SMART Designs

KELLEY M KIDWELL, PHD

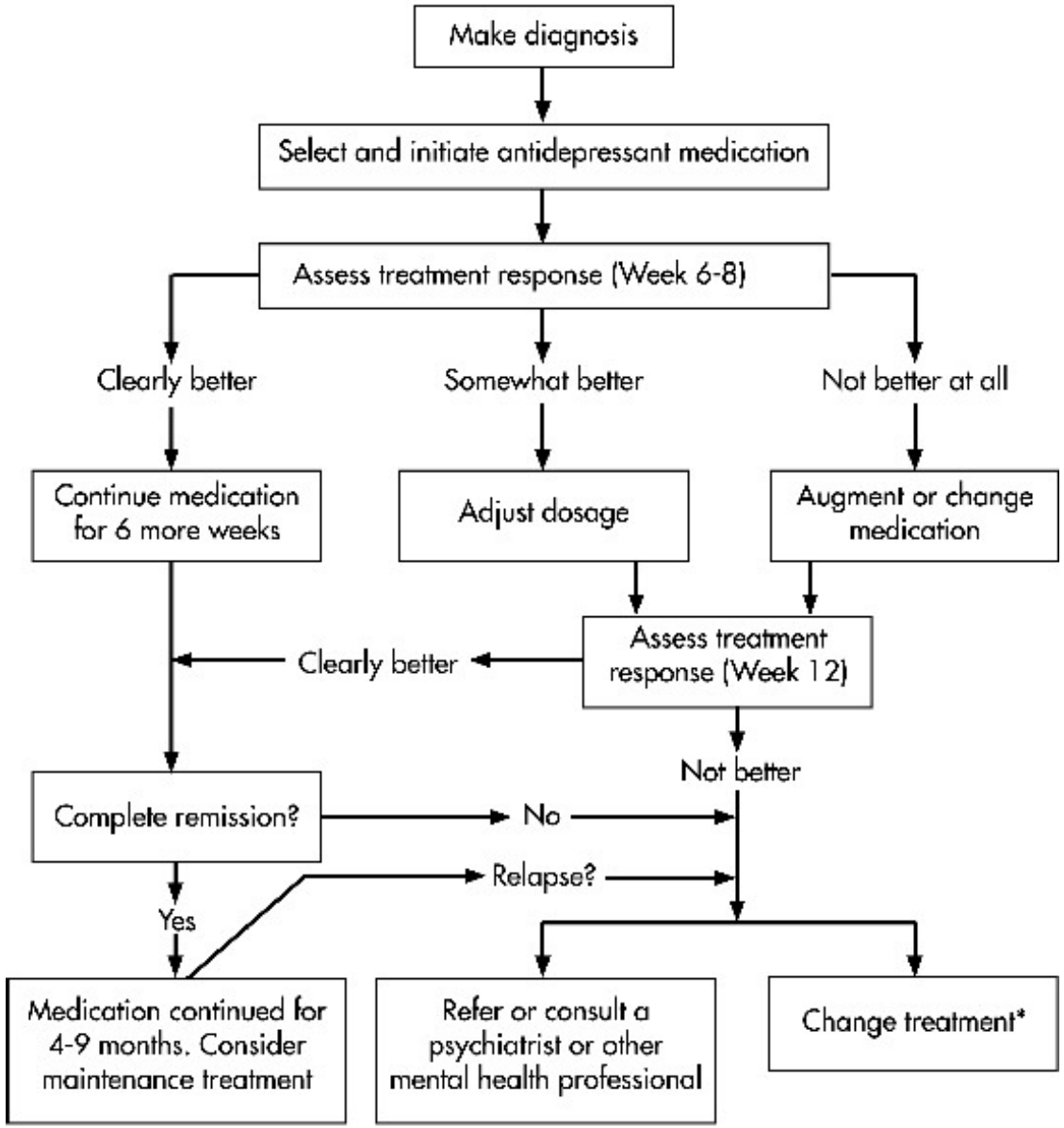
UNIVERSITY OF MICHIGAN

ACSTAT 2017

BALTIMORE, MD

How do providers treat individuals in real life?

- Ongoing care and follow up
- Therapies are not set in stone
- Therapies can be changed, intensified, discontinued
- Treatment decisions can be based on health progress, treatment adherence, side effects, and patient choice
- Follow-up therapy based on experience, guidelines, clinical trials



*Short-term behavior counseling has been shown to be just as effective as anti-depression medication

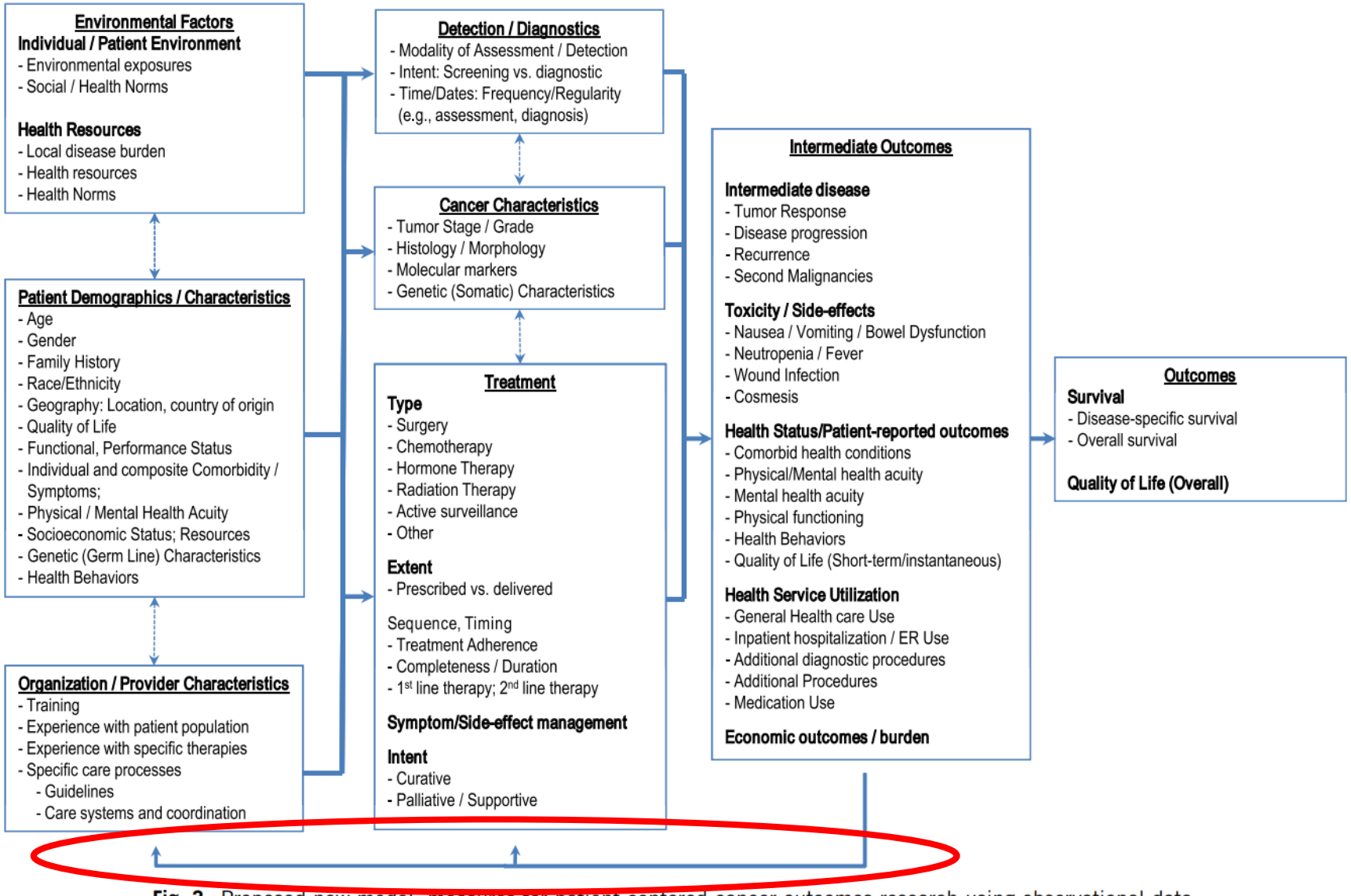


Fig. 3. Proposed new model: measures for patient-centered cancer outcomes research using observational data.

Dynamic Treatment Regimens (DTRs)

- a.k.a. adaptive intervention, adaptive treatment strategy, stepped care, treatment policies

