

Getting SMART about Developing Adaptive Interventions: Individualizing Sequences of Treatment

Society for Behavioral Medicine, Philadelphia, PA – April 23, 2014

12NOON - 6PM, Room Franklin 10, 4th Floor, Downtown Marriot

Instructors: Daniel Almirall and Billie Nahum-Shani (University of Michigan)

Session	Time	Topic
INTRO	12:00-12:15PM	15 Minute Introductions, Course Outline & Structure
MODULE 1 (BILLIE)	12:15-01:05PM (50 min)	Introduction to Adaptive Interventions <ul style="list-style-type: none"> • What are adaptive interventions (AI)? • What are the components of an AI? • Compare simple versus deeply-tailored AIs. • Why are AIs needed? • The use of theory in designing an AI • How AIs can be used to inform clinical practice?
BREAK	01:05-01:15PM	10 Minute Break (Bathroom, Coffee)
PRACTICUM	01:15-02:05PM (50 min)	Practice Exercises with Discussion & Feedback <ul style="list-style-type: none"> • Develop some example adaptive interventions in your area of study • List 4 critical questions that need to be answered to develop a high-quality AI
MODULE 2 (DANNY)	02:05-02:55PM (50 min)	Sequential Multiple Assignment Randomized Trials (SMARTs) <ul style="list-style-type: none"> • What are SMARTs? Why do we need SMARTs? • Utilizing theory to plan a SMART • SMARTs versus using a multiple-RCT approach to develop an AI • Discuss SMART design principles • What are typical primary and secondary aims in a SMART? • Sample size considerations in designing a SMART
BREAK	02:55-03:05PM	10 Minute Break (Bathroom, Coffee)
MODULE 3 (BILLIE)	03:05-03:55PM (50 min)	Case Studies: Six SMART Studies Completed or in the Field <ul style="list-style-type: none"> • Present two example SMARTs in autism, child ADHD, women who are pregnant and abuse substances, adult alcohol use, depression, implementation science • For each SMART: Present and discuss the rationale for the SMART, primary and secondary scientific questions, range of treatments/components examined, type of SMART (e.g., restricted, balanced), sample size
BREAK	03:55-04:05PM	10 Minute Break (Bathroom, Coffee)
PRACTICUM	04:05-04:55PM (50 min)	Practice Exercises with Discussion & Feedback <ul style="list-style-type: none"> • Review questions from Practicum 1 and sketch a SMART design to address them • Outline/sketch the argument/rationale for your SMART design
MODULE 4 (DANNY)	04:55-05:45PM (50 min)	Primary Aims Data Analysis: Comparing Adaptive Interventions using SMART <ul style="list-style-type: none"> • Review what it means to compare adaptive interventions in a SMART • Present an easy-to-use weighted-and-replicated regression method • Illustrate the method by reviewing two published studies that utilized weighted-and-replicated regression to compare AIs in child ADHD and in autism
BUFFER	05:45-06:00PM	15 Minute Buffer (to ensure we end on time)

Adaptive Interventions

Module 1

Getting SMART about Developing Adaptive Interventions: Individualizing Sequences of Treatment

SBM 2014; Philadelphia, PA

Daniel Almirall, Billie Nahum-Shani, Susan A. Murphy



50 minutes

Introduction to Adaptive Interventions

What are adaptive interventions (AI)?

What are the pieces that make up an AI?

Examples of AIs: Compare simple versus deeply-tailored AIs.

Discuss why AIs are needed

Utilizing theory to design an AI

How AIs can be used to inform clinical practice

Module 1 Learning Goals

- To understand all of the components that make up an Adaptive Intervention (AI).
- To understand settings in which Adaptive Interventions may be necessary (rationale)
- To understand the difference between a well-operationalized vs a poorly operationalized AI
- To begin to identify critical scientific questions (gaps in your area of study) related to developing a high-quality AI

We will have practice exercises (leading to interesting discussions) at the end of this module aimed at helping to ensure you achieve these learning goals.

Outline

- What are Adaptive Interventions?
- Why use Adaptive Interventions?
- Adaptive Intervention Design Goals
- What does an Adaptive Intervention include?
- Summary & Discussion

Other names for adaptive interventions: adaptive health interventions, adaptive treatment strategies, dynamic treatment regimes, treatment algorithms, stepped care models, treatment protocols, individualized interventions.

Related Ideas: Stepped care models are a special case of adaptive interventions. Adaptive interventions are used in decision support systems where they are often called expert systems. Structured treatment interruptions in the treatment of HIV/AIDS are an old form of adaptive intervention. To our knowledge, structured treatment interruptions are not longer in use in the treatment of HIV/AIDS. Just-in-time-adaptive-interventions (or JITAIs) are different types of adaptive interventions used in mobile health (mHealth) applications, which are able to provide treatment anytime, anywhere and respond immediately to momentary information about the individual/patient. In this workshop, we focus exclusively on time-varying adaptive interventions, in which there are multiple critical decision points (in temporal order) at which an individualized treatment decision is required. In personalized medicine, investigators/clinicians focus on non-time-varying (or point-treatment) adaptive interventions. These ideas become more clear as we go on in the workshop.

Definition of an Adaptive Intervention

- An adaptive intervention is a sequence of individually-tailored decision rules that specify whether, how or when to alter the dosage (frequency or duration), type, or delivery of pedagogical, psychosocial, behavioral, or pharmacological treatments at critical decision points during intervention.
- Sounds much like actual clinical, policy or public health practice? Yes!
- Adaptive Interventions go by many different names.

Other names for adaptive interventions: adaptive health interventions, adaptive treatment strategies, dynamic treatment regimes, treatment algorithms, stepped care models, treatment protocols, individualized interventions.

Related Ideas: Stepped care models are a special case of adaptive interventions. Adaptive interventions are used in decision support systems where they are often called expert systems. Structured treatment interruptions in the treatment of HIV/AIDS are an old form of adaptive intervention. To our knowledge, structured treatment interruptions are not longer in use in the treatment of HIV/AIDS. Just-in-time-adaptive-interventions (or JITAs) are different types of adaptive interventions used in mobile health (mHealth) applications, which are able to provide treatment anytime, anywhere and respond immediately to momentary information about the individual/patient. In this workshop, we focus exclusively on time-varying adaptive interventions, in which there are multiple critical decision points (in temporal order) at which an individualized treatment decision is required. In personalized medicine, investigators/clinicians focus on non-time-varying (or point-treatment) adaptive interventions. These ideas become more clear as we go on in the workshop.

Components of an Adaptive Intervention

- The components of an adaptive intervention are
 - a sequence of critical treatment decision points, where at each decision point, there is
 - a set of possible treatment options
 - a set of tailoring variables
 - a decision rule, which inputs the value of the tailoring variables and outputs an individualized treatment recommendation(s).
- The objective of an AI is to provide a guide for clinical policy or public health practice.

AIs provide a paradigm by which to improve clinical, policy, and public health practice which by its nature is often adaptive.

Individualization/personalization/tailoring is achieved by use of a decision rules at each decision point. Each decision rule takes accumulated, ongoing information about the unit (e.g., individual) including past response, adherence, burden, etc., and outputs a recommended, individualized treatment tailored to the circumstances of that unit.

A scientist first develops an AI. Later, they are used by clinicians to guide their thinking in actual clinical practice.

We use the term AI but others might use the terms: dynamic treatment regimes, treatment algorithms, stepped care models, expert systems, adaptive interventions, treatment protocols.

Example: An Adaptive Aftercare Intervention for Alcoholic Individuals

- **Population:** alcohol dependent individuals who have graduated from an intensive outpatient program
- **Overall goal:** prevent relapse to alcohol abuse
- **Critical treatment decisions:** which txt to provide first? which txt to provide second?
- **Tailoring variable:** number of heavy drinking days (measured weekly)

Example Adaptive Intervention (in words)

Upon graduation, all alcohol dependent individuals are provided Naltrexone along with Medical Management.

IF an individual experiences 3 or more heavy drinking days prior to 8 weeks

THEN the individual's Naltrexone treatment is augmented with Combine Behavioral Intervention.

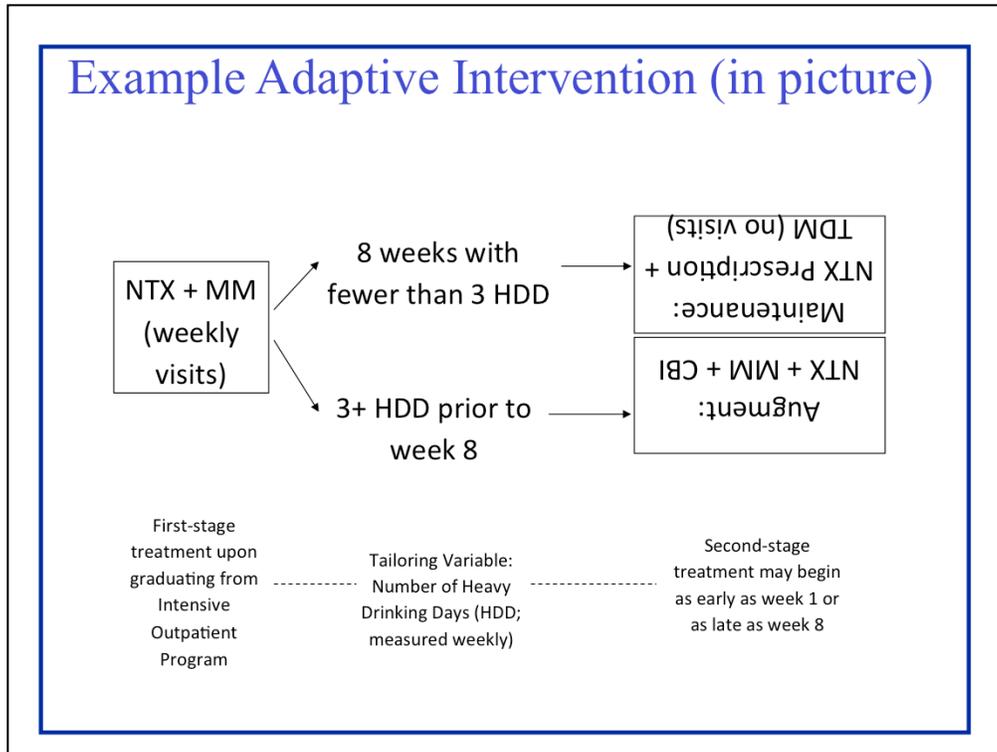
ELSE IF the individual successfully completes 8 weeks with fewer than 3 heavy drinking days

THEN the individual is provided a prescription to Naltrexone along with Telephone Disease Management.

