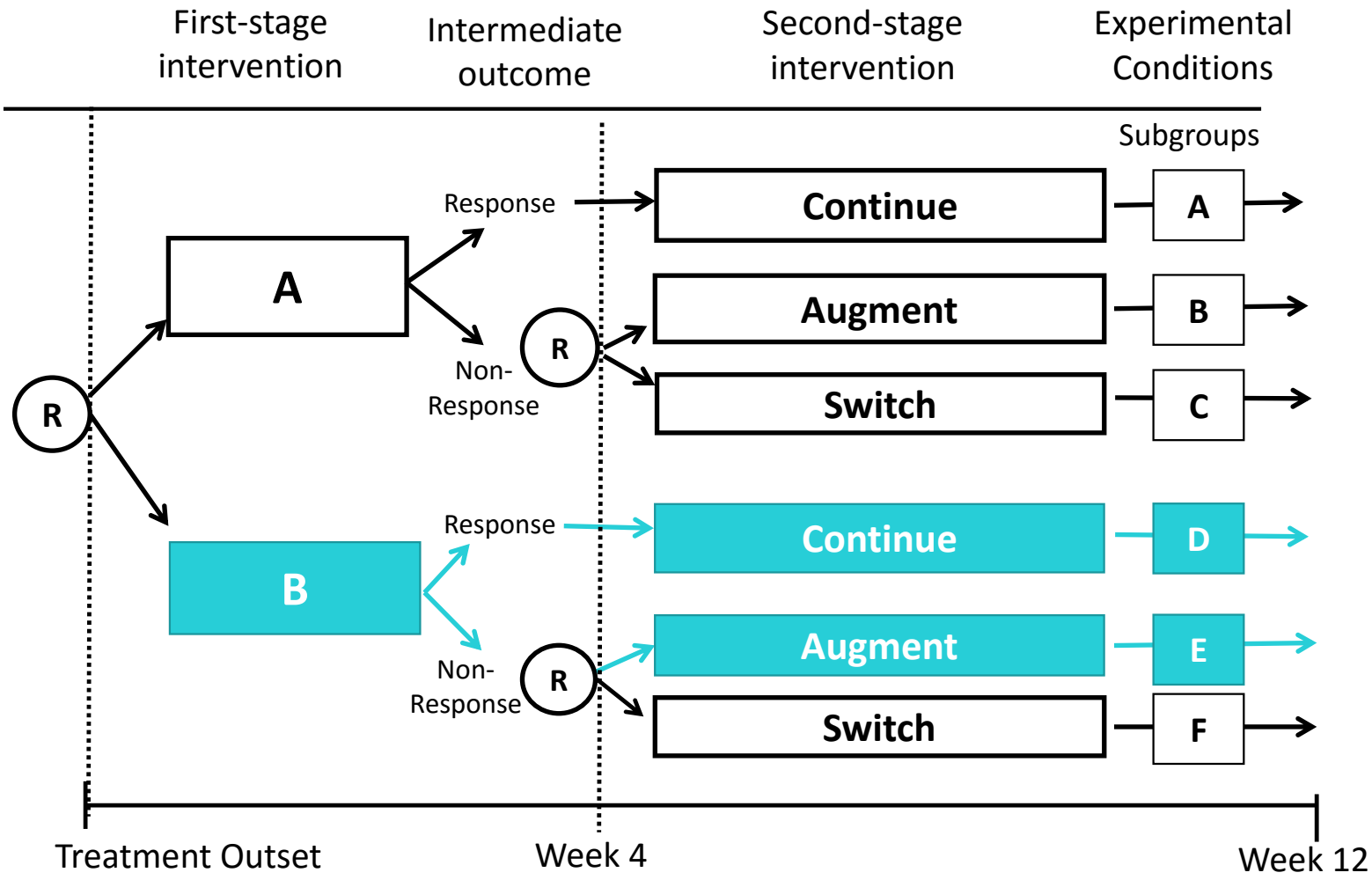
DTRs are Embedded in SMARTs



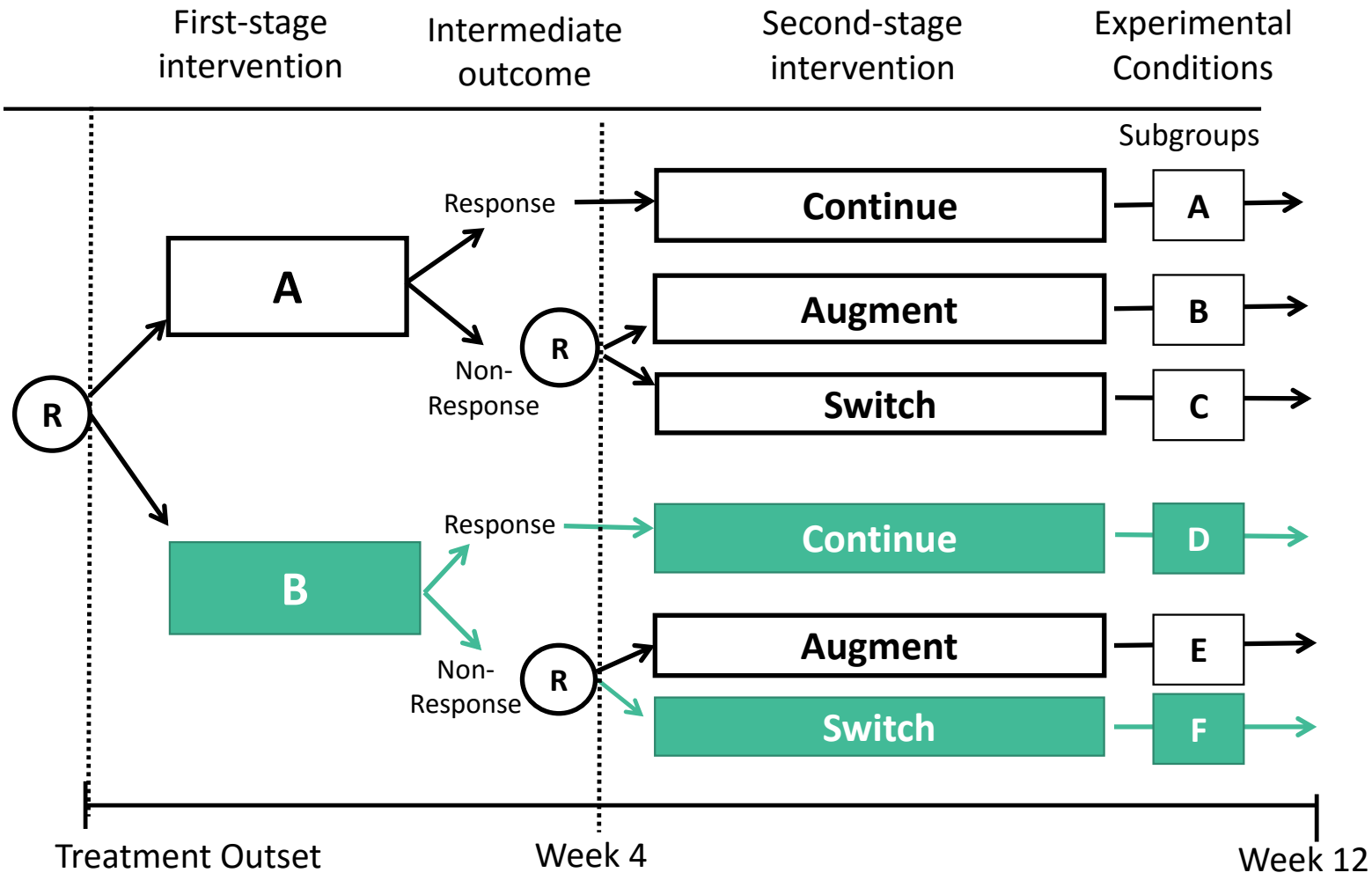
DTRs are Embedded in SMARTs



DTRs are Embedded in SMARTs