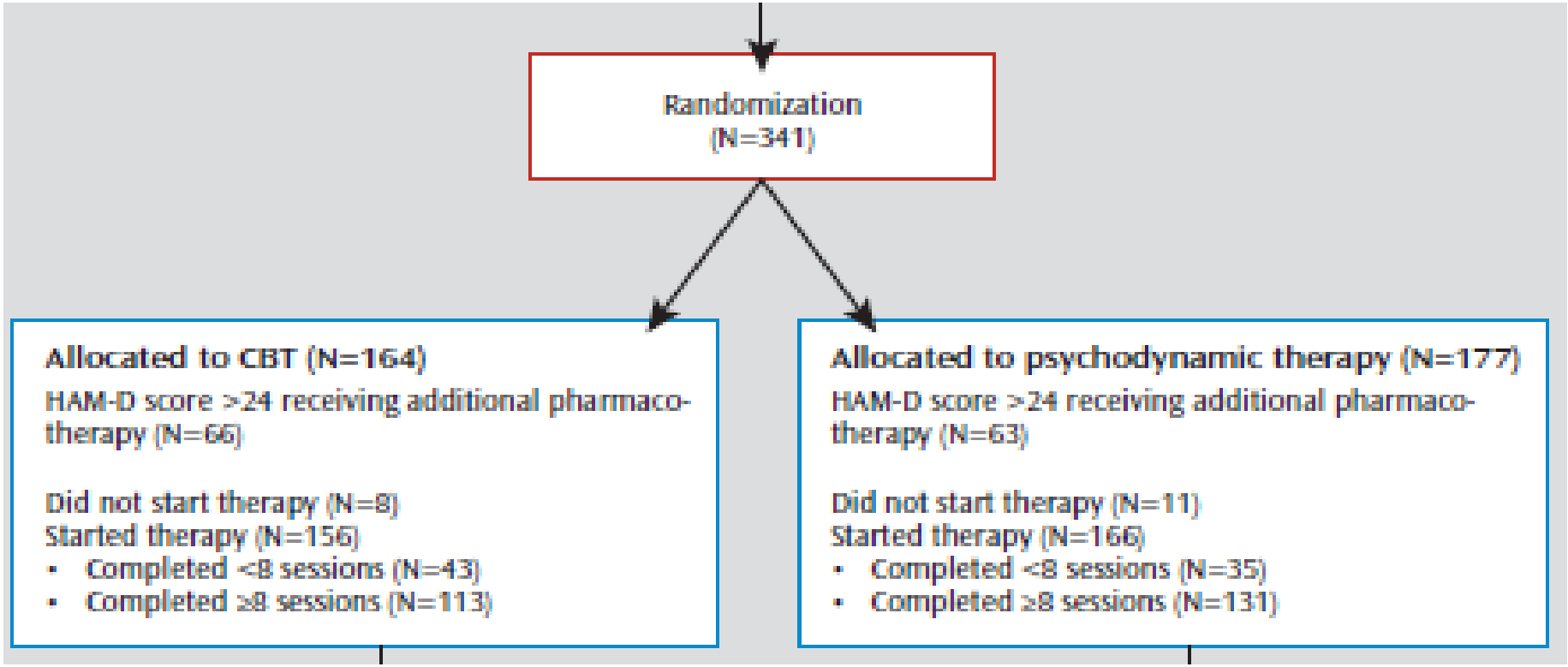
- Sequence of **individually tailored decision rules** that specify whether, how and/or when to alter the intensity, type, dose or delivery of treatment at critical decision points in the course of care
- **Guide/Formula for treatment**

- Goal: operationalize sequential decision making with the aim of improving clinical practice

Dynamic Treatment Regimens (DTRs)

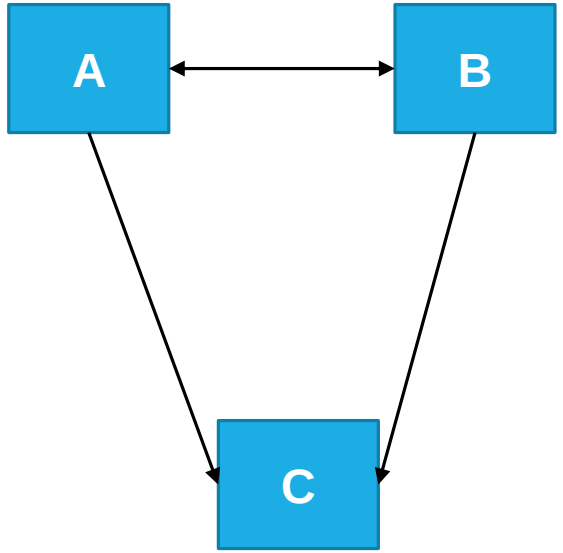
- **Prostate Cancer:**
  - First receive combination chemotherapy paclitaxel + estramustine + etoposide (TEC).
  - If successful (a decrease of 40% or more in PSA from diagnosis, with no evidence of disease progression at any site) at 8 weeks, then continue TEC;
  - otherwise switch to cyclophosphamide + vincristine + dexamethasone (CVD).
- **Alcohol Abuse**
  - First take naltrexone and receive in person medical management
  - After 2 weeks, but before 8 weeks, if the individual has 2 or more heavy drinking days then add cognitive behavioral therapy
  - After 2 weeks, but before 8 weeks, if the individual has less than 2 heavy drinking days, replace in person medical management with telephone disease management

How do we often study treatments?



Some Consequences

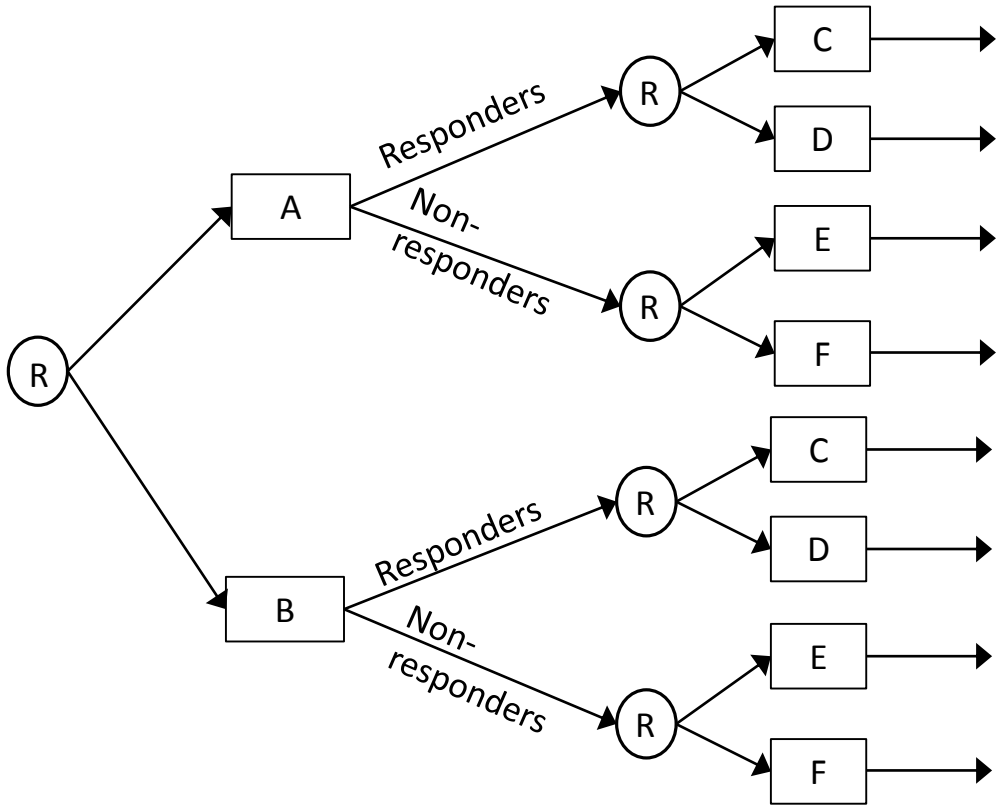
- Compare treatments A vs. B for first-line treatment
  - Response Rates
    - A: 60%
    - B: 50%
  - Treatment A wins
- Test efficacy of second-line treatment C for non-responders
  - Response Rates
    - A followed by C: 10%
    - B followed by C: 40%
  - Treatment B followed by C wins
- Overall sequence
  - A,C: 64% (60%+40%*10%)
  - B,C: 70% (50%+50%*40%)



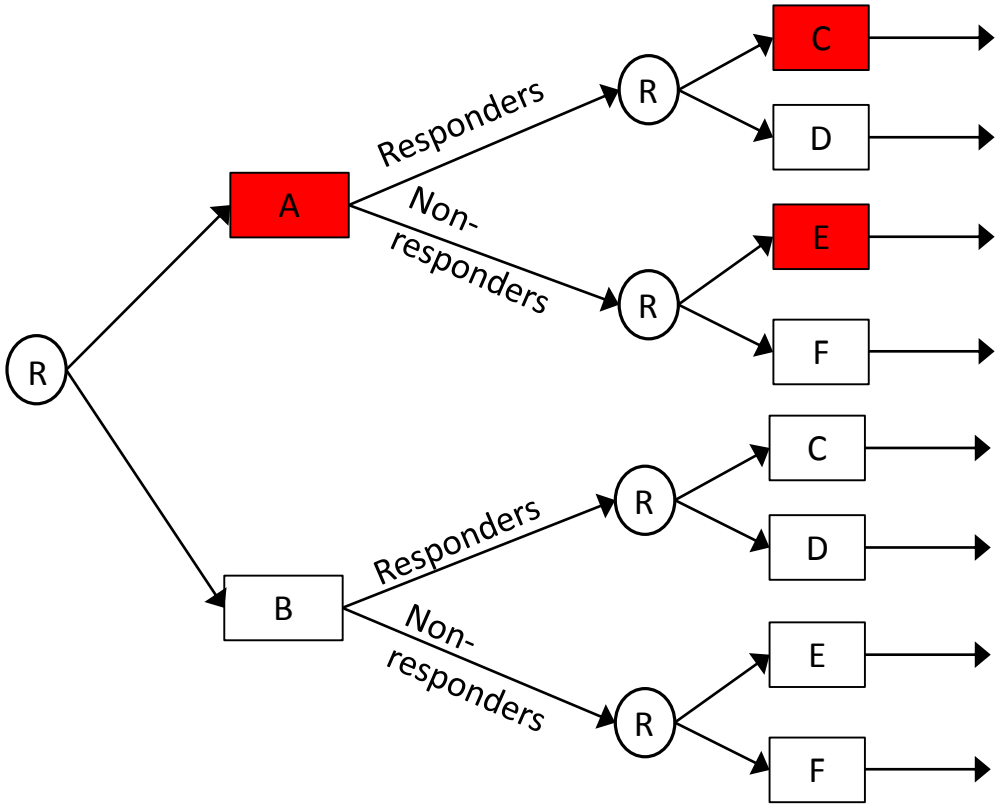
Sequential Multiple Assignment Randomized Trial

- A type of **multi-stage randomized design**
- Trial participants are randomized to a set of treatment options at critical decision points over the course of treatment
- **All individuals** participate in all stages of the trial
- Subsequent randomization is based on information leading up to that point (tailor treatment)
- DTRs embedded in design
- Goal: **develop effective DTRs**