Individuals have weekly medical management visits

naltrexone medication (opiate antagonist—reduces the reinforcing or pleasurable effects of alcohol)
+ MM is standard treatment

CBI is combine behavioral intervention this is motivational enhancement and cognitive behavioral therapy—incorporates pharmacotherapy



Individuals have weekly medical management visits

naltrexone medication (opiate antagonist—reduces the reinforcing or pleasurable effects of alcohol)
+ MM is standard treatment

CBI is combine behavioral intervention this is motivational enhancement and cognitive behavioral therapy—incorporates pharmacotherapy

Adaptive Interventions

- From the individual/patient/client's point of view: a sequence of (individualized) treatments
- From the clinician's point of view: a sequence of decision rules that recommend one or more treatments at each critical decision.

More examples of critical treatment decisions

- How long should we use the first treatment before transitioning to a maintenance/relapse prevention treatment? And which treatment should this be?
- How long should we try the first treatment before declaring non-response and moving to another treatment? And which treatment should this be?
- How should a treatment be delivered?
- How do we re-engage patients who are non-adherent?

Other critical decisions: The individual's participation in treatment (e.g., who should set health-related goals, the participant or the care provider?), the location of the intervention offered (e.g., is it better to offer treatment at home or at the clinic?), the provider of the intervention (e.g., should the parent or the teacher intervene?), the mode of delivery (e.g., is face-to-face delivery better than Internet-based delivery?), or the timing of treatment (e.g., is it better to intervene immediately or at some later point?)

More examples of tailoring variables

- Age, Severity of illness, Presence of comorbid mental or physical conditions, Quality of family support, Past failed treatments
- Adherence to present treatment, Side effects while on present treatment, Symptoms while on present treatment
- Candidate tailoring variables include moderators, mediators or short-term outcomes or even proximal measures of the ultimate outcome of interest

Other tailoring variables are genetics, family background, proteomics

Example: Adaptive Drug Court Program

- **Population:** drug abusing offenders assigned to drug court
- **Overall goal:** minimize recidivism and drug use
- **Critical treatment decisions:** which treatment to provide first?; which treatment to provide second?
- Marlowe et al. (2008, 2009, 2012)

Criminal Justice Review 2008; 33; 343 Douglas B. Marlowe, David S. Festinger, Patricia L. Arabia, Karen L. Dugosh, Kathleen M.

Benasutti, Jason R. Croft and James R. McKay

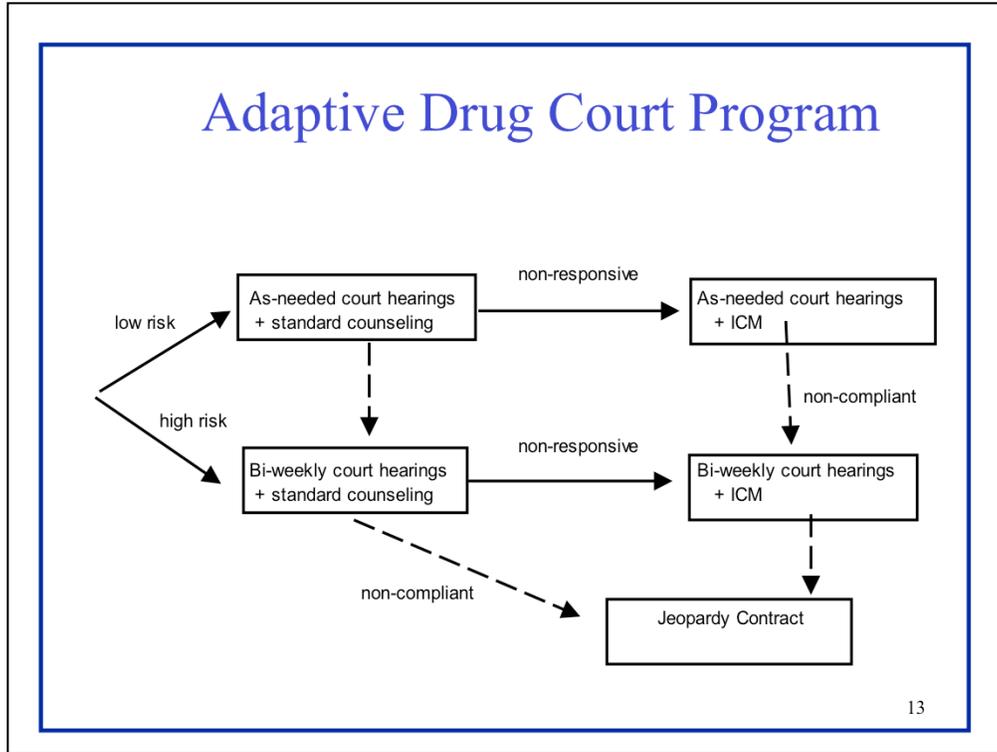
Adaptive Interventions in Drug Court: A Pilot Experiment

Adaptive interventions may optimize outcomes in drug courts: a pilot study. Marlowe DB, Festinger DS, Arabia PL, Dugosh KL, Benasutti KM, Croft JR. *Curr Psychiatry Rep.* 2009 Oct;11(5):370-6.

Adaptive Programming Improves Outcomes in Drug Court: An Experimental Trial by Douglas B. Marlowe, David S. Festinger, Karen L. Dugosh, Kathleen M. Benasutti, Gloria Fox, and Jason R. Croft *Criminal Justice and Behavior*; April 2012; vol. 39, 4: pp. 514-532.

minimize recidivism and drug use is operationalized by graduating from the drug court program

To graduate offender must attend 12 counseling sessions; provide 14 consecutive weekly negative drug urine specimens; remain arrest-free; obey program rules and procedures; pay 200 dollar court fee



All movement between steps or stages is operationalized.

High risk: ASPD or history of drug treatment otherwise low risk

These are assessed monthly::

Noncompliance: is(1) fails to attend 2 or more counseling sessions or (2) fails to provide 2 or more scheduled urine specimens

Nonresponsive = (1) is attending sessions and completing program requirements, **and** (2) is not committing new infractions, **but** (3) provides 2 or more drug-positive urine specimens.

(from Marlowe paper:) A jeopardy contract involves “zero tolerance” for further violations of the rules of the program. Any further violation leads to a termination hearing, at which the participant is terminated from the program and sentenced on the original charge or charges unless he or she can provide a good-cause reason to be given another chance. The decision whether or not to permit another chance is within the discretion of the judge and is generally granted in approximately 30% of cases

Adaptive Drug Court Program Tailoring Variables

- Stage 1 Tailoring Variables: ASPD, Prior formal drug abuse treatment
- Stage 2 Tailoring Variables: Attendance at counseling sessions, Infractions, Providing scheduled urine screens, Positive urine specimens

Adaptive Drug Court Program Decision Rules

- Stage 1 Decision Rule: Provide group-based drug abuse counseling to all. **If** ASPD or Prior formal drug abuse treatment **then** provide bi-weekly court hearings. **Else** provide as-needed court hearings.
- Stage 2 Decision Rule: **If** committed an infraction or missed 2 or more counseling sessions or missed 2 or more urine screens **then** step up court supervision. **Else if** 2 or more positive urine specimens **then** step up treatment to ICM. **Else** continue on stage 1.

ICM is intensive case management, includes individual counseling as well as help with other aspects of life (housing, etc.)

Other Examples of Adaptive Interventions

- Brooner et al. (2002, 2007) Treatment of Opioid Addiction
- McKay (2009) Treatment of Substance Use Disorders
- HIV-Causal Collaboration (2011) Treatment of HIV
- Rush et al. (2003) Treatment of Depression

Brooner uses a two component adaptive txt strategy, one component has to do with txt and the other with encouragement to adhere.

One steps up/down intensity and type of counseling sessions based on negative urines and adherence

One steps up/down behavioral contingencies based on adherence to counseling sessions.

Rules are explicit.

McKay has a book on this topic– see **Treating Substance Use Disorders With Adaptive Continuing Care (Hardcover)** by [James R. McKay](#)

When to initiate combined antiretroviral therapy to reduce mortality and AIDS-defining illness in HIV-infected persons in developed countries: an observational study.