DTRs are Embedded in SMARTs



Other Common SMART Designs

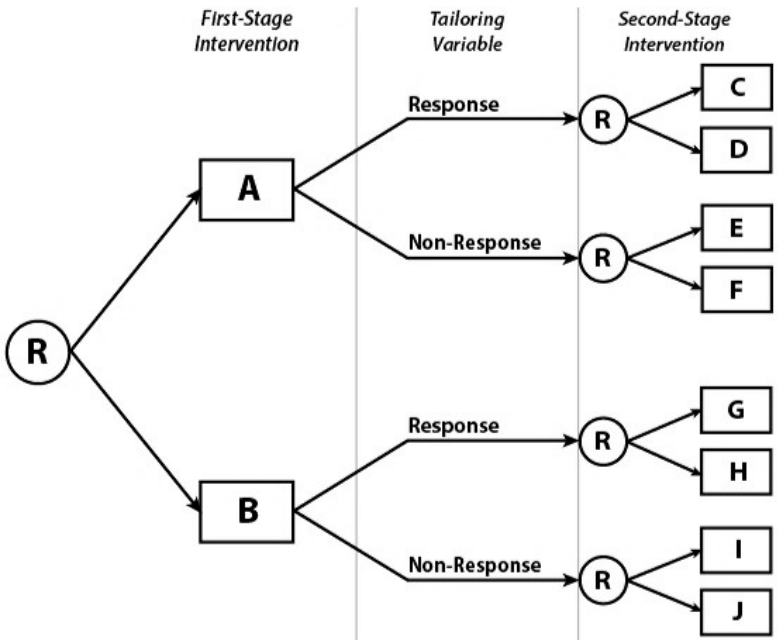


Figure: Re-randomize all participants. 8 embedded DTRs.