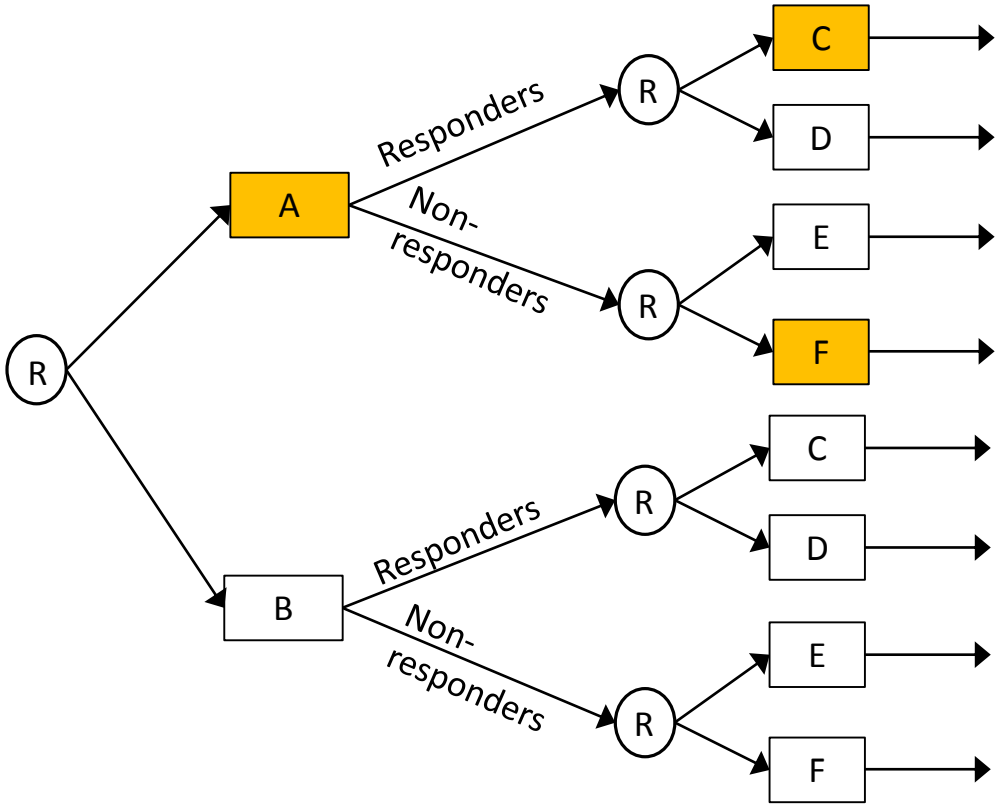
SMART Example



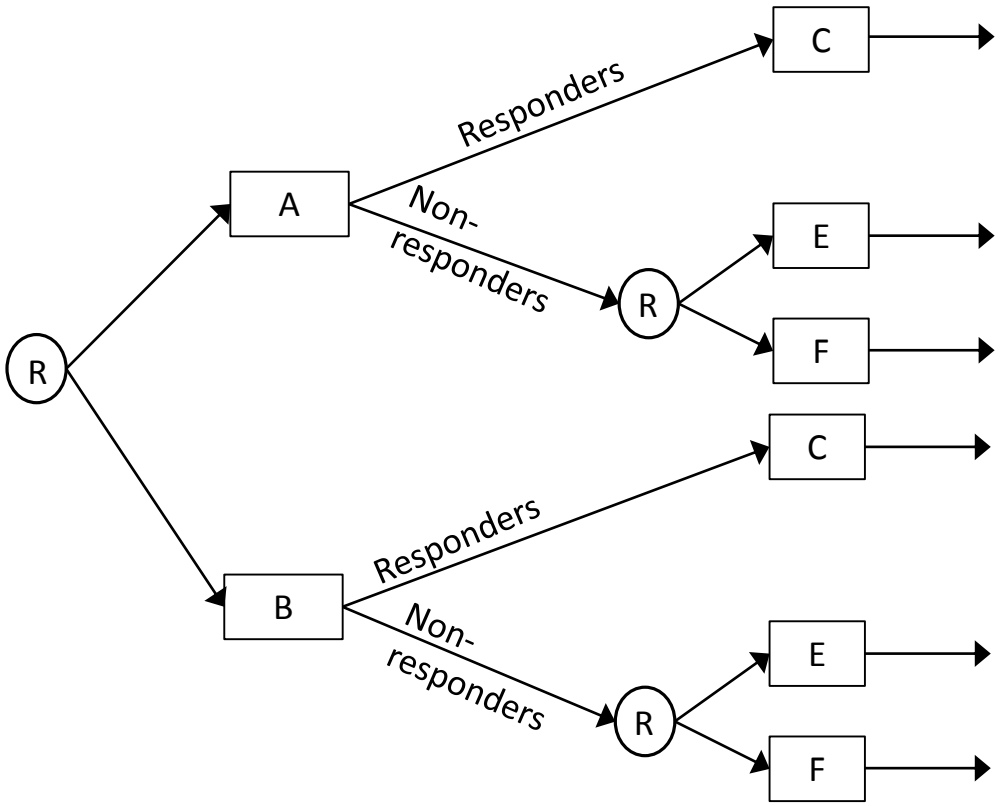
Embedded DTRs



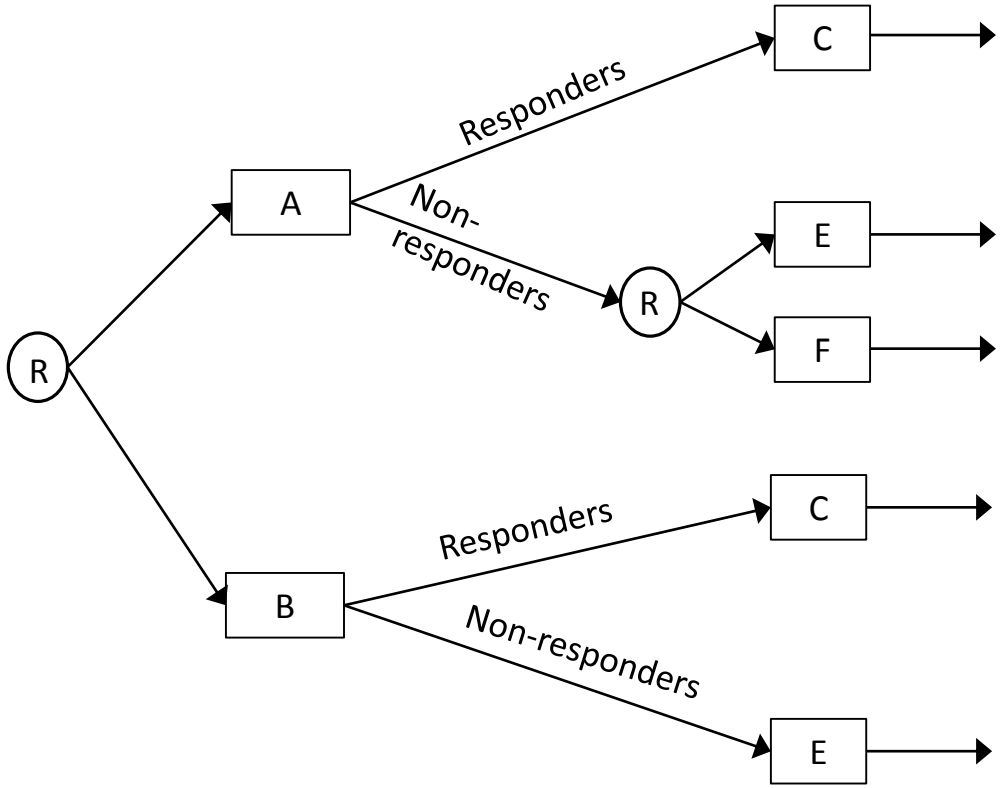
Embedded DTRs



SMART Example 2



SMART Example 3



SMARTs in the Field

- Oncology
- Drug abuse
- ADHD
- Alcoholism
- Obesity
- OCD
- Autism
- Schizophrenia
- Depression
- Insomnia
- Bipolar
- Conduct problems
- Smoking cessation
- Suicide prevention