[HIV-CAUSAL Collaboration](#), [Cain LE](#), [Logan R](#), [Robins JM](#), [Sterne JA](#), [Sabin C](#), [Bansi L](#), [Justice A](#), [Goulet J](#), [van Sighem A](#), [de Wolf F](#), [Bucher HC](#), [von Wyl V](#), [Esteve A](#), [Casabona J](#), [del Amo J](#), [Moreno S](#), [Seng R](#), [Meyer L](#), [Perez-Hoyos S](#), [Muga R](#), [Lodi S](#), [Lanoy E](#), [Costagliola D](#), [Hernan MA](#). [Ann Intern Med](#). 2011 Apr 19;154(8):509-15.

The decision rules used by Brooner et al and McKay are quite detailed, and based on explicit actions by patient, whereas in contrast the Rush et al study (Texas Medication Algorithm Project) appears to be more loosely structured; the clinician uses clinical judgment to decide if depression levels are clinically significant and thus an augmentation or switch in treatment intensity is needed. The particular secondary treatment is chosen out of a set of specified alternatives and depends on clinical judgment/patient preference.

Outline

- What are Adaptive Interventions?
- Why use Adaptive Interventions?
- Adaptive Intervention Design Goals
- What does an Adaptive Intervention include?
- Summary & Discussion

Why Adaptive Interventions?

- 1) High heterogeneity in need for or response to any one treatment

What works for one person may not work for another, thus often need a sequence of treatments just to obtain an acute response

18

This is really “why do we need to consider a sequence of treatments?”

Why Adaptive Interventions?

2) Chronic or Waxing and Waning Course

Improvement often marred by relapse

Intervals during which more intense treatment is required alternate with intervals in which less treatment is sufficient

19

Why Adaptive Interventions?

3) Treatment is burdensome

Treatment required over long time periods is burdensome

Non-adherence leads to relapse or loss of positive effect

20

Why not combine all possible efficacious therapies and provide all of these to the patient now and in the future?

- Treatment incurs side effects and substantial burden, particularly over longer time periods.
- Problems with adherence:
 - Variations of treatment or different delivery mechanisms may increase adherence
 - Excessive treatment may lead to non-adherence
- Treatment is costly (Would like to devote additional resources to patients with more severe problems)

More is not always better!

21

Why not give a universal intervention to all for a sufficiently long time??

These are all reasons why you should not provide MORE treatment than is needed. Only provide MI to people who need motivation to adhere.

That is a multi-component fixed treatment is not practical or is too costly or would not result in good adherence

A principle of adaptive tx strategies is to provide no more than needed to accomplish desired result!

Outline

- What are Adaptive Interventions?
- Why use Adaptive Interventions?
- Adaptive Intervention Design Goals
- What does an Adaptive Intervention include?
- Summary & Discussion

Adaptive Intervention Design Goals

- Maximize the strength of the adaptive intervention
 - by well chosen tailoring variables, well measured tailoring variables, well conceived decision rules & well implemented decision rules

CLARIFICATION NOTE: Here we are discussing the design of the adaptive intervention (hence “treatment design”). We are not discussing the design of a trial to inform the development of an AI—that’s the next module on “trial design”.

Use behavioral/social/biological theory, clinical experience, expert opinion, consultation with clinical staff, review of extant literature to help select the tailoring variables and form the decision rules.

Adaptive Intervention Design Goals

- Maximize replicability in future experimental and real-world implementation conditions
 - by clearly defining the treatment strategy & by fidelity of implementation
 - example: AIs require us to think carefully about having a contingency plan for non-standard scenarios that may arise

To achieve this goal, AI should be explicit. We have the most confidence in an adaptive intervention when its effects are replicable with different experimenters, different clinical staff, different locations, etc

Recall

- **Adaptive interventions** are individually tailored time-varying treatments composed of
 - a sequence of critical treatment decisions
 - tailoring variables
 - decision rules, one per critical decision; decision rules input tailoring variables and output individualized treatment recommendation(s).

Tailoring is achieved by use of a decision rules. Takes ongoing info (past response, adherence, burden, etc) and outputs txt level type

Considerations re Critical Decisions

- Need to decide which treatment decisions are critical and need to be guided (e.g. manualized, structured)? Not saying that all decisions need to be guided. Knowing which are guided and which are not is, in my view, part of manualizing.
- Consider which decisions are likely influenced by variance unrelated to the individual?
- Consider which decisions likely influenced by systematic bias (racial/age bias lead to diff txt)?

variance: different staff would provide the same individual with different treatments

Non-systematic variance: this variance is due to issues unrelated to the individual (staff member is in a hurry, staff member is tired, last patient of the day, etc.)

Systematic variance: this variance is due to (unconscious) bias on the part of the staff member. One staff member connects to the individual whereas the other staff member does not. Racial or gender or age bias lead to different treatment recommendations.

Adaptive Intervention Design Considerations

- Choice of the Tailoring Variable
- Measurement of the Tailoring Variable
- Decision Rules linking Tailoring Variables to Treatment Decisions
- Implementation of the Decision Rules

In order to understand how to achieve our design goals it is important to understand what constitutes the treatment.

Aspects of the site such as individual staff, schools, treatment sites, etc. are not part of the intervention. Rather, they are sources of extraneous variance

Measurement is particularly an issue if you have a theory based adaptive txt strategy.

This bundle (tailoring variable→ decision rule→ implementation) denotes one txt. Condition

Considerations re Tailoring Variables

- We can think of two broad types of tailoring variables
- Type 1: A variable which, at certain values, different treatments altogether must or ought to be considered.
- Type 2: A variable which is useful in pinpointing either what treatment to provide to the individual

Actually it is the optimal txt varies by individual characteristics.
To help understand this consider the following example.

Considerations re Tailoring Variables

- Type 1: A variable which, at certain values, different treatments altogether must to be considered.
 - Example: Suppose it is well established that after 2 weeks, if the patient is not responding, a different treatment must be provided (or different treatment options became unavailable once a patient exhibits an event).

Actually it is the optimal txt varies by individual characteristics.
To help understand this consider the following example.

Considerations re Tailoring Variables

- Type 2: Significant differences in effect sizes in a comparison of fixed treatments as a function of characteristics.
 - That is, some values of the tailoring variable should indicate a particular treatment decision is best while other values of the tailoring variable should indicate that a different treatment decision is best.

Actually it is the optimal txt varies by individual characteristics.
To help understand this consider the following example.

Adaptive Aftercare for Alcohol Dependent Individuals

- Hypothetical Study: Alcohol dependent individuals on NTX; after 8 weeks randomize individuals to continue on NTX or to an augment of NTX with CBI
- Result of hypothetical study: Among individuals who had returned to heavy drinking, NTX+CBI performs better than NTX only. However there is little or no difference for individuals who were maintaining a more sober lifestyle.

Adaptive Aftercare for Alcohol Dependent Individuals

- Individuals who return to heavy drinking while on Naltrexone (NTX) need additional help to maintain a non-drinking lifestyle.
- Tailoring variable is heavy drinking
- Providing CBI to individuals who are maintaining a non-heavy drinking lifestyle is costly.
- Implication: Provide NTX +CBI to individuals who are drinking heavily. NTX only is sufficient for individuals who are maintaining a non-heavy drinking lifestyle.

tailoring variable: proximal measure of heavy drinking –a proximal value of primary outcome!

This is one of those cases where a cost might be incorporated into the response, Y.

Technical Interlude!

S =tailoring variable (heavy drinking)

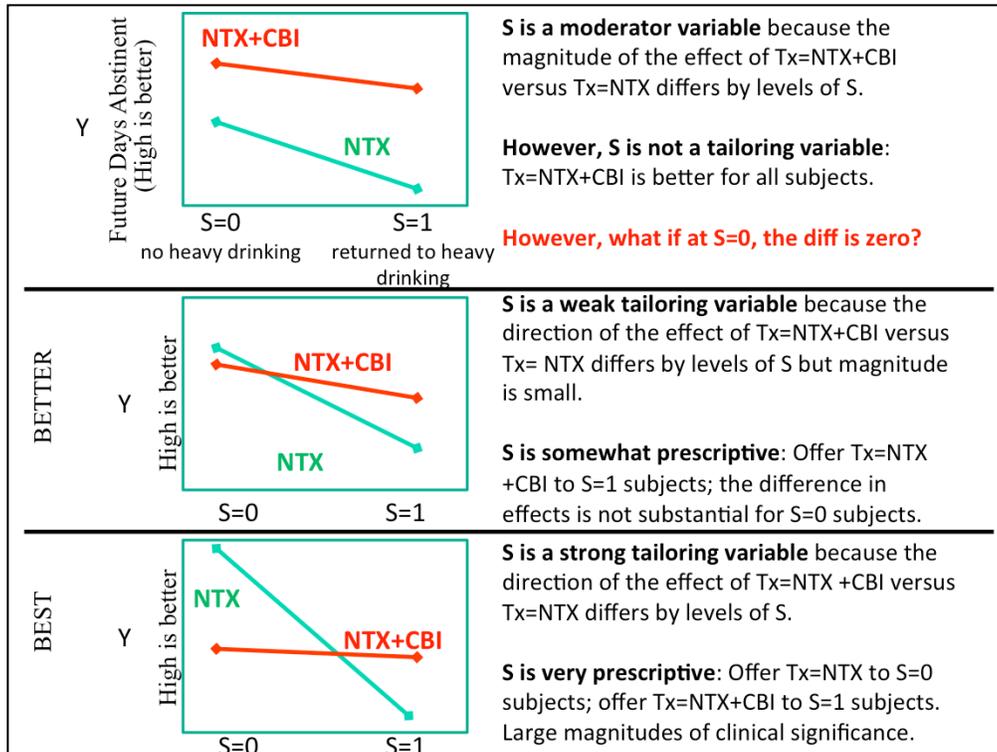
Tx =treatment type (NTX vs NTX+CBI)

Y =primary outcome (days abstinent, high is preferred)

$$Y = \beta_0 + \beta_1 S + \beta_2 Tx + \beta_3 S Tx + error$$

$$= \beta_0 + \beta_1 S + (\beta_2 + \beta_3 S)Tx + error$$

If $(\beta_2 + \beta_3 S)$ is zero or negative for some S and positive for others then S is a tailoring variable.