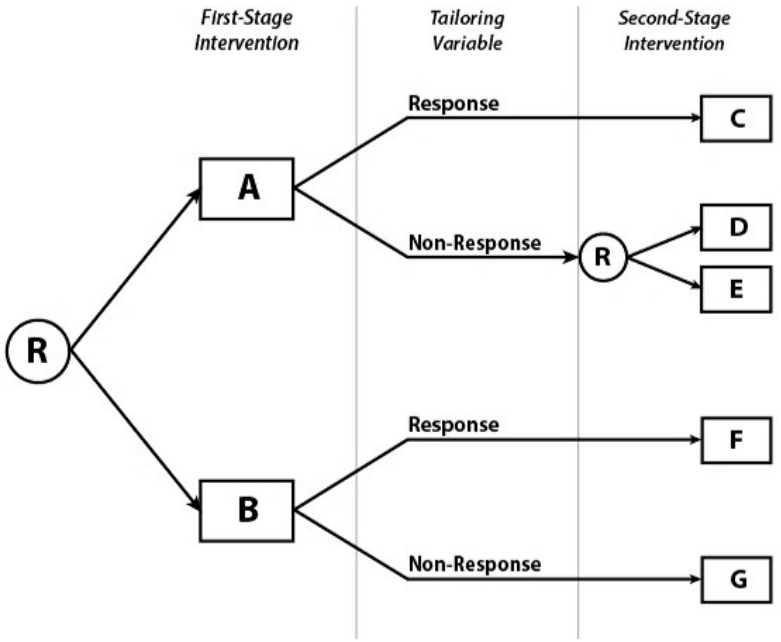
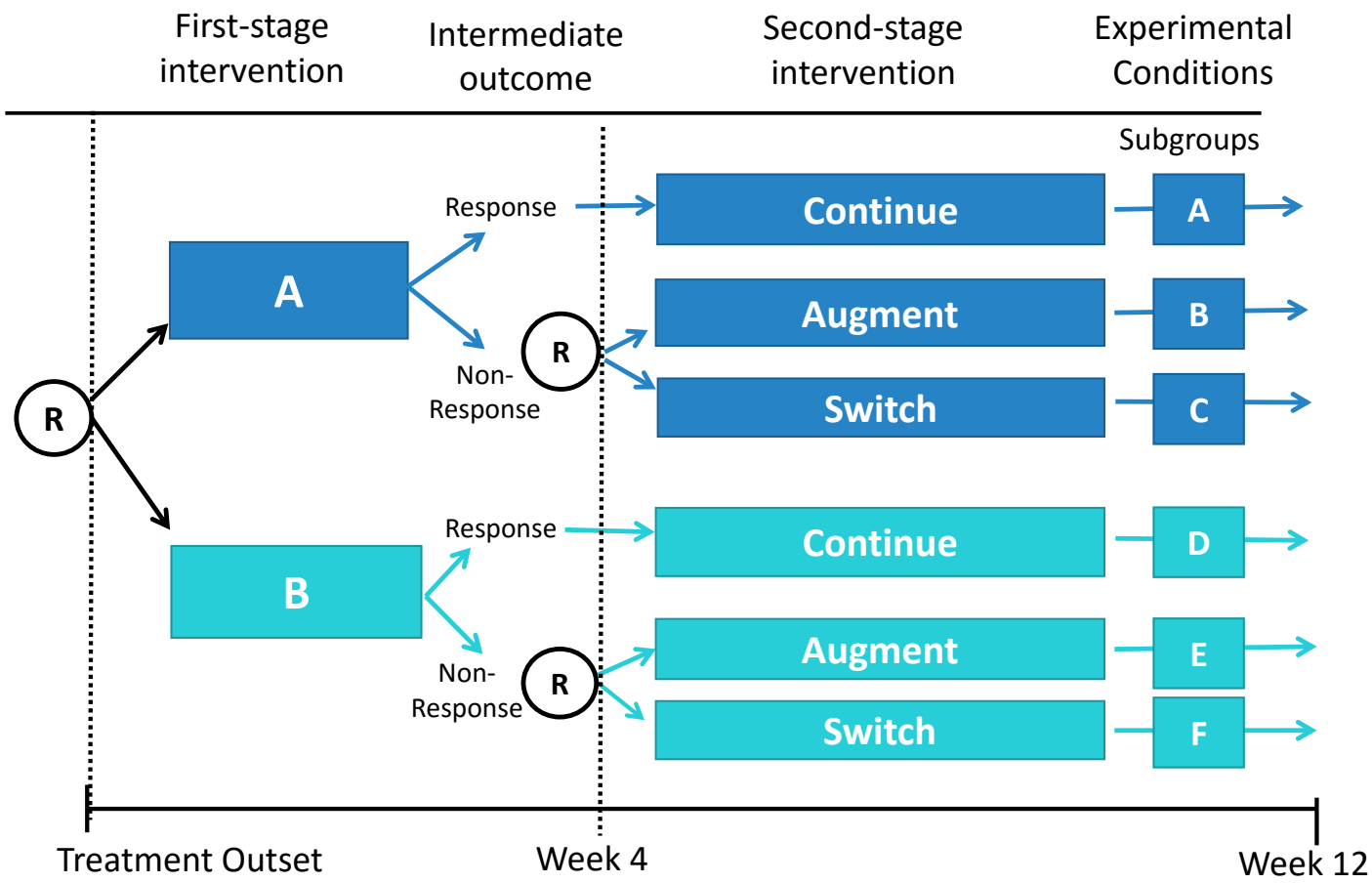


Figure: Re-randomize only non-responders to treatment A. 3 embedded DTRs.