https://methodology.psu.edu/ra/adap-inter/projects

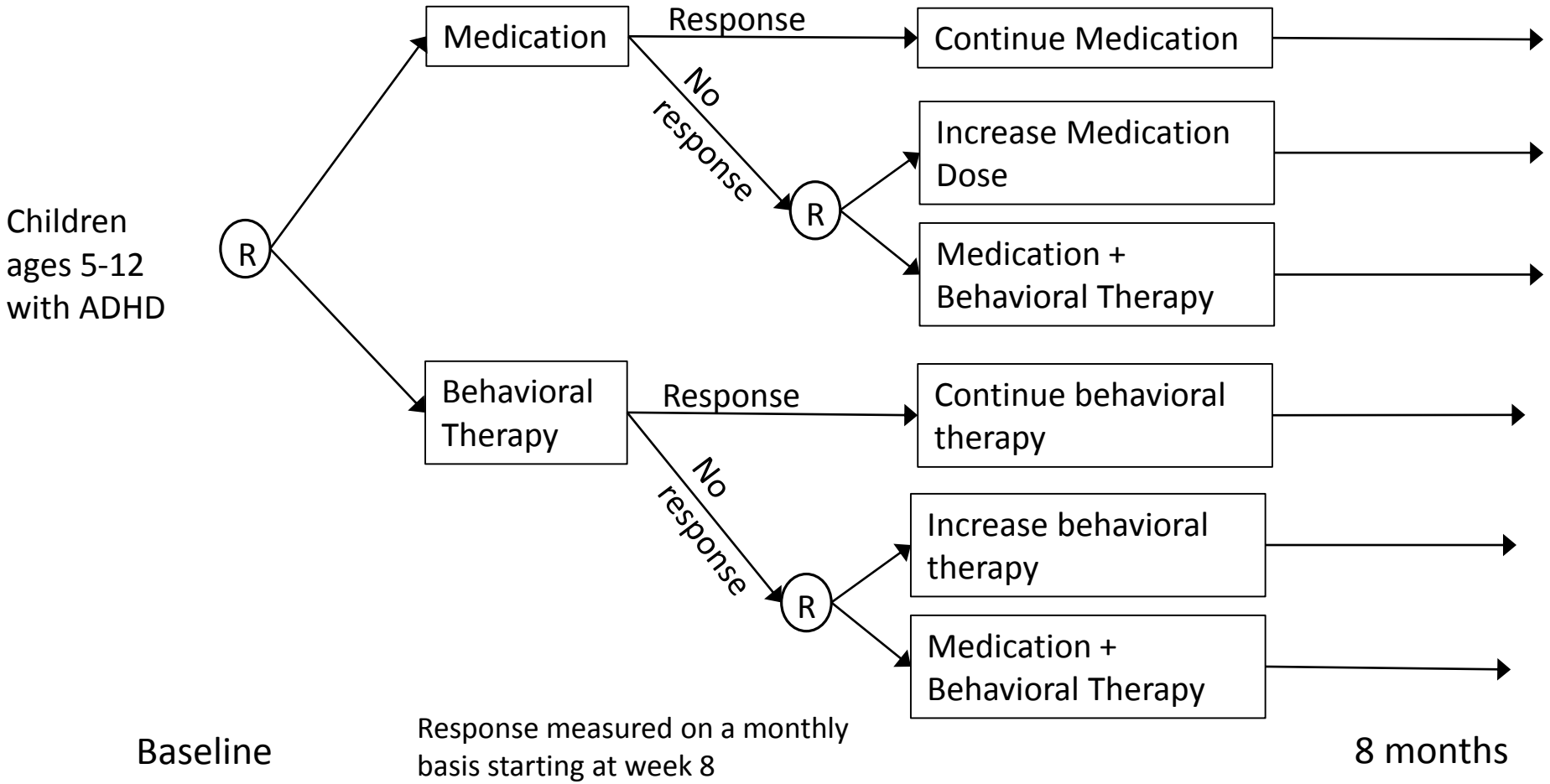
SMART Benefits

- **Delayed Effects** – treatment synergies or antagonisms
- **Prescriptive Effects** – initial treatment may elicit symptoms to better match individual to subsequent treatment
- **Sample Selection Effects** – individuals who enroll in, remain in or are adherent in a SMART may be different from those in other designs

Case Study: ADHD SMART

- PI: Bill Pelham
- Children with **Attention Deficit Hyperactivity Disorder** ages 6-12
- American Psychological Association recommended behavioral therapy first (2007) whereas American Academy of Child and Adolescent Psychiatry recommended using medication first (2007)
- 20%-50% of children can be expected to insufficiently respond to the initial intervention
- **Study goal**: To understand whether to begin with medication or behavioral therapy for children with ADHD, and whether to intensify or augment initial treatment for children who do not respond to treatment

Case Study: ADHD SMART



Case Study: ADHD

- **Outcome:** school performance from the Impairment Rating Scale (IRS) at 8 months

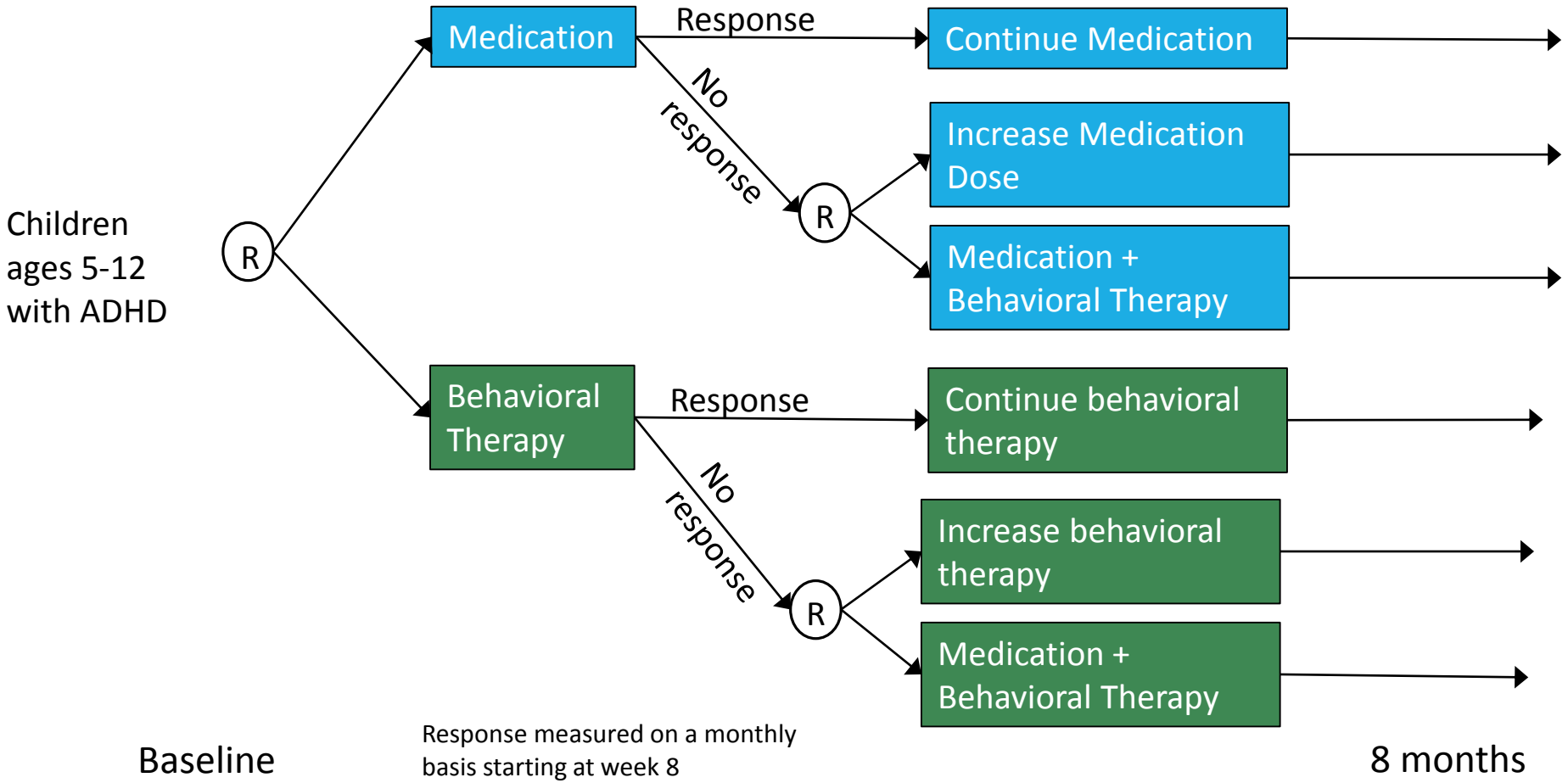
- **Response**: Individualized List of Target Behaviors (ITB) and IRS
  - Non-response: An average performance <75% on the ITB and a rating of impairment in at least one domain in the IRS

- **Primary Aim**: Is beginning with low-dose medication versus beginning with low-intensity behavioral therapy best?

Case Study: ADHD

- **Secondary Aims**:
  - Is it best to have non-responders intensify or add treatment?
  - What is the most effective DTR?
- **Exploratory Aims:** How do baseline variables (e.g. severity of impairment, comorbid child psychopathology and prior medication history) influence
  - the difference between the two initial treatments
  - the difference between the second stage intervention options for nonresponding children
  - the difference between the 4 embedded dynamic treatment regimens

ADHD SMART: Primary AIM

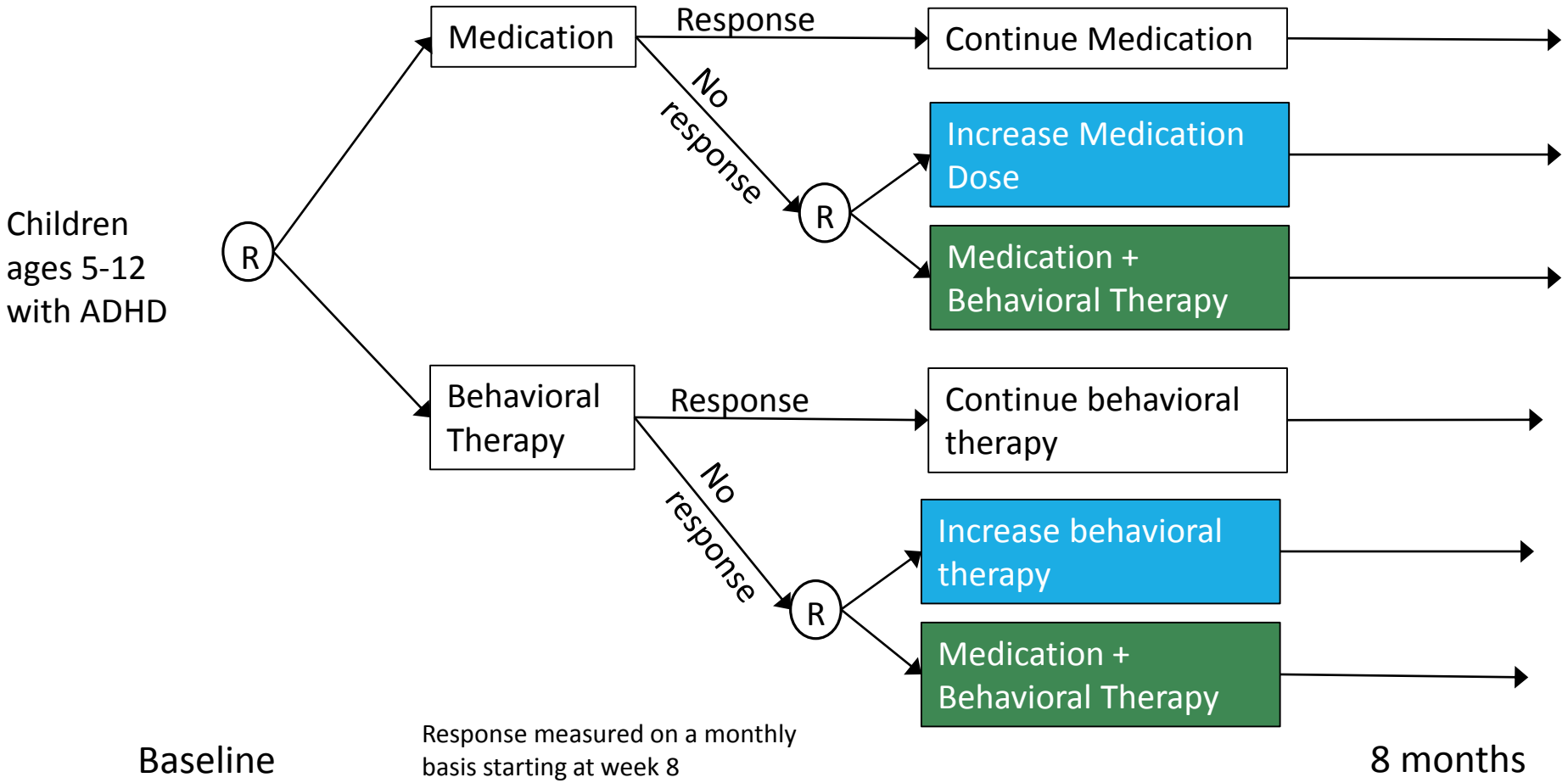


ADHD SMART: Primary Aim

Is it best to start with medication or behavioral therapy?