Tailoring variables

- Tailoring variables are moderators but they may also be
 - Baseline variables
 - Mediators
 - Short-term outcomes
 - Proximal measures of the ultimate outcome of interest.

Measurement of Tailoring Variables

- Unreliability means that you might be making unsystematic assignment of dose – getting close to random assignment, especially at the “cut point”.
- An invalid measure will weaken intervention effect (assuming your theory is correct) as you will be systematically assigning the wrong dose.

Reliability – high signal to noise ratio

Validity – unbiased.

Unreliability means that you are making unsystematic assignment of dose – getting close to random assignment.

Invalid measure will weaken intervention effect (assuming your theory is correct) as you will be systematically assigning the wrong dose.

Alcohol aftercare study included weekly self report, but biological and from collaterals is not weekly –oh no!.

Self-report: Time-Line Follow-Back (TLFB).

Biological: Carbohydrate Deficient Transferrin (CDT).

Timing of Tailoring Variable Collection

- Tailoring variable should be assessed at sufficiently frequent intervals so that non-response is detected in a timely manner.
- Too infrequent and an individual's condition may deteriorate so much that readily available rescue options are ineffective.
- Too frequent assessment may result in dependence or non-adherence

How frequently to measure a tailoring variable may be a critical decision!

Adaptive Aftercare for Alcohol Dependent Individuals

- Example:** The tailoring variable is heavy drinking days. Should we measure this variable weekly or twice a week?

Derivation of Decision Rules

- Articulate a theoretical model for how treatment effect on key outcomes should differ across values of the tailoring variable.
- Use prior clinical experience.
- Use prior experimental and observational studies.
- Discuss with research team and clinical staff, “What dosage would be best for people with this value on the tailoring variable?”

In order to achieve a particular desired treatment effect different amounts or types of treatment may be needed by different individuals

In alcohol aftercare study they know from prior studies that people who relapse to heavy drinking while on naltexone within first two months rarely recover.

Derivation of Decision Rules

- Good decision rules are objective, are operationalized.
- Strive for comprehensive rules for realistic scenarios that may arise (this is really hard but important and often not considered carefully in standard “fixed intervention” thinking!).
- Cover situations that can occur in practice, including when the tailoring variable is missing or unavailable. This is unbelievably important!

Use staff to help brainstorm about operationalizing the rules.

Operationalize the Decision Rules

- Bad:** Individuals who are drinking excessively are nonresponders and are switched to NTX +MM+CBI
- Better:** Individuals who experience 3 or more heavy drinking days are nonresponders and are switched to NTX +MM+CBI.

An even better example: As soon as 3 or more heavy drinking day occur within weeks 3-8 the person is declared a nonresponder and switched to NTX+MM +CBI

Adaptive Aftercare for Alcohol Dependent Individuals

- **Example:** Suppose an individual misses his weekly clinic visit. Then the number of heavy drinking days in the prior week is missing.
- Should we wait until the following week to decide if the individual is a non-responder or should we call the individual a non-responder immediately?

Derivation of Decision Rules

How should the decision rules include clinical judgment? How should clinical judgment be structured? (Note: We are not saying clinical judgment should not be part of the decision rules.)

- Via structured measurements that enter into tailoring variable?
- Via a choice among a restricted set of options?

Clinical judgment is used to inform the development of the decision rules and to produce structured measurements of tailoring variable

Should clinical judgment be used to select among a limited set of dosages (Although using clinical judgment to inform dosage decisions in this way may seem useful from a clinical standpoint, it is important to consider that this procedure renders clinical judgment a part of the decision rules, and therefore a part of the overall treatment.)

Is the following desirable?: Decision rules may include the less structured clinical judgment, e.g. Clinician selects one treatment in a set of recommended treatments for non-responders who are non-adhering
Clinician can declare non-response only after 6 weeks (with guidelines for what constitutes non-response)

In clinical judgment—how can local knowledge be used in a replicable way? Should local knowledge be used to choose between equivalent treatments?

Implementation in an Intervention Trial

- Try to implement decision rules universally, applying them consistently across subjects, time, site & staff members.
- We want to avoid treating some subjects differently from others due to factors that are not in the decision rules.
 - The non-systematic (eg, staff hurry) component introduces random error and increases variance.
 - The systematic component (eg, racial bias) harms replicability by increasing the plausibility of alternative explanations for the (in)effectiveness of the adaptive intervention.

If rules are not implemented universally, some persons are treated differently from others, because the dosage assignment is based in part on factors that do not figure in the decision rules and may be unique to a certain individual, time, or situation.

Implementation in an Intervention Trial

- Try to implement rules universally, applying them consistently across subjects, time, site & staff members.
- Staff may be resistant to implementing the rules universally because
 - Missing but needed tailoring variables
 - Measured tailoring variable lacks validity
 - The way the tailoring variable weighs different criteria may be questioned.
 - Decision rules are ambiguous
 - Insufficient training

Staff perceive dosage rules are inappropriate in a particular case
To the extent that individuals with the same tailoring variable values are assigned dosages by relying on ad hoc procedures rather than the established dosage assignment rules, there will be problems with replicability.

The rule is like the manual in a manualized therapy.

Implementation

- Exceptions to the rules should be made only after group discussions and with group agreement.
- If it is necessary to make an exception, document this so you can describe the implemented treatment.
- Document the value of the tailoring variable.

If it is a big deal to make an exception then staff must come up with a cogent argument that you can use to help plan future implementations.

This helps you

- 1) Future revision of rule
- 2) Indicates if there is a need for further staff training
- 3) May indicate that you need to be clearer in articulating the purpose of a txt component.

Summary & Discussion

- Adaptive interventions are attractive alternatives to fixed treatments
 - if in a comparable fixed treatment, significant variation in treatment effect would be expected as a function of identifiable tailoring variables, across participants and/or within participants over time

Summary & Discussion

Adaptive interventions enhance the potency of the treatment if

- by increasing salience and negative effects, they improve adherence
- by reducing waste it becomes possible to devote additional resources to higher-risk individuals who can benefit from them.
- by increasing the total number of individuals benefitting by taking advantage of heterogeneity

Summary & Discussion

- Research is needed to build a theoretical literature that can provide guidance:
 - in identifying tailoring variables,
 - in the development of reliable and valid indices of the tailoring variables that can be used in the course of repeated clinical assessments
 - on when/how to allow clinical judgment.

Questions?

More information

L.M Collins, S.A. Murphy and K.A. Bierman (2004), A Conceptual Framework for Adaptive Preventive Interventions, *Prevention Science* 5:185-196.

S.A. Murphy & J.R. McKay (2004), Adaptive Interventions: an Emerging Approach for Improving Treatment Effectiveness. *Clinical Science* (Newsletter of the American Psychological Association Division 12, section III: The Society for the Science of Clinical Psychology) Winter 2003/Spring 2004

L.M. Collins, S.A. Murphy, V. Nair & V. Strecher (2005), A Strategy for Optimizing and Evaluating Behavioral Interventions, *Annals of Behavioral Medicine*. 30:65-73.

S.A. Murphy, L.M. Collins, A.J. Rush (2007). Customizing Treatment to the Patient: Adaptive Interventions. *Drug and Alcohol Dependence*,. 88(2):S1-S72.

Discussion & Practice Exercise

Exercise 1: *Identify 2-3 critical decision points in your area of study.*

Exercise 2: *At each decision point, identify 2-3 treatment options (may vary by treatment type, duration, timing, modality, dose, etc.).*

Exercise 3: *At each decision point, identify 2-3 candidate tailoring variables.*

(2 more exercises on the next page)

Discussion & Practice Exercise

Exercise 4: *At each decision point, create a decision rule linking the tailoring variable to one or more of the treatment options.*

Exercise 5: *List some critical questions you think must be addressed to develop a high-quality (e.g., to improve/optimize longer-term outcomes) AI in your area?*

Sequential, Multiple Assignment, Randomized Trials

Module 2

Getting SMART about Developing Adaptive Interventions:
Individualizing Sequences of Treatment

SBM 2014; Philadelphia, PA

Daniel Almirall, Billie Nahum-Shani, Susan A. Murphy



40 minutes

Sequential Multiple Assignment Randomized Trials (SMARTs)?

What are SMARTs?

Why do we need SMARTs?

Discuss the role of critical decisions and treatment options to plan and provide the rationale for a SMART

Utilizing theory to plan a SMART

Compare SMARTs to using a multiple-RCT approach

Discuss SMART design principles

What are typical primary and secondary aims in a SMART?