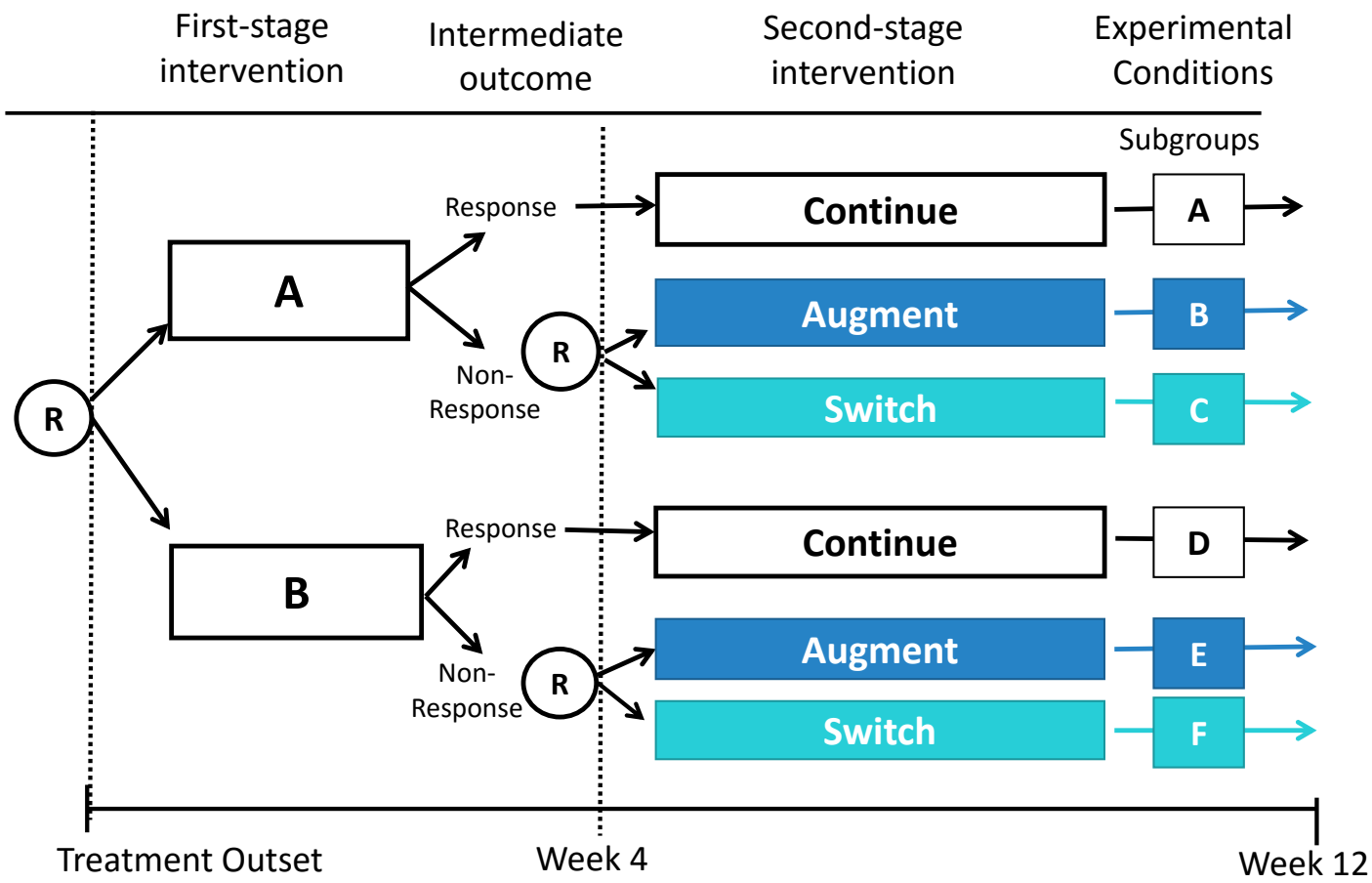
Common Primary Aims

Compare Initial Treatments



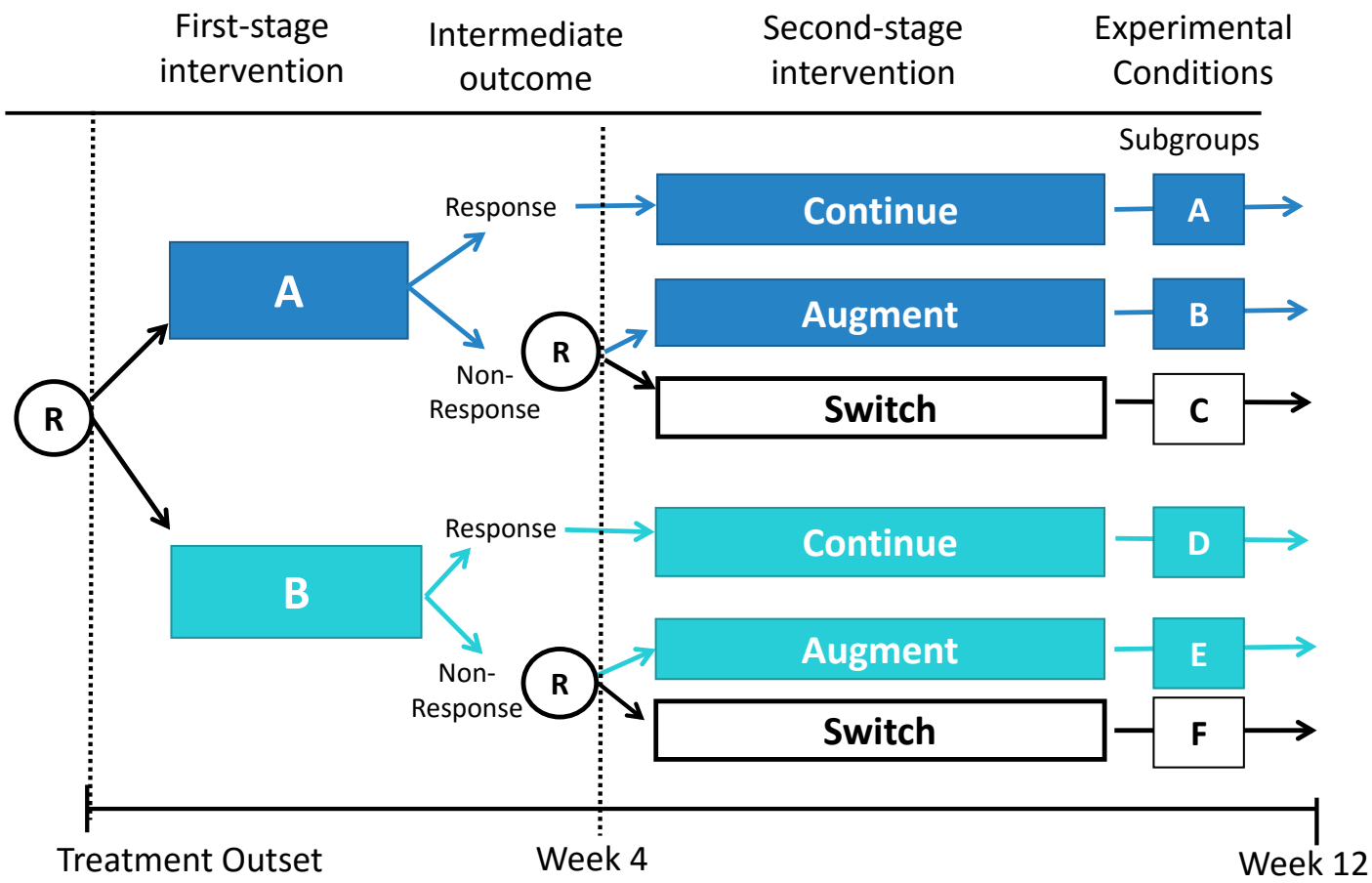
Common Primary Aims

Compare Second-Stage Treatments among Non-Responders



Common Primary Aims

Compare Two Embedded DTRs



Connecting Notation

Definition

A **decision rule** d_j is a function that takes intermediate outcomes prior to time j, $\bar{S}_j = \{S_1, \ldots, S_j\}$ and previous decisions $\{a_1, \ldots, a_{j-1}\}$ and outputs a treatment decision a_j.

- DTRs are a sequence of decision rules.
- Our DTRs of interest can be expressed as the sequence

$$\left\{ d_1(S_1),\ d_2(\bar{S}_2, a_1) \right\}$$

Connecting Notation

$$\left\{ d_1(S_1),\ d_2(\bar{S}_2, a_1) \right\}$$

- S_1: Baseline covariates
- a_1: Initial treatment
- S_2: Intermediate outcome after initial treatment
- a_2: Second-stage treatment
- S_3: Binary outcome

Recall that $\bar{S}_j = \{S_1, \ldots, S_j\}$

- O: Baseline covariates
- X_1: Indicator for initial treatment
- R: Indicator for response status
- X_{2NR}: Indicator for second-stage treatment among non-responders
- Y: Binary outcome