- **Sample Size Calculation: 2 arm comparison (t-test)**
  - With 68/group (136 total), we have 82% power to find a medium standardized effect size of 0.5 between the two groups.
- **Analysis**
  - Standard 2 group comparison: t-test, linear regression controlling for pre-planned covariates
- **Why do a SMART**
  - Obtain information regarding second-stage treatment and DTRs even if not powered for that can influence subsequent trials

ADHD SMART: Secondary Aim

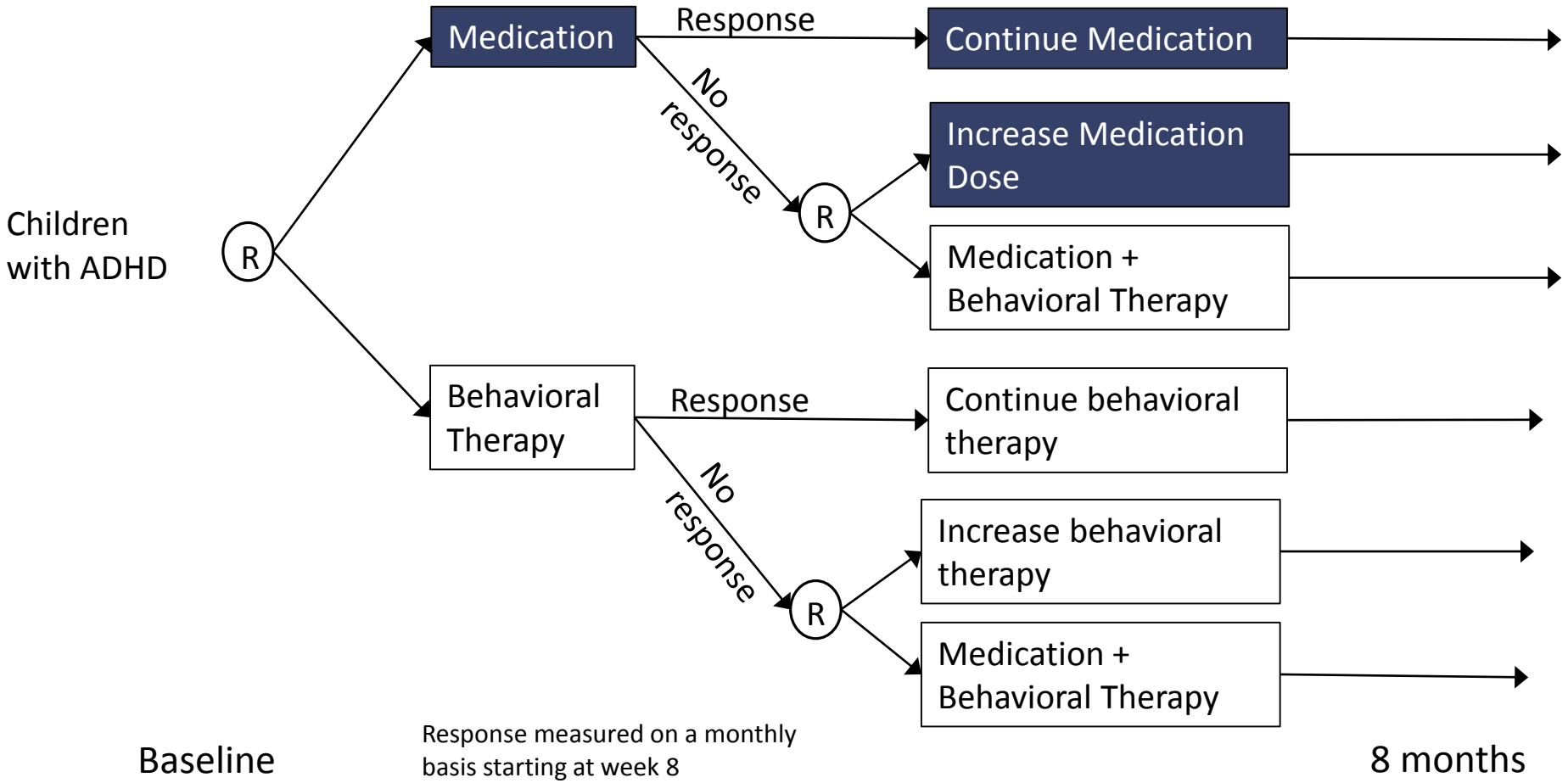


ADHD SMART: Secondary Aim

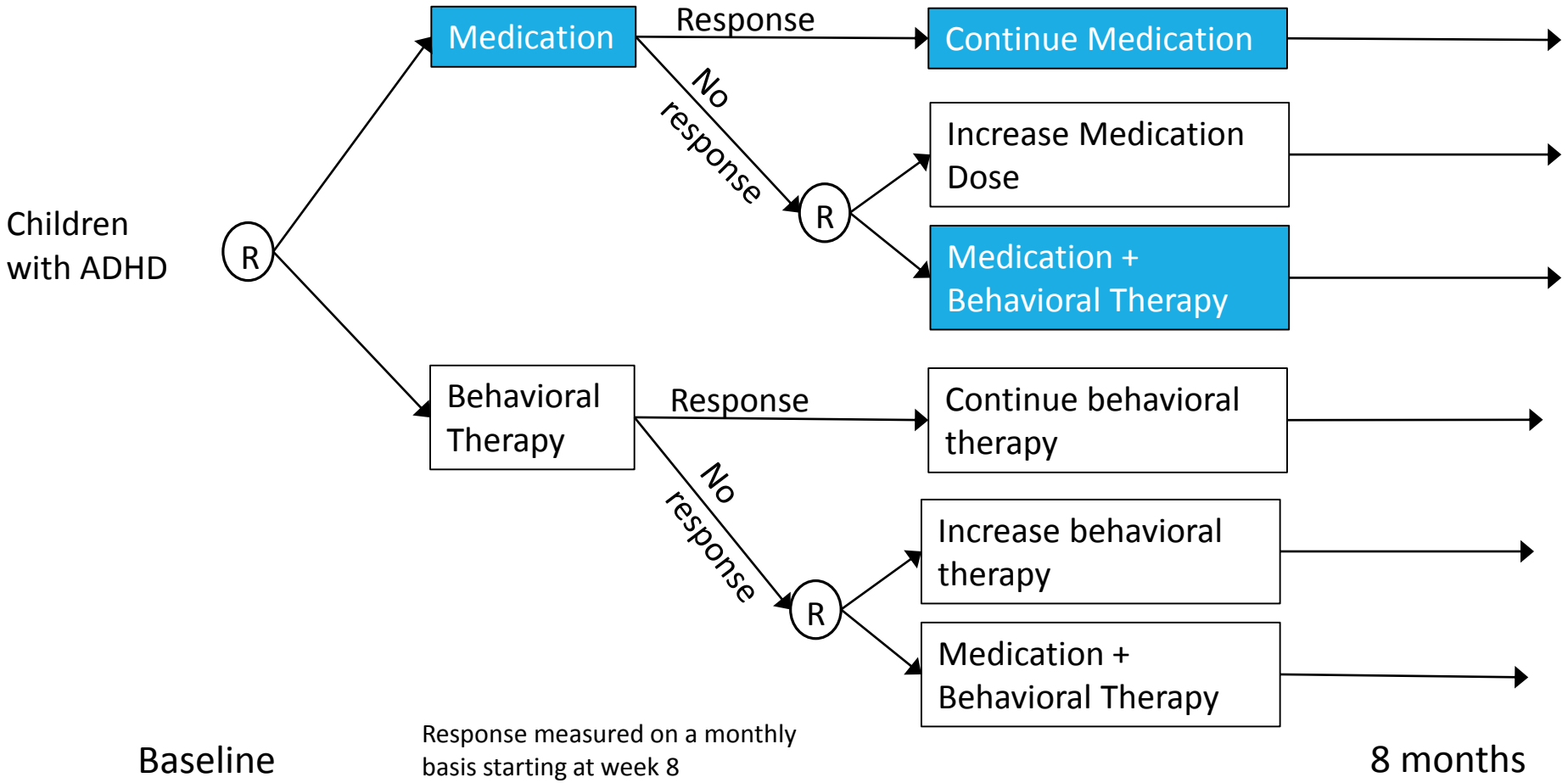
Is the best strategy for non-responders to increase initial stage therapy or add therapy?