Sample size considerations

De-bunk misconception that SMARTs necessarily require large sample sizes.

Module 2 Learning Goals

- To understand the prototypical structure of a SMART
- To understand the alternatives to a SMART design
- To understand some general SMART design principals

We will have practice exercises (leading to interesting discussions) at the end of this module aimed at helping to ensure you achieve these learning goals.

Some Critical Questions in Adaptive Intervention Development

- What is the best sequencing of treatments?
- What is the best timings of alterations in treatments?
- What information do we use to make these decisions?
(how do we *individualize* the sequence of treatments?)

The purpose of the SMART study is to provide high quality data for addressing these questions.

Take a broad view of what constitutes therapies: changing intensity, switching medication, augmenting medication, behavioral contingencies, monitoring schedules, motivational therapy, support networks, form of treatment delivery.

Outline

- What are Sequential Multiple Assignment Randomized Trials (SMARTs)?
- Why SMART experimental designs?
- Trial Design Principles
- Summary & Discussion

What is a SMART Study?

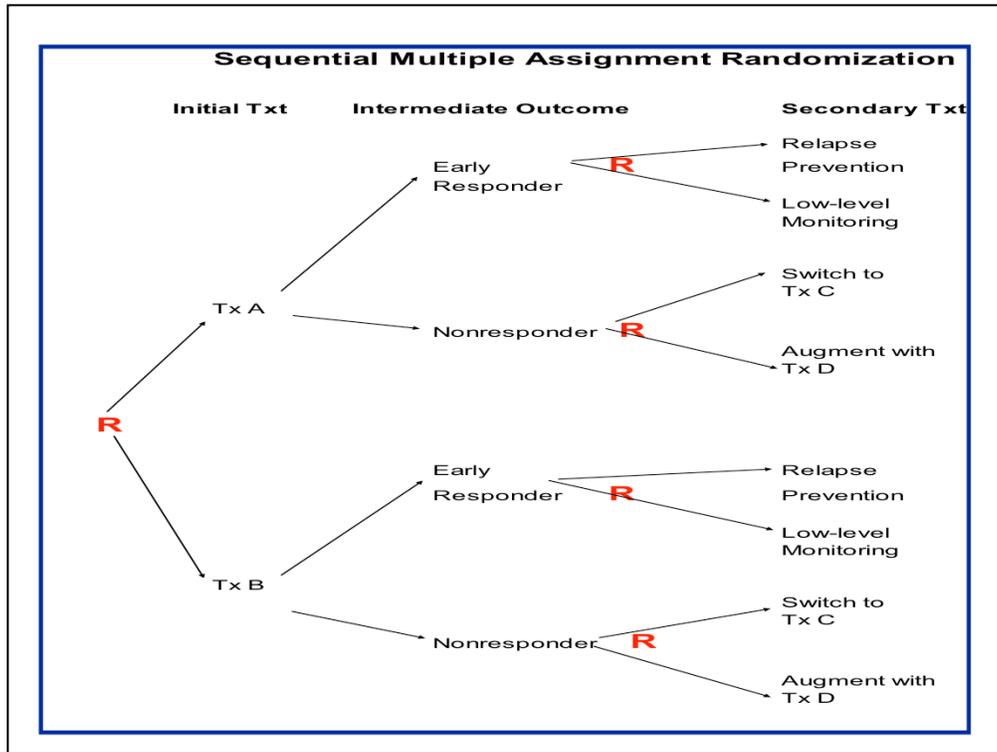
What is a sequential multiple assignment randomized trial (SMART)?

These are multi-stage trials; each stage corresponds to a critical treatment decision and a randomization takes place at each critical decision.

Goal is to inform the construction of adaptive interventions.

In stat. people may call these multistage trials (the randomization at each stage is assumed)

The randomizations at each stage allow us to learn what the best treatment is for that stage.



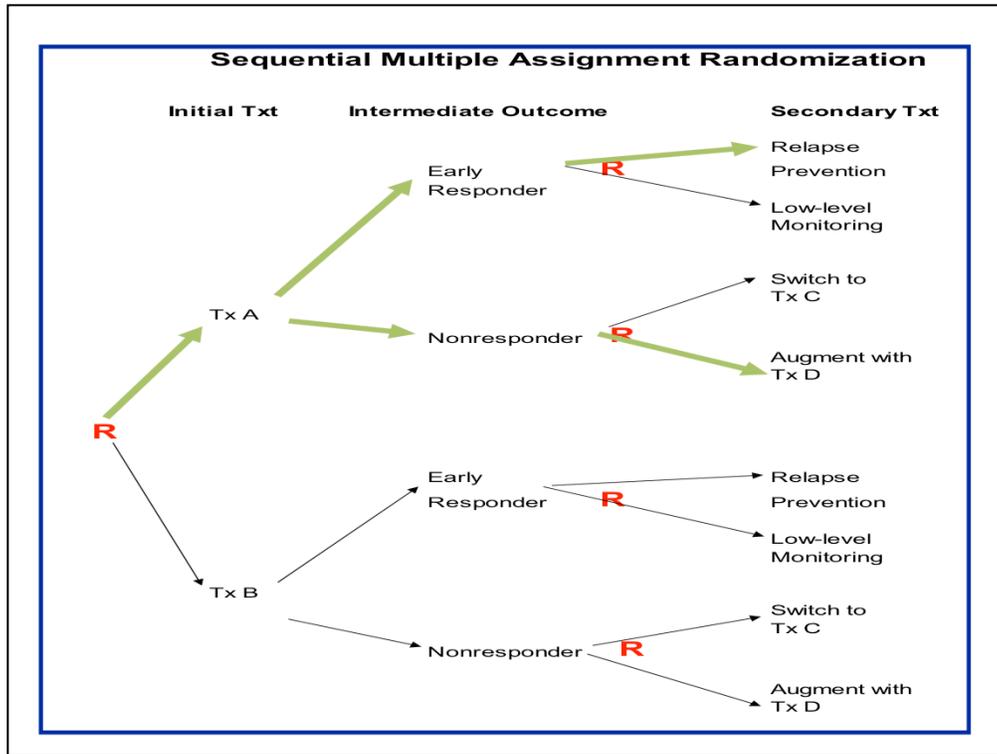
Hypothetical trial: Outcome is not shown but is on far right. The randomizations can take place up front.

Equal randomization

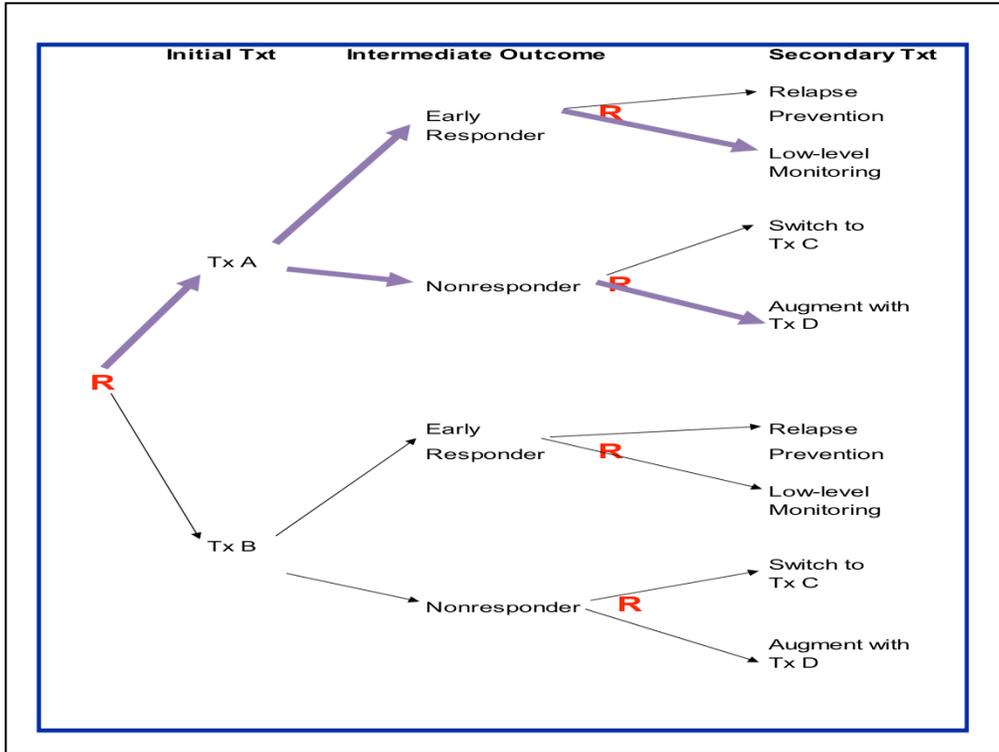
Usual reaction is (1) I'm worried about sample size and

(2) This looks awfully complicated.

In reality both of these problems are less worrisome than one might think—see following slides.



An embedded adaptive intervention



Another embedded adaptive intervention!

Outline

- What are Sequential Multiple Assignment Randomized Trials (SMARTs)?
- Why SMART experimental designs?
- Trial Design Principles
- Summary & Discussion

Challenges in constructing Adaptive Interventions

- Delayed, Prescriptive & Sample Selection Effects

---sequential multiple assignment randomized trials (SMART)

- Adaptive Interventions are Multi-component Treatments

---series of screening/refining randomized trials prior to confirmatory trial (MOST).