Initial SMART Data Structure

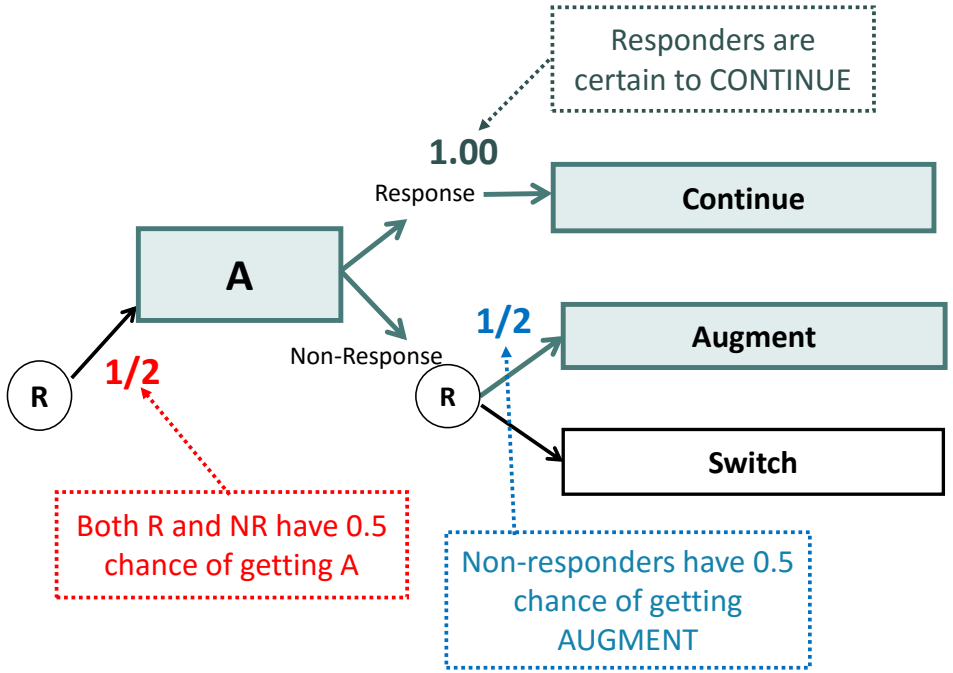
ID	Y	X_1	R	X_{2NR}	X_1X_{2NR}
1	1	1	1	NA	NA
2	0	0	0	0	0
3	0	0	0	1	0
4	1	0	1	NA	NA
5	1	1	0	1	1
6	0	1	0	1	1
⋮	⋮	⋮	⋮	⋮	⋮

- Y: Binary outcome
- X_1: Indicator for first-stage treatment
- R: Indicator for response status
- X_{2NR}: Indicator for second-stage treatment among non-responders

Inverse-Probability-of-Treatment Weighting

Unequal randomization creates imbalance. Here, responders are over-represented in each DTR we want to estimate.

- Responders have a $1/2 \times 1 = 1/2$ chance of getting their own DTR.
- Non-responders have a $1/2 \times 1/2 = 1/4$ chance of getting their own DTR.



Inverse-Probability-of-Treatment Weighting

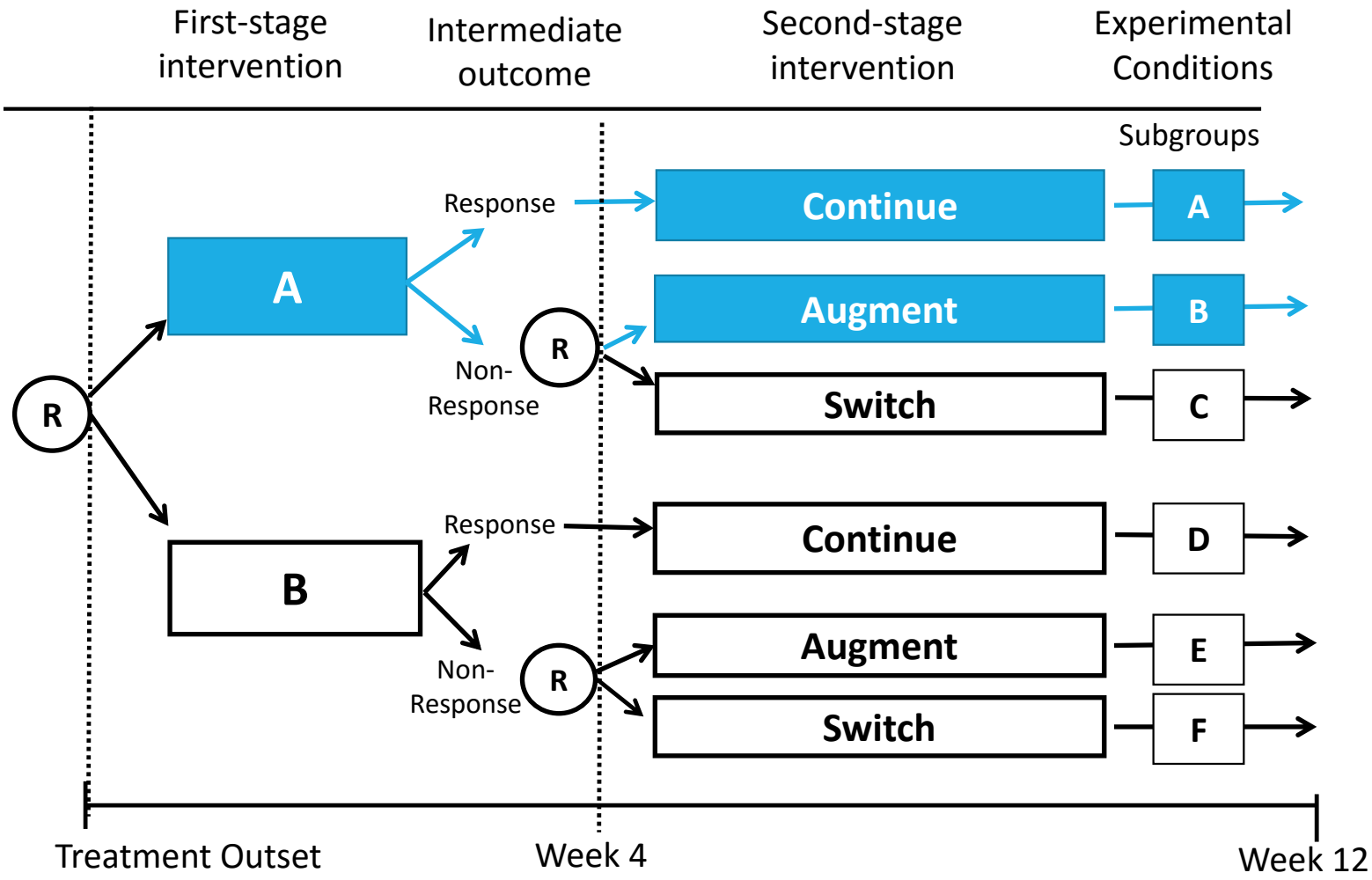
- Assign each observation a weight inversely proportional to its probability of receiving its own DTR.
  - Responders get weight $W = 1/(1/2) = 2$.
  - Non-responders get weight $W = 1/(1/4) = 4$.
- Distribution of weights depends on response rate (unknown *a priori*).
  - Robust (sandwich) variance estimation accounts for this