- **Sample Size Calculation: 2 arm comparison (t-test)**
  - Assume: 50% non-response
  - With 40/group of non-responders (136 total), we have 80% power to find a medium to large standardized effect size of 0.7 between the two groups.
  - To find an effect size of 0.5, 256 total children are needed so that there are 64 non-responders per group
- **Analysis**
  - Standard 2 group comparison: t-test, linear regression controlling for pre-planned covariates and first-stage intervention
- **Why do a SMART**
  - Obtain information regarding DTRs even if not powered for that can influence subsequent trials
- **Potential Problems**
  - Less likely to be interested in strategy and more likely to be interest in best specific treatment/DTR

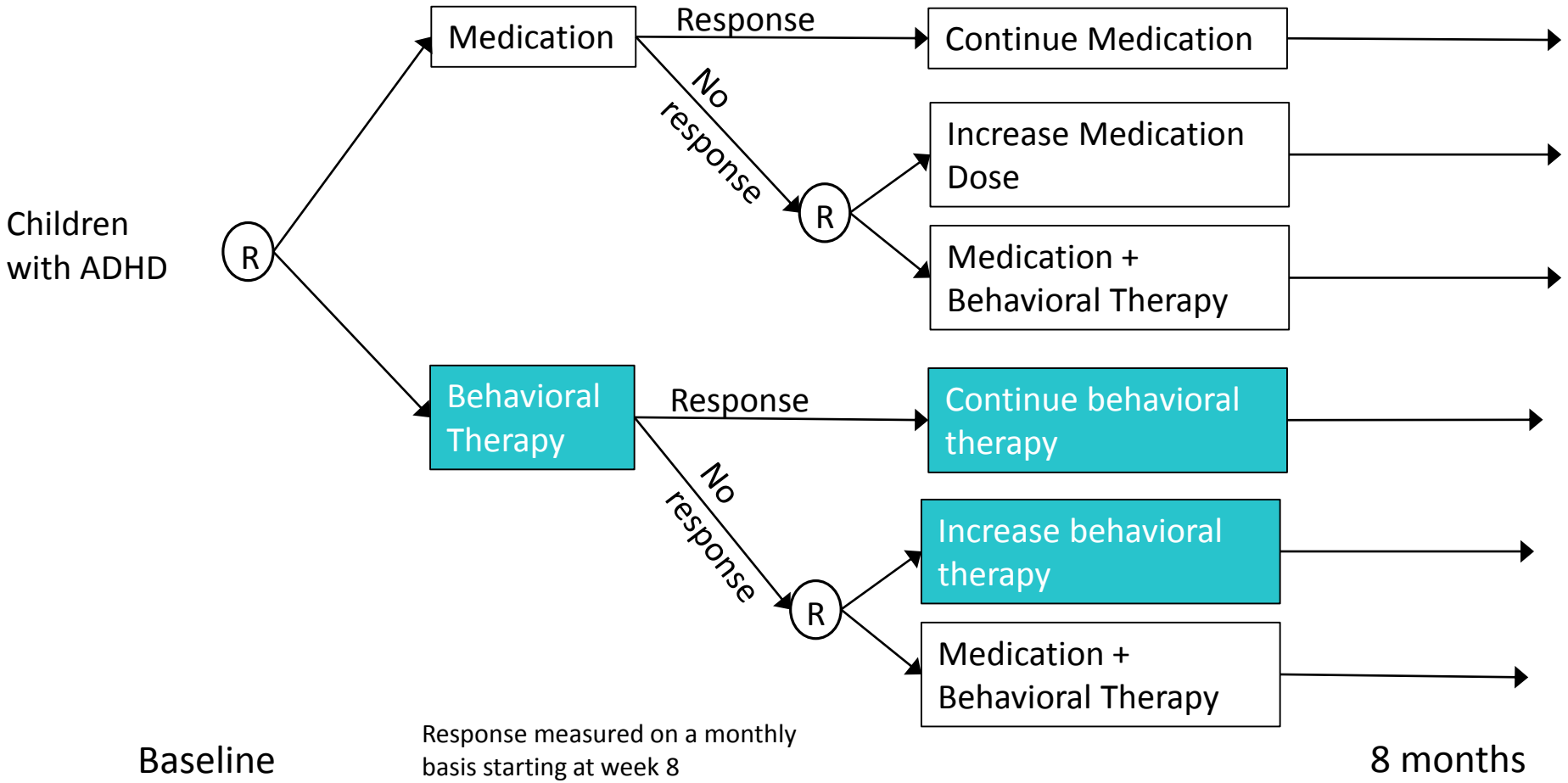
ADHD SMART: Compare DTRs



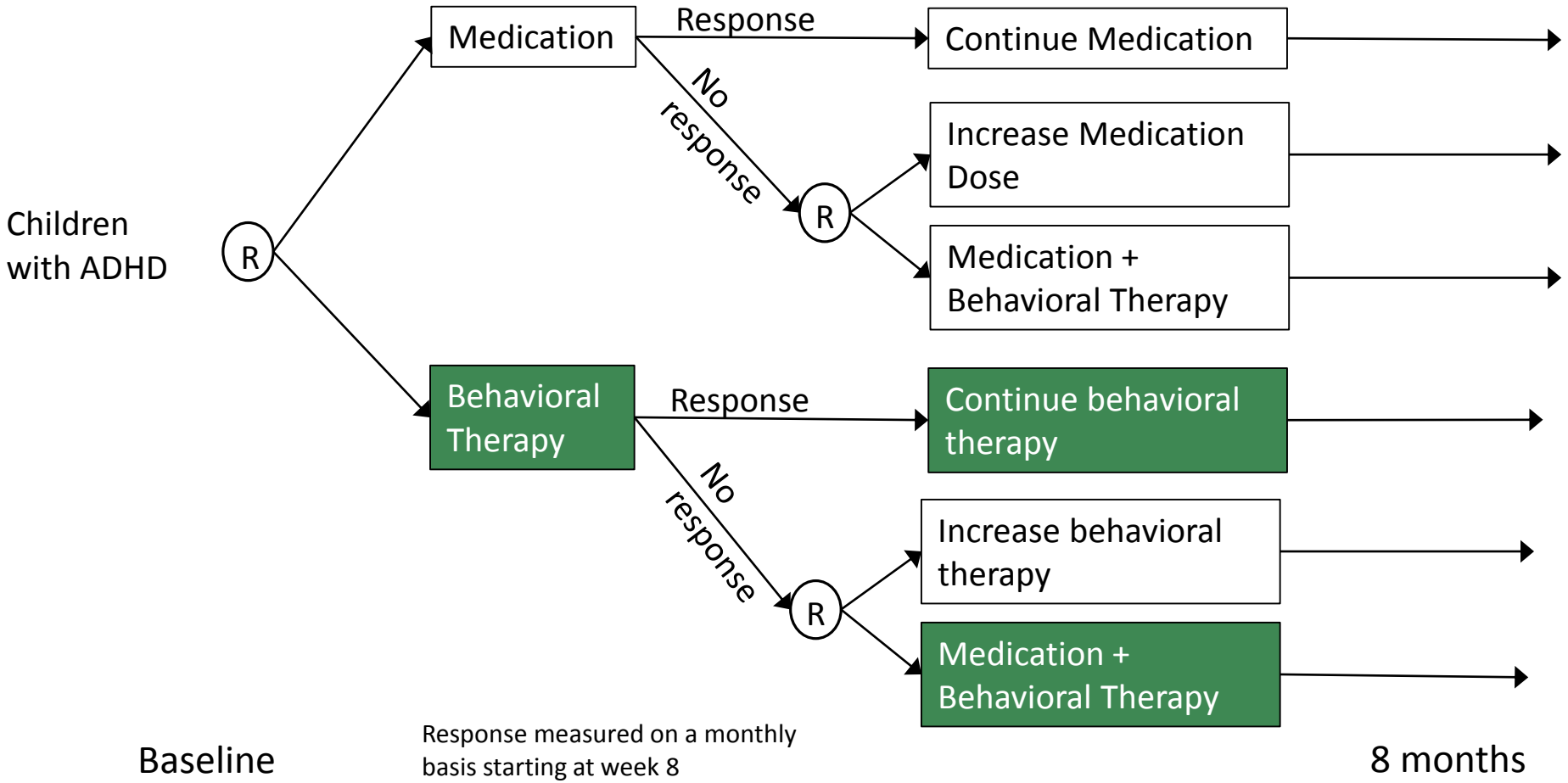
ADHD SMART: Compare DTRs



ADHD SMART: Compare DTRs



ADHD SMART: Compare DTRs



Secondary Aim: Compare DTRs

What is the most effective DTR?

- **Sample Size Calculation:** effect size of 0.5, 50% response
  - **Pairwise comparisons** of DTRs that begin with different treatment and multiple comparison correction (6): n=1199
  - **Global Hypothesis test:** n=315
  - **Estimation:** n=57