10

L.M. Collins, S.A. Murphy, V. Strecher (2007). The Multiphase Optimization Strategy (MOST) and the Sequential Multiple Assignment Randomized Trial (SMART): New Methods for More Potent e-Health Interventions. *American Journal of Preventive Medicine* , 32(5S):S112-118

Alternate Approach I to Constructing an Adaptive Intervention

- Why not use data from multiple trials to construct the adaptive intervention?
- Choose the best initial treatment on the basis of a randomized trial of initial treatments and choose the best secondary treatment on the basis of a randomized trial of secondary treatments.

Particularly attractive since potential initial treatment may have been evaluated in prior trials. So you propose a responder study or you propose a nonresponder study.

Or, why choosing the best initial treatment on the basis of a randomized trial of initial treatments and choosing the best secondary treatment on the basis of a randomized trial of secondary treatments is not the best way to construct an adaptive intervention

Delayed Therapeutic Effects

Why not use data from multiple trials to construct the adaptive intervention?

Positive synergies: Treatment A may not appear best initially but may have enhanced long term effectiveness when followed by a particular maintenance treatment. Treatment A may lay the foundation for an enhanced effect of particular subsequent treatments.

counseling and then if respond, monitoring with low level telephone counseling.

But this type of delayed effect may not be discovered/seen in the results of a standard randomized trial. If it is, say using follow-up data from the trial, often this is observational and strong assumptions must be made to reach the right conclusions concerning the treatment sequences in question.

Delayed Therapeutic Effects

Why not use data from multiple trials to construct the adaptive intervention?

Negative synergies: Treatment A may produce a higher proportion of responders but also result in side effects that reduce the variety of subsequent treatments for those that do not respond. Or the burden imposed by treatment A may be sufficiently high so that non-responders are less likely to adhere to subsequent treatments.

treatment of psychosis: a medication may result in many immediate responders but Some patients are not helped and/or experience abnormal movements of the voluntary muscles (TDs). The class of subsequent medications is greatly reduced.

Or the kind of response produced may not be sufficiently strong so that patients can take advantage of maintenance care.

A negative delayed effect would occur if the initial treatment overburdens an individual, resulting decreased responsiveness to future treatment; see Thall et al. (2007) for an example of the latter in cancer research.

A Consequence of Delayed Therapeutic Effects

- Comparisons of initial treatments based on a acute 3 month outcome may result in a different result from a comparison of these two initial treatments based on a 6 month outcome.
- Restricting to 6 month outcomes, a comparison of initial treatments followed by usual care in months 4-6 may differ from a comparison of initial treatments followed by one of several maintenance therapies in months 4-6.

In the second bullet, there are also issues related to observational study design since the follow-up or usual care is often not an explicit part of the study.

Harnessing Delayed Therapeutic Effects

- Our goal is to ensure that the subsequent treatment builds on gains achieved by prior treatments even when the participant initially appears non-responsive.
- We want large positive delayed effects (i.e. large positive cross-over effects)
- We want to prevent negative delayed effects.

Harnessing Delayed Therapeutic Effects

Using data from multiple trials to construct the adaptive intervention is less helpful in harnessing delayed therapeutic effects because we need to assess the combined effect of a sequence of treatments.

Prescriptive Effects

Why not use data from multiple trials to construct the adaptive intervention?

Treatment A may not produce as high a proportion of responders as treatment B but treatment A may elicit symptoms that allow you to better match the subsequent treatment to the patient and thus achieve improved response to the sequence of treatments as compared to initial treatment B.

Consider the issue of motivation as expressed via adherence; if tx A has provides less adherence support than tx B, then patients who require the adherence support will exhibit adherence problems during tx with A but not during tx with B. This is useful information as we then know that these patients, even if they respond will potentially need an enhanced adherence support during the maintenance or aftercare phase.

Sample Selection Effects

Why not use data from multiple trials to construct the adaptive intervention?

Subjects who *will enroll in*, who *remain in or* who *are adherent in* the trial of the initial treatments may be quite different from the subjects in SMART.

Consider the issue of adherence; in many historical trials subjects were assigned a fixed treatment, that is, there were no options besides non-adherence for subjects who were not improving. This often leads to higher than expected drop-out or non-adherence. This is particularly the case in longer studies where continuing treatments that are ineffective is likely associated with high non-adherence. As a result the subjects who remained in the historical trial may be quite different from the subjects that remain in a SMART trial, which by design provides alternates for non-improving subjects. David Oslin made this point to me.

A Different Example of Sample Selection Effects

True story: A scientist who has conducted non-responder trials comparing treatment A versus B decides to conduct a SMART. The scientist reports that when conducting the SMART he discovers that a large fraction of the non-responders do not want to be randomized to either treatment A or B.

What has happened? Great learning experience!

Consider the issue of motivation. Nonresponder trials recruit individuals who are not responding to their present treatment, say Med A. An important consideration is whether these nonresponders represent the population of individuals who do not respond to Med A or whether the nonresponders recruited into the trial are more motivated. Such selection bias will prevent us from realizing that we might need a behavioral intervention to encourage nonresponders to start again with treatment.

Summary (two key points RE alternative approach I):

- When evaluating and comparing initial treatments, *in a sequence of treatments*, we need to take into account the effects of the secondary treatments. SMARTs are useful for this.
- Standard one-stage randomized trials may yield information about different populations from SMART trials. In a SMART, the same group of participants participate in all stages of treatment.

Just because an initial txt looks best when looking at intermediate outcomes does not mean that it is best initially in an adaptive txt strategy

Alternate Approach II to Constructing an Adaptive Intervention

- Theory, clinical experience and expert opinion are **critical** in the development of adaptive interventions
- However, why not use theory, clinical experience and expert opinion to do the following?
 - 1) **completely construct** the adaptive intervention
 - 2) **evaluate** this adaptive intervention against an appropriate alternative in a confirmatory randomized two group trial?

21

Why constructing an adaptive intervention and then evaluating the strategy against a standard alternative is not always the answer.

- We do not know why your adaptive intervention worked or did not work. Did not open black box.
- We do not know what components of the adaptive intervention are (in)active. Is the first stage treatment or the second treatment or the tactical decisions regarding the criterion for nonresponse or the timing of assessment of nonresponse sequence effective?

22

Meeting the Challenges

Delayed/Prescriptive/Sample Selection Effects:
SMART

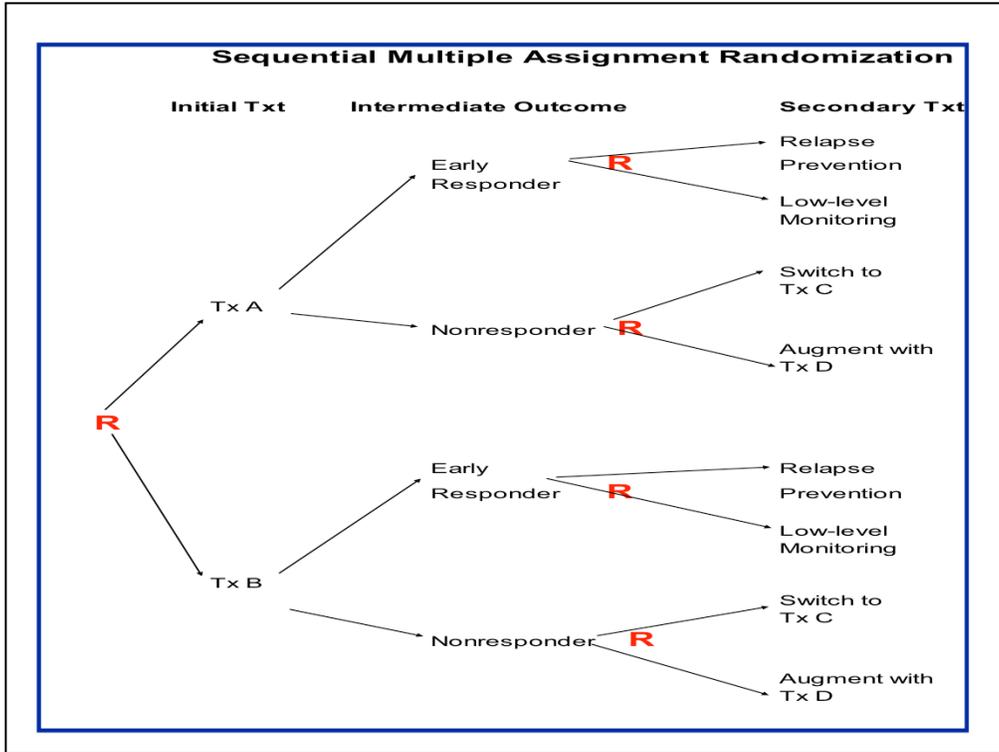
Developing Multi-Component Interventions:
Screening/refining randomized trials prior to a
confirmatory trial (MOST). Linda Collins' work.

The SMART design is one of the screening/
refining randomized trials in MOST

23

confirmatory trial is to compare the developed adaptive intervention versus an appropriate alternative—this is the standard randomized two group trial.