Updated SMART Data Structure

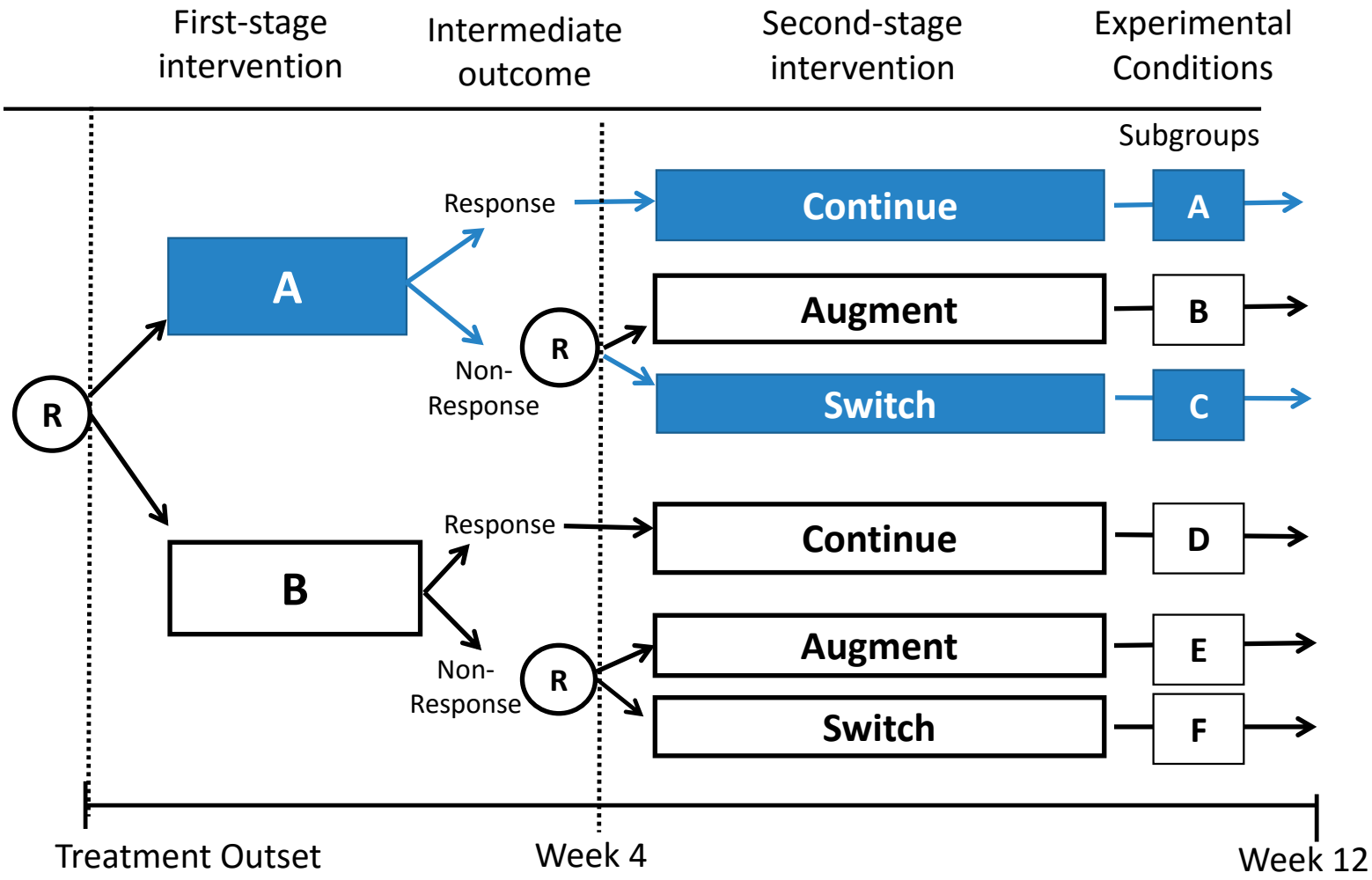
ID	Y	X_1	R	X_{2NR}	X_1X_{2NR}	W
1	1	1	1	NA	NA	2
2	0	0	0	0	0	4
3	0	0	0	1	0	4
4	1	0	1	NA	NA	2
5	1	1	0	1	1	4
6	0	1	0	1	1	4
⋮	⋮	⋮	⋮	⋮	⋮	⋮

- To compare DTRs, we need to know which observations are consistent with each DTR.
- But, by design, responders are consistent with more than one DTR!