- **Analysis**: Weighted and Replicated Regression

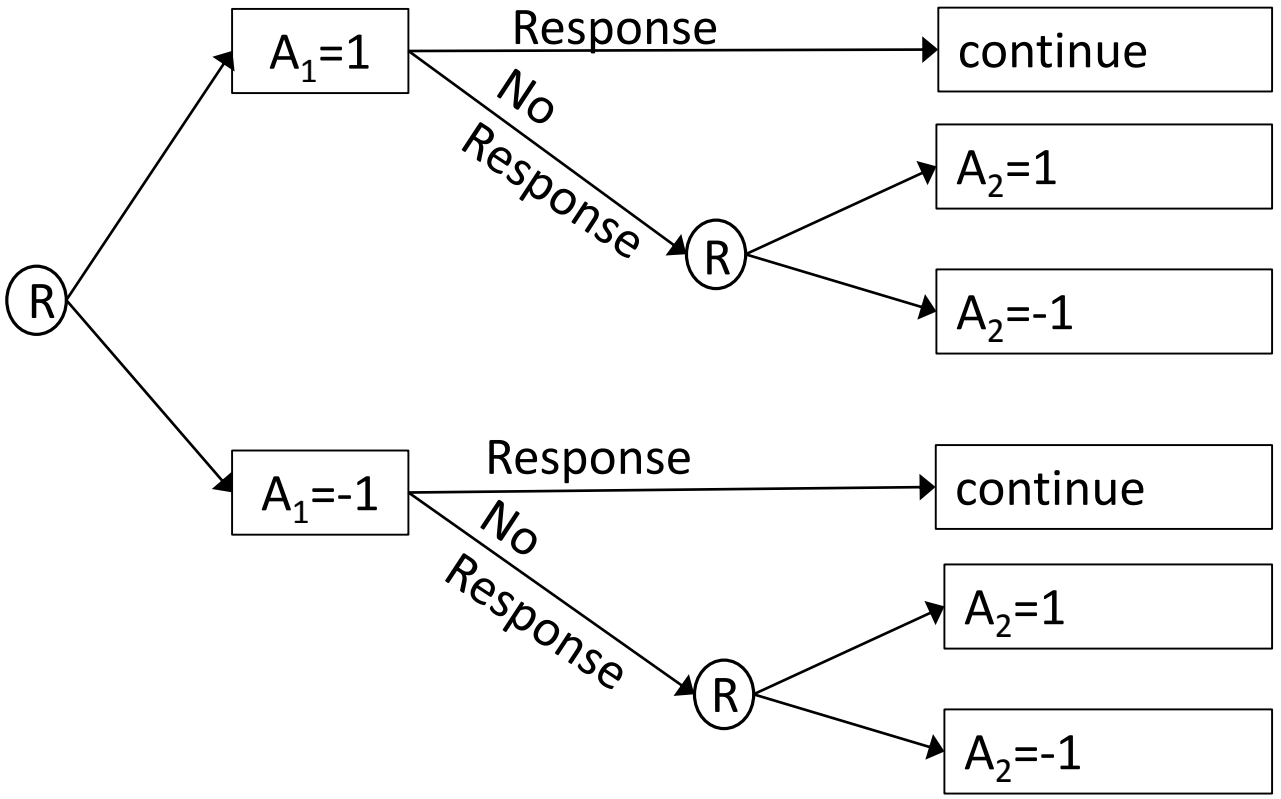
Sample Size Comparison

Aim	N	Analytic Method	Resource
Compare First-stage treatments (effect size 0.5)	136	Standard	Standard
Compare Non-responder Strategy (effect size 0.5, response prob=0.5)	256	Standard	Standard
6 Pairwise comparisons of DTRs (effect size 0.5 between DTRs 1, 3 and 4; .25 between DTRs 1 and 2; 0 between DTRs 3 and 4)	1199	Weighted and Replicated Regression	Nahum-Shani et al. Psych Methods 2012 https://sites.google.com/a/umich.edu/kidwell/home/tools-for-design-and-analysis
Global DTR comparison (same as above)	315	Weighted and Replicated Regression	Ogbagaber, Karp and Wahed, SIM 2015 (simulation)
Estimate best DTR (effect size 0.5, 3 DTRs have the same mean and 1 has the largest)	57	Weighted and Replicated Regression	http://methodologymedia.psu.edu/smart/samplesize

Weighted and Replicated Regression

- **Analysis:** Weighted and Replicated Regression
  - **Weights**: accommodate over/under representation due to restricted randomization scheme (or unequal randomization); For ADHD: W=2 for responders and W=4 for non-responders (depends on design)
  - **Robust standard errors** needed to account for sample to sample variation in the distribution of weights for appropriate inferences with weighted averages
  - **Replication**: trick to have standard software simultaneously estimate all DTRs; replicate those consistent with more than one DTR (responders) and fill in missing second-stage information balanced between both
  - **GEE** estimation with independent covariance matrix

SMART Setup


R
A_1=1
Response
continue
No Response
R
A_2=1
A_2=-1
A_1=-1
Response
continue
No Response
R
A_2=1
A_2=-1

End-of-Study Outcome

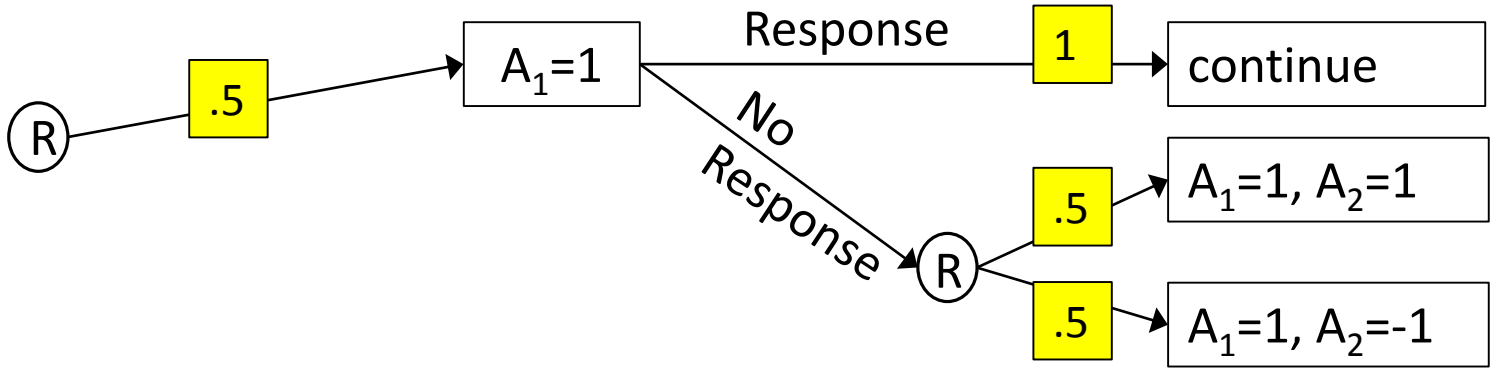
- Simpler analysis
- Compare end of study means (Y)
- Good for conditions/diseases with clear end targets or events
- Basic marginal model (design dependent equation):

$$E_{(a_1,a_2)}(Y|X) = \eta X + \beta_0 + \beta_1 a_1 + \beta_2 a_2 + \beta_3 a_1 a_2$$

- Where X is a vector of centered co-variates, a_1 is the first randomization assignment, and a_2 is the second randomization assignment.

Weighting

- Imbalance by design
- Imbalance will cause bias in estimates if we do not account for it, so we assign weights derived from the probability of being given a particular assignment
- Weights are assigned according to inverse of probabilities: Since $P(A_1)=.5$ and $P(A_2|\text{nonresponse})=.5$, $W_R=2$ and $W_{NR}=4$



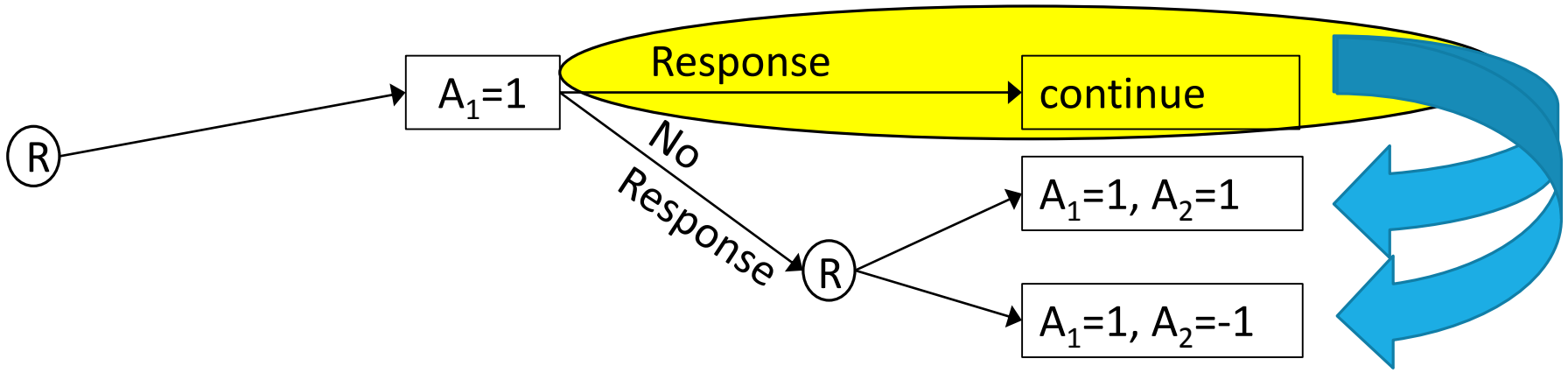
Gaining efficiencies in weighting

- In addition to design weights (previous slide)
- Robins and colleagues (1995), Hirano et al (2003) and others describe potential gains in statistical efficiency by estimating the weights using auxiliary baseline (L_1) and time-varying (L_2) covariate information
- Example:
  - The observed data is $\{L_{1i}; X_i; A_{1i}; L_{2i}; A_{2i}; Y_i\}$
  - Use logistic regression to get $\hat{p}_{1i} = \hat{P}(A_1 | L_1, X)$
  - Use logistic regression to get $\hat{p}_{2i} = \hat{P}(A_2 | L_1, X, L_2, \textit{non-response})$
  - Use $W_{Ri}=1/\hat{p}_{1i}$, $W_{NRi}=1/(\hat{p}_{1i}\,\hat{p}_{2i})$
  - Key is to use L's that are highly correlated with Y

Review of weighting and replication

Those who responded to A_1=1, are consistent with 2 DTRs:

- Offer A_1=1. If patient responds, then continue.
  If disease progresses, then offer A_2=1
- Offer A_1=1. If patient responds, then continue.
  If disease progresses, then offer A_2=-1



Review of weighting and replication

- Responders were not given treatment assignments in A_2. Their rows are duplicated. One row is assigned A_2 =1, and the other is assigned A_2 =-1. Thus, each row corresponds to one DTR. Responder data is now associated with both DTRs where A_1=1.
- Non-responders will have a single row associated with the DTR consistent with the treatment sequence they experienced.