MOST multistage optimization strategy



Examples of SMART designs:

- Kasari (completed) Alternative Augmentative Communication Devices for Children with Autism
- Pelham (analyzing data) Treatment of ADHD
- Oslin (analyzing data) Treatment of Alcohol Dependence
- McKay (2 SMARTs; primary analysis) Treatment of Alcohol and Cocaine Dependence
- Jones (in field) Reinforcement Behavioral Therapy for Pregnant Women Who are Drug Dependent
- Kasari (in the field) Improving Spoken Communication in Autism

After lunch we will discuss some of these designs in some detail!

Outline

- What are Sequential Multiple Assignment Randomized Trials (SMARTs)?
- Why SMART experimental designs?
- Trial Design Principles
- Summary & Discussion

Critical Decisions

- Choose two or three critical decisions to address.
- Examples of critical decisions
 - Sequencing decisions: Which treatment to try first? Which treatment to try if individual shows signs of nonresponse? Which treatment to try if the individual is doing well?
 - Timing decisions: How soon do we declare nonresponse? How soon do we declare response?
- Which decisions are most controversial or need investigation? Which decisions are likely to have the biggest impact on the outcome?

In the use of naltrexone for alcohol dependence different researchers and clinicians use different criteria for non-response ranging from at least 5 heavy drinking days to at least 2 heavy drinking days. Yet 8 weeks of little to no heavy drinking is a common criterion for response.

So one of the critical decisions to investigate was the heavy drinking days trigger for nonresponse. We decided that it was less important to investigate the best duration of little to no heavy drinking before declaring response.

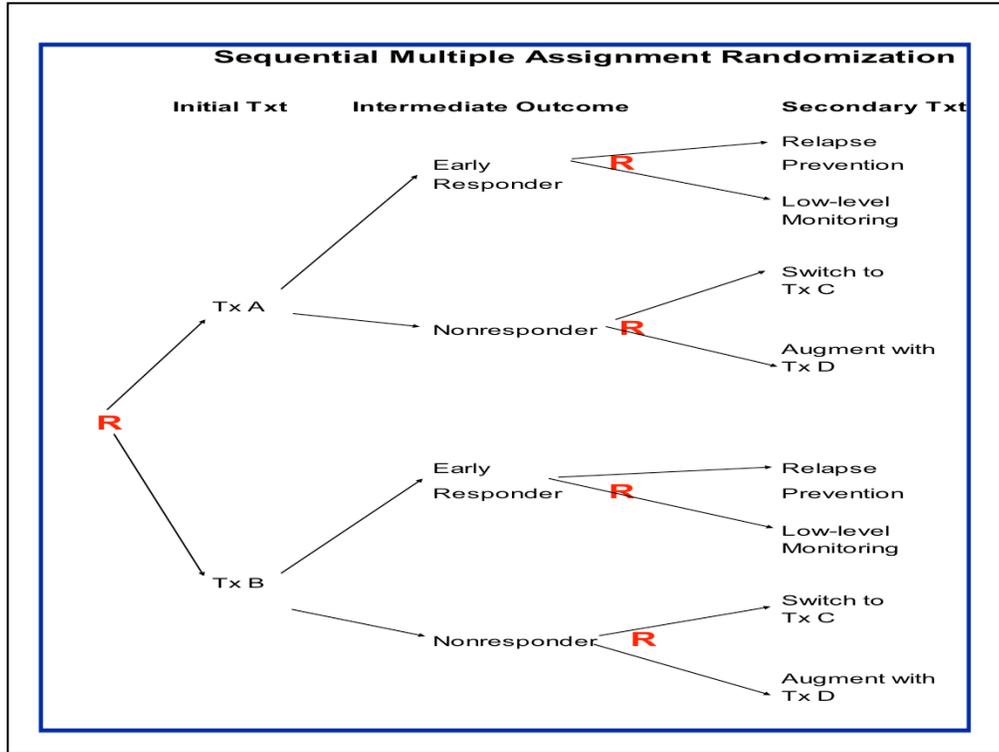
Critical Decisions

- In planning the study of Naltrexone for alcohol dependence, we realized that different researchers and clinicians use different criteria for non-response ranging from at least 5 heavy drinking days to at least 2 heavy drinking days.
 - This timing decision became one of the critical decisions to investigate.
- Other critical decisions involved which maintenance treatment to provide responders and which treatment to provide nonresponders.

See H. Lei, I. Nahum-Shani, K. Lynch, D. Oslin and [S.A. Murphy](#)
[A SMART Design for Building Individualized Treatment Sequences.](#)
The Annual Review of Clinical Psychology (2012), Review in Advance
first posted online on December 12, 2011 for greater detail.

SMART Treatment Stages

- Each treatment stage (i.e., phase) in the SMART corresponds to a critical decision.
- We randomize participants at each treatment stage among different treatment options.
- The first stage of the alcohol dependence study involved randomization to either a “ ≥ 5 HDD nonresponse definition” or a “ ≥ 2 HDD nonresponse definition.”



What are the critical decisions in this hypothetical trial? What are the stages?

SMART Design Principles

- **KEEP IT SIMPLE:** At each stage (critical decision point), restrict class of treatments only by ethical, feasibility or strong scientific considerations. Use a low dimension summary (responder status) instead of all intermediate outcomes (adherence, etc.) to restrict class of next treatments.
- Collect intermediate outcomes that might be useful in ascertaining for whom each treatment works best; information that might enter into the adaptive intervention.

Note we considered different txt' s for the responders as compared to the nonresponders. A SMART does not need to restrict the class of treatments by responder status.

Collect information on adherence, symptoms, side effects, problems with co-occurring disorders, etc.

SMART Design Principles

- Choose primary hypotheses that are both scientifically important and aid in developing the adaptive intervention.
 - Power trial to address these hypotheses.
- Choose secondary hypotheses that further develop the adaptive intervention and use the randomization to eliminate confounding.
 - Trial is not necessarily powered to address these hypotheses and doesn't have to be.

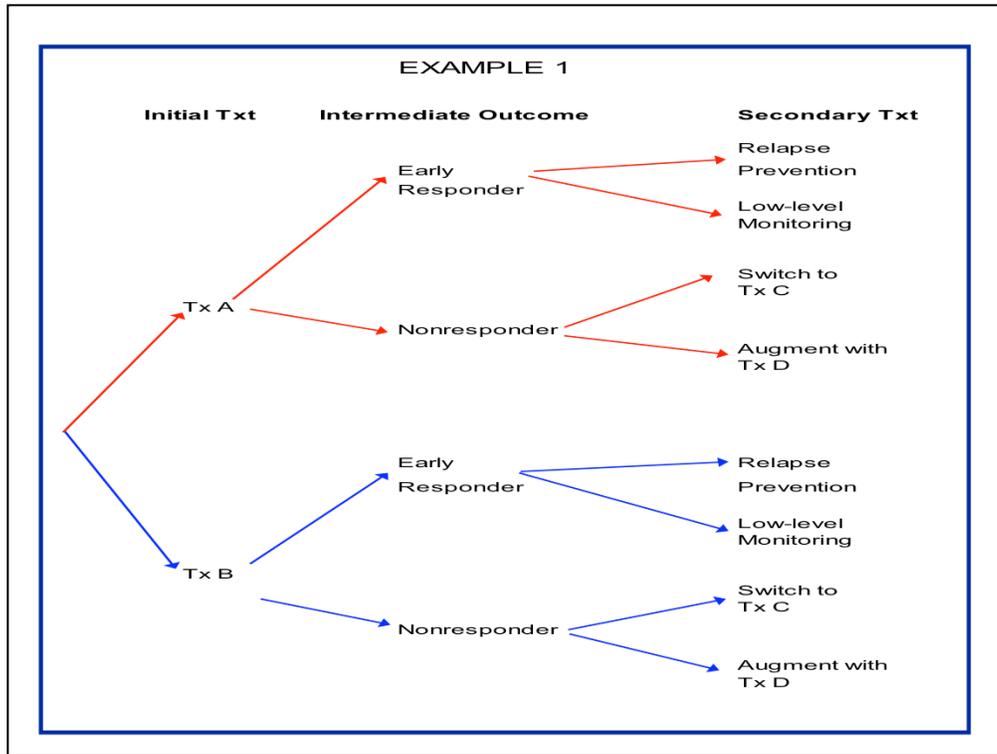
SMART Designing Principles: Primary Hypothesis

- EXAMPLE 1: (*sample size is highly constrained*):
Hypothesize that controlling for the secondary treatments, the initial treatment A results in lower symptoms than the initial treatment B.
- EXAMPLE 2: (*sample size is less constrained*):
Hypothesize that among non-responders a switch to treatment C results in lower symptoms than an augment with treatment D.