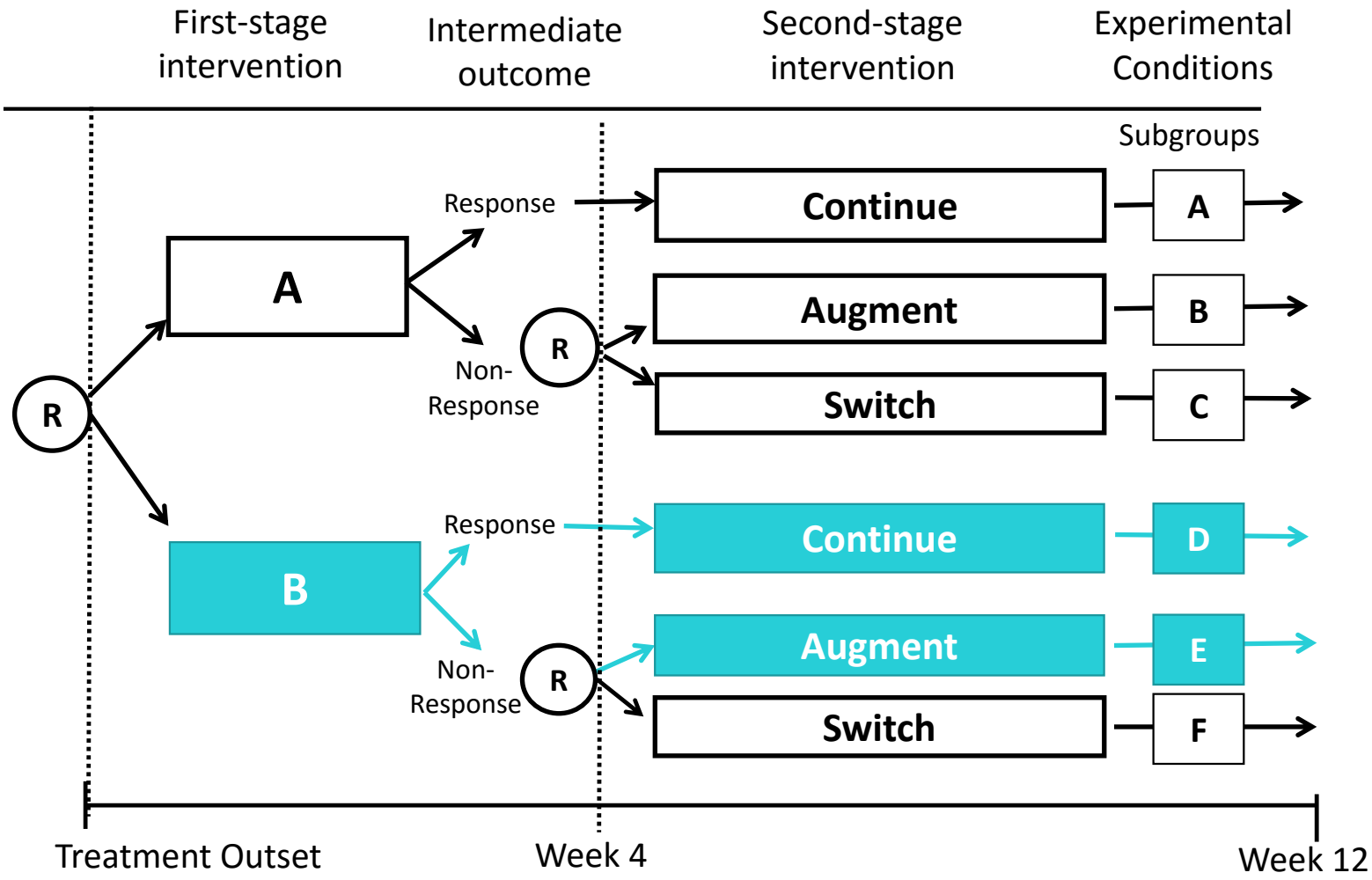
DTRs Share Responders



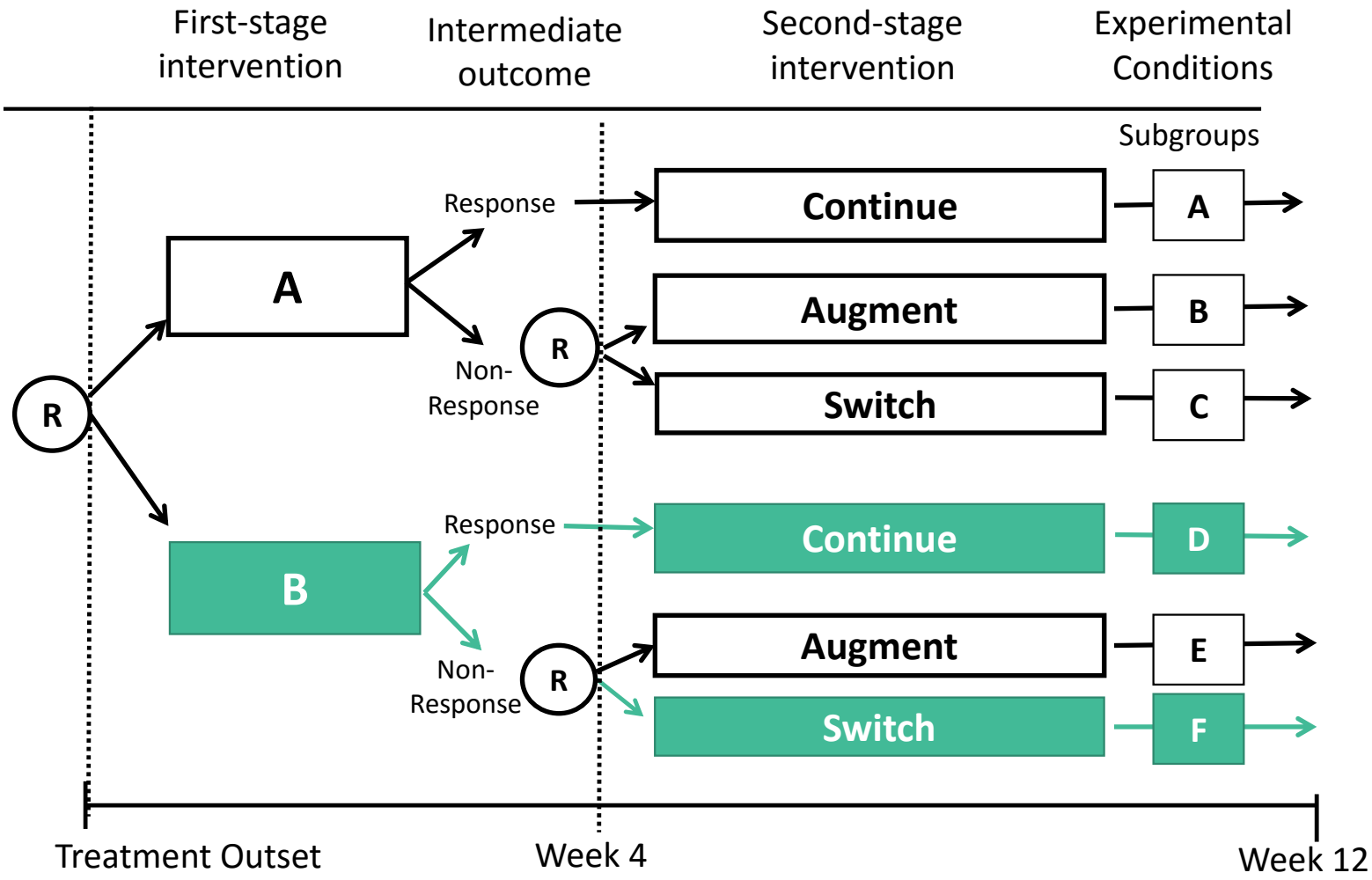
DTRs Share Responders



DTRs Share Responders



DTRs Share Responders



Recycling of Observations

- To use standard software, we need to somehow share responders between DTRs.
- We *recycle* observations of those participants consistent with more than one DTR (responders).
  - One observation gets $X_{2NR} = 0$, and the other gets $X_{2NR} = 1$.
- Robust (sandwich) variance estimation accounts for this.

SMART Data Structure for Analysis

ID	Obs.	Y	X_1	R	X_{2NR}	X_1X_{2NR}	W
1	1	1	1	1	0	0	2
1	2	1	1	1	1	1	2
2	1	0	0	0	0	0	4
3	1	0	0	0	1	0	4
4	1	1	0	1	0	0	2
4	2	1	0	1	1	0	2
5	1	1	1	0	1	1	4
6	1	0	1	0	1	1	4
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Data Analysis: Overview

We use weighted logistic regression. Our model is of the form

$$\text{logit}\left[P\left(Y \mid X_1, X_{2NR}\right)\right] = \beta_0 + \beta_1 X_1 + \beta_2 X_{2NR} + \beta_3 X_1 X_{2NR}.$$

- Interaction allows for non-additive effects between treatments in the DTR.
- Robust (sandwich) standard errors are needed to account for variation in weights across samples.
- Use `geepack::geeglm` in R, or `proc genmod` in SAS.
  - Independent covariance structure; may need to request robust variance estimator

Data Analysis: Example

$$\text{logit}\left[P\left(Y \mid X_1, X_{2NR}\right)\right] = \beta_0 + \beta_1 X_1 + \beta_2 X_{2NR} + \beta_3 X_1 X_{2NR}.$$

DTRs can be uniquely identified by a linear combination of Xs.

- **DTR 1**: "Give A; if response, continue; if non-response, augment.": $(1, 1, 1, 1)$
- **DTR 2**: "Give B; if response, continue; if non-response, augment.": $(1, 0, 1, 0)$

Estimate the difference in log-odds of success with a contrast matrix.

Data Analysis: Example

$$\text{logit}\left[P\left(Y \mid X_1, X_{2NR}\right)\right] = \beta_0 + \beta_1 X_1 + \beta_2 X_{2NR} + \beta_3 X_1 X_{2NR}.$$

The contrast matrix for DTR 1 - DTR 2 is

$$\mathbf{C} = (1,1,1,1) - (1,0,1,0) = (0,1,0,1).$$

Suppose we fit this model, and get the following estimate of β:

$$\hat{\beta} = (2.95, 0.13, -0.23, -0.19)'.$$

The estimated difference in log-odds is

$$\boldsymbol{C}\hat{\beta} = 0.13 - 0.19 = -0.06.$$

Sample Size: Preliminaries

- Sample size for clinical trials is chosen to address the primary aim.