Responder

Non-responder

ID	A1	A2	Y	Med before Stage 1	BL ODD Dx	BL ADHD Severity
101	1	-1	4	1	0	2
101	1	1	4	1	0	2
102	1	1	1	0	1	0
103	1	-1	0	1	0	2
103	1	1	0	1	0	2
104	-1	-1	2	1	0	1
105	1	-1	3	0	1	2
105	1	1	3	0	1	2
106	-1	-1	1	1	0	3
106	-1	1	1	1	0	3

Estimating Equation for End of Study Analysis

The weighted and replicated estimator for β is obtained by solving the following estimating equation:

$$0 = \sum_{i=1}^{N} \sum_{(a_1,a_2)} I\{a_1, a_2\} d(Xi, a_1, a_2) Wi \cdot (Y_8 - \mu_8 (Xi, a_1, a_2; \beta, \eta))$$

Where…

- $I\{a_1,a_2\}$: treatment sequence of individual *i* consistent with protocol $\{a_1,a_2\}$:
- Y_8: The outcome at month 8.
- μ_8 : The marginal mean of Y_8.
- W_i: Inverse-proportioned weights for each individual.
- $d(X_i, a_1, a_2)$ Derivative with respect to β.

Modeling and Estimating in R

```
GEE_endofstudy  <- geeglm(formula = Y ~ blODDDx + blADHDsev + MEDb4  +
                a1+ a2 + a1:a2,
                id=id,
                weights = KnownWeight,
                data=   DataForAnalysis,
                family=  gaussian,
                corstr =  "independence")
```

baseline covariates

Measures the differences between the DTRs

Accounts for replicated data

Incorporates weights

Accounts for type of data and robust errors

Estimating means and contrasts

Simple linear combinations with beta coefficients: (design dependent equation)

$$E_{(a_1,a_2)}(Y|X) = \eta X + \beta_0 + \beta_1 a_1 + \beta_2 a_2 + \beta_3 a_1 a_2$$

With the above formula, and coding of $a_1 = \{1, -1\}$ and $a_2 = \{1, -1\}$, and zero-centered covariates, the following linear combinations can be used:

Estimated mean for

- DTR {1,-1} = $\beta_0 + \beta_1 - \beta_2 - \beta_3$
- DTR {-1,1} = $\beta_0 - \beta_1 + \beta_2 - \beta_3$
- DTR {-1,-1} = $\beta_0 - \beta_1 - \beta_2 + \beta_3$

Contrasts are found by subtracting. So DTR{1,1} – DTR{1,-1} is $2\beta_2 + 2\beta_3$

Analysis: Results

Table 2
Results (Parameter Estimates) for Model 1

Parameter	Estimate	Robust *SE*	95% confidence limit LL	95% confidence limit UL	Z	Pr > |Z|
Intercept	3.43	0.23	2.97	3.89	14.63	<.0001
Baseline: ODD diagnosis	0.37	0.18	0.02	0.72	2.07	.0384
Baseline: ADHD symptoms	0.57	0.14	0.29	0.85	3.95	<.0001
Baseline: Medication before Stage 1	−0.61	0.25	−1.10	−0.13	−2.47	.0134
A1	0.07	0.09	−0.11	0.24	0.75	.4555
A2	0.02	0.08	−0.13	0.18	0.26	.7924
A1 * A2	−0.12	0.08	−0.27	0.04	−1.46	.1436

Table 4
Estimated Differences Between the Four Adaptive Interventions Based on the Estimated Regression Coefficients in Table 2

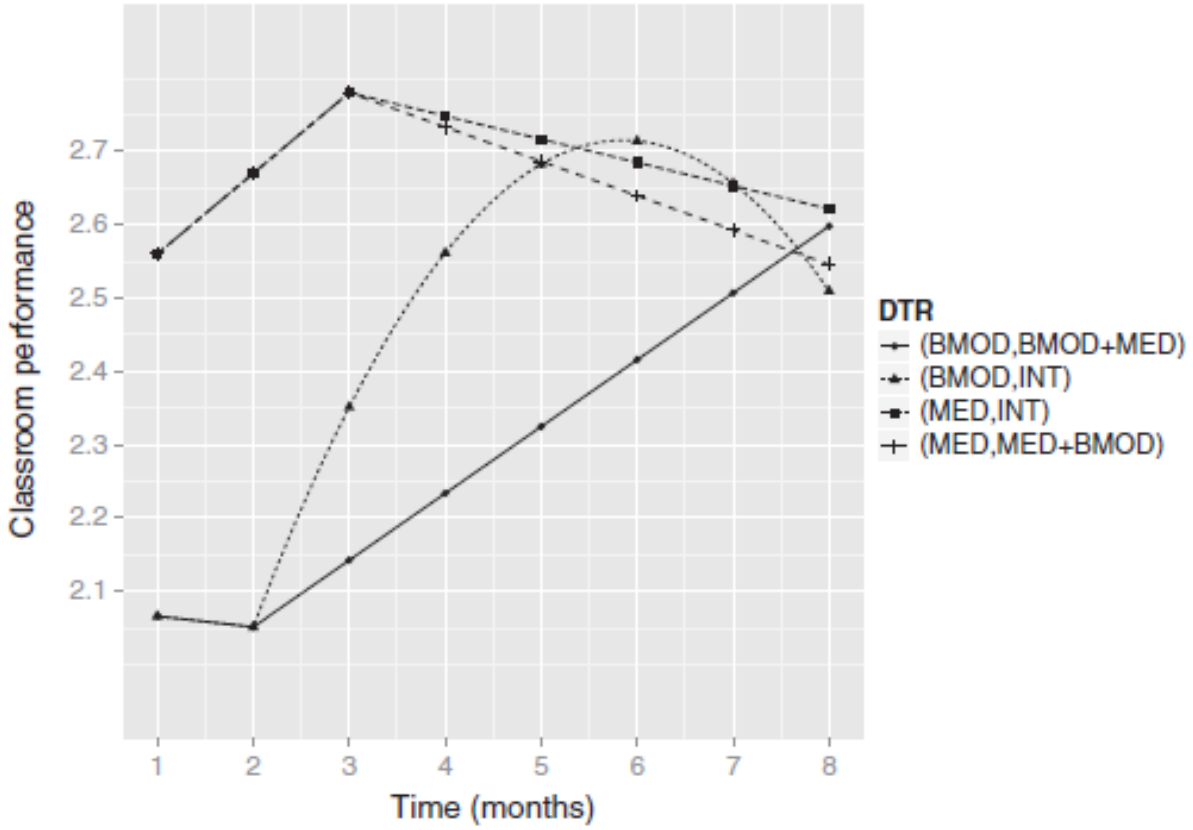
Label	Estimate	95% confidence limit LL	95% confidence limit UL	Robust *SE*	χ^2	Significance
Difference between (1, −1) and (−1, −1)	0.36	−0.06	0.79	0.22	2.82	.0932
Difference between (1, 1) and (−1, 1)	−0.10	−0.60	0.40	0.26	0.16	.6920
Difference between (1, −1) and (1, 1)	0.19	−0.32	0.70	0.26	0.55	.4600
Difference between (1, −1) and (−1, 1)	0.09	−0.31	0.49	0.20	0.20	.6563
Difference between (−1, −1) and (1, 1)	−0.17	−0.70	0.35	0.27	0.42	.5161
Difference between (−1, −1) and (−1, 1)	0.27	−0.08	0.63	0.18	2.26	.1328

Note. LL = lower limit; UL = upper limit.

Tertiary Aim: Q-learning

- Find more tailored DTRs (similar to subgroup analysis)
- Backward regression
- Example:
  - If medication was not used in the prior year, then begin with medication; otherwise select either medication or behavioral therapy. If the child is non-responsive and was non-adherent then add on to present treatment; else if the child is non-responsive and was adherent, then select either intensify or add on to current treatment.
  - https://caps.ucsf.edu/wordpress/wp-content/uploads/2011/02/CAPS.10.24.13.pdf
- Software: http://methodology.psu.edu/downloads or R packages: DynTxRegime, iqlearn, qLearn

Analysis: Longitudinal



Method from Lu et al, SIM 2016
DOI: 10.1002/sim.6819

Figure 5. Estimated mean trajectories under the embedded dynamic treatment regime of the ADHD SMART.

Summary

- **Dynamic treatment regimens** are guidelines for clinical practice
- A **SMART** is a clinical trial design that can build better and compare DTRs
- The sample size of a SMART is highly dependent on the primary aim
- R packages and applets are available to help in the design and analysis plan for a SMART, but often simulation is required

Resources

- Lei H, Nahum-Shani I, Lynch K, Oslin D, Murphy SA. *A "SMART" design for building individualized treatment sequences.* The Annual Review of Clinical Psychology, 2012. 8:21-48.
- Almirall, D., Nahum-Shani, I., Sherwood, N.E., Murphy, S.A. *Introduction to SMART designs for the development of adaptive interventions: with application to weight loss research*. Translational behavioral Medicine, 2014. 4(3):260-274.
- Chakraborty, B. and Murphy, S. A. *Dynamic treatment regimes*. Annual Review of Statistics and its Applications. 2014. 1:447-464.
- Nahum-Shani I, et al. *Experimental design and primary data analysis methods for comparing adaptive interventions*. Psychological Methods. 2012. 17:457-477.
- Kidwell, K.M. *SMART designs in cancer research: past, present and future*. Clinical Trials. 2014. 11(4): 445-456.
- Lavori, P.W., Dawson, R. *Introduction to Dynamic Treatment Strategies and sequential multiple assignment randomization*. Clinical Trials. 2014. 11(4): 393-399.

Textbooks

- **Adaptive Treatment Strategies in Practice: Planning Trials and Analyzing Data for Personalized Medicine**. Ed. Kosorok & Moodie. 2016. ASA-SIAM.
- **Statistical Methods for Dynamic Treatment Regimes: Reinforcement Learning, Causal Inference, and Personalized Medicine**. Chakraborty and Moodie. 2013. Springer.

Dynamic Treatment Regimens (DTRs)

- Sequences of treatments are relevant:
- Waxing and waning of disease/disorder
- No widely effective treatment
- Treatments may be costly or burdensome
- Adherence problems
- Within and between person heterogeneity