These are main effects a la' ANOVA

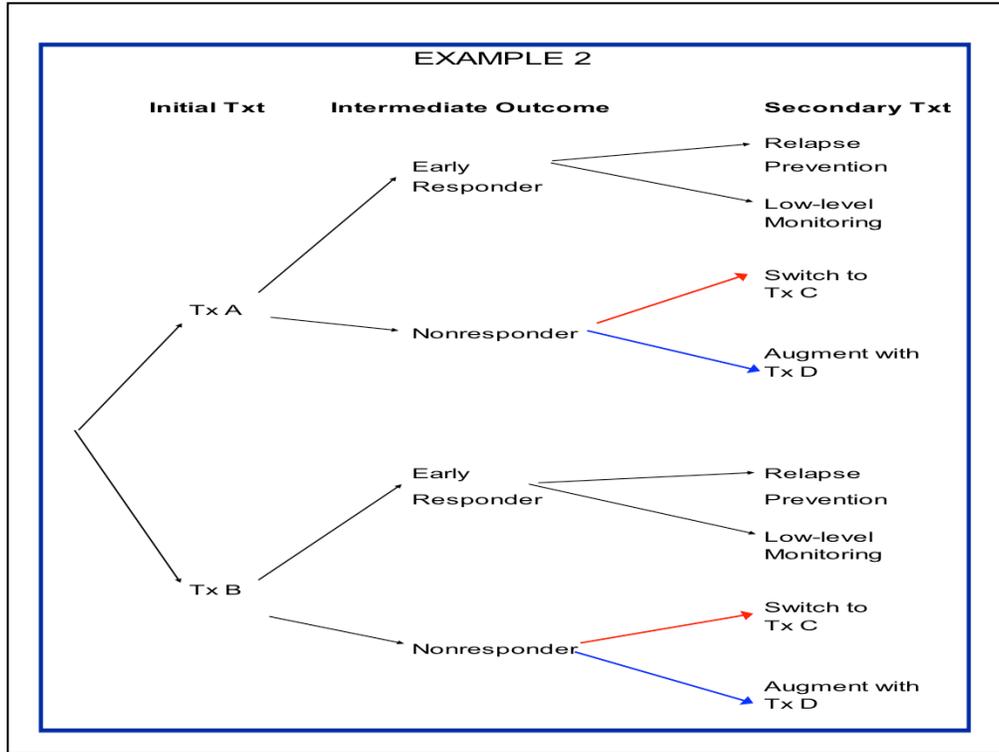
The second would be appropriate if you initially wanted to run a trial for non-responders and are now considering SMART

Example 1: Effects of secondary treatments are controlled by experimental design –not by statistical analysis



A study of initial tx' s in which subsequent tx' s are controlled.

Here you can use a variety of analyses, growth curve models, survival analysis, etc.

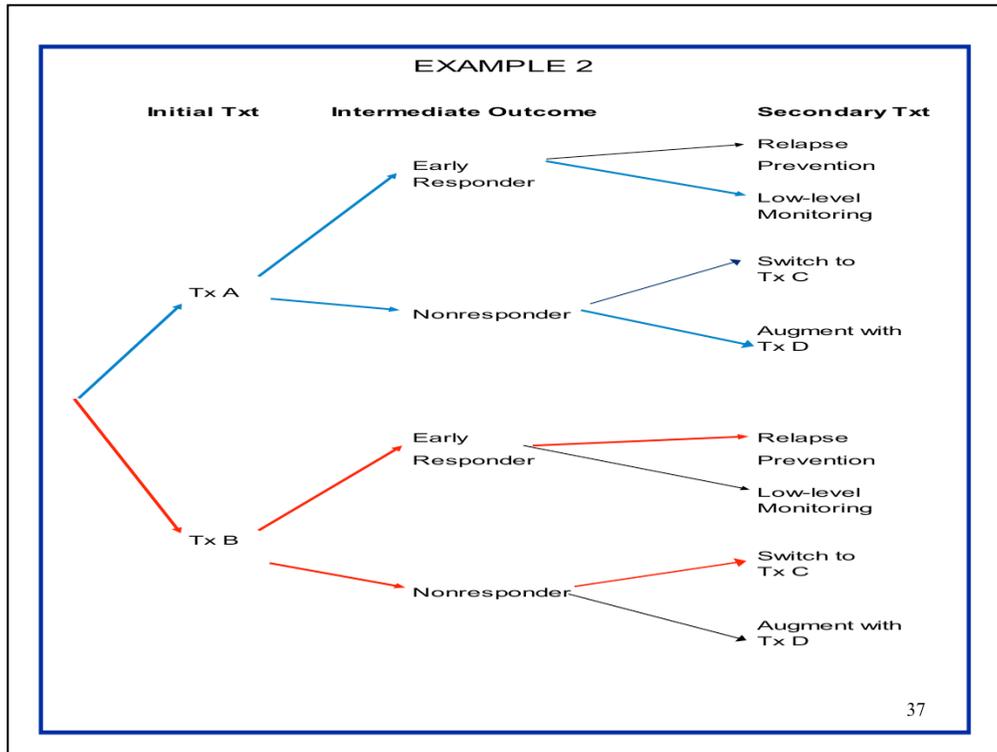


A study of nonresponders in which one controls the tx's to which people don't respond to.

SMART Designing Principles: Primary Hypothesis

- EXAMPLE 3: (*sample size is less constrained*):
Hypothesize that embedded adaptive intervention 1 (in blue) results in improved symptoms as compared to embedded adaptive intervention 2 (in red)

These are main effects a la' ANOVA



Sample size formula for this SMART to compare the red versus blue embedded adaptive interventions is given in [S.A. Murphy \(2005\), An Experimental Design for the Development of Adaptive Interventions., Statistics in Medicine. 24:1455-1481](#)

Requires a weighted analysis Murphy et al (2001)

SMART Designing Principles: Sample Size Formula

- EXAMPLE 1: (lower sample size needed): Hypothesize that given the secondary treatments provided, the initial treatment A results in lower symptoms than the initial treatment B. *Sample size formula is same as for a two group comparison.*
- EXAMPLE 2: (larger sample size needed): Hypothesize that among non-responders a switch to treatment C results in lower symptoms than an augment with treatment D. *Sample size formula is same as a two group comparison of non-responders.*

These are main effects a la' ANOVA

Example Sample Sizes

N=trial size

	Example 1	Example 2
$\Delta\mu/\sigma = .3$	N = 402	N = 402/initial nonresponse rate
$\Delta\mu/\sigma = .5$	N = 146	N = 146/initial nonresponse rate

$\alpha = .05,$ power = $1 - \beta = .85$

Sigma for example 1 is the std of primary outcome of patients initially assigned tx A (or B)

Sigma for example 2 is the std of primary outcome of non-responding patients who are assigned a switch (or augment)

Throughout working assumptions are equal variances and normality

Sample sizes calculated on the website: http://hedwig.mgh.harvard.edu/sample_size/quant_measur/para_quant.html

In the case of example 3, multiply N by 2. Sigma for example 3 is the std of the primary outcome of patients assigned the blue adaptive intervention (or red adaptive intervention).

An analysis that is less useful in the development of adaptive interventions:

Decide whether treatment A is better than treatment B by comparing proportion of early responders.

This is interesting and certainly helps with the story (e.g., to explain a delayed effect) but often not a primary or secondary aim in a SMART.

40

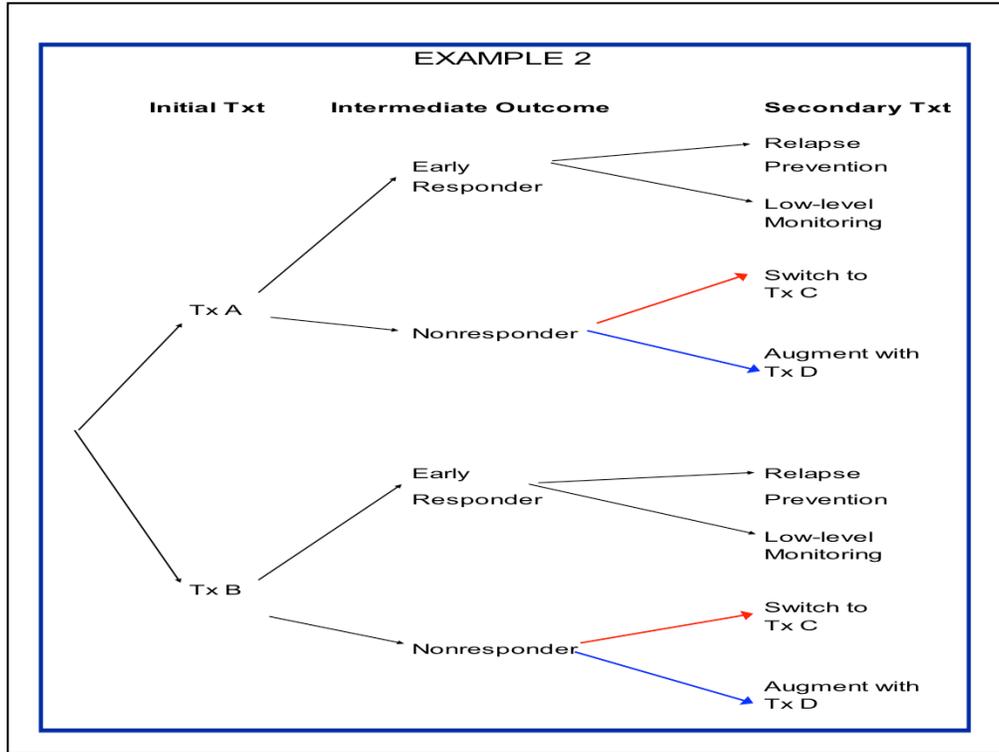
It is interesting but not as useful in the development of adaptive interventions

SMART Designing Principles

- Choose secondary hypotheses that further develop the adaptive intervention and use the randomization to eliminate confounding.
- EXAMPLE: Hypothesize that *non-adhering* non-responders will exhibit lower symptoms if their treatment is augmented with D as compared to an switch to treatment C (e.g. augment D includes motivational interviewing).

Confounding::: alternative explanations other than txt effect for the observed comparisons

Use analysis of covariance or regression.



Just use nonresponders' data. For example with a continuous outcome we might use a regression that