- For primary aims comparing first- or second-stage treatments, standard methods apply.
- To compare two embedded DTRs that start with different treatments, we need special methods.
  - Analogous to comparing two independent treatment groups.
  - Choose size based on Wald test, then modify to account for SMART design.

Hypothesis Testing for the Primary Aim

A comparison of two embedded DTRs is a comparison of two linear combinations of βs. For this, we use a Wald test.

$$H_0 : E\left[\theta_{1,1}\right] - E\left[\theta_{0,1}\right] = 0,$$

where $\theta_{x_1,x_{2NR}} = \text{logit}\left(P(Y = 1 \mid X_1 = x_1, X_{2NR} = x_{2NR})\right)$.
The test statistic is

$$\frac{\left(\boldsymbol{C}\hat{\beta}\right)^2}{\hat{\text{Var}}\left(\boldsymbol{C}\hat{\beta}\right)}$$

Sample Size

Sample Size Formula

The sample size N for a SMART with a binary outcome where the primary aim is to compare two embedded DTRs is given by

$$N = 2\left(z_\beta + z_{\alpha/2}\right)^2 \frac{(1+A)^2 B + (1+AB)^2}{AB \ln^2 B} \times (2(1-r)+r)$$

- z_p: Standard normal $1-p$ quantile
- $1-\beta$: Target power for test
- α: Significance level of test
- A: Odds of success for "reference" DTR
- B: Odds ratio of two DTRs of interest
- r: Hypothesized response rate

Web-Based Sample Size Calculation

- Key for improving accessibility of SMARTs
- Often can be quite difficult to use

`https://nseewald1.shinyapps.io/SMARTsizeBeta`

Acknowledgements

- Alex Giessing, Jingshen Wang
- Inbal Nahum-Shani, Survey Research Center
- Susan A. Murphy, Department of Statistics
- Jeremy M.G. Taylor, Department of Biostatistics

- N.J.S. acknowledges funding from NIH, grant 5T32CA083654-12
- D.A. acknowledges funding from NIDA, grant P50DA010075; NIMH, grant R03MH097954; and NICHD, grant R01HD073975.

References

- Murphy, S.A. (2005). "An experimental design for the development of adaptive treatment strategies", *Statistics in Medicine*, 24, 1455-1481.
- Kasari, C., Kaiser, A., Goods, K., Nietfeld, J., Mathy, P., Landa, R., ... & Almirall, D. (2014). "Communication Interventions for Minimally Verbal Children With Autism: A Sequential Multiple Assignment Randomized Trial", *Journal of the American Academy of Child & Adolescent Psychiatry*, 53(6), 635-646.
- Nahum-Shani, I., et al. "Experimental Design and Primary Data Analysis Methods for Comparing Adaptive Interventions", *Psychological Methods*, 17(4): 457-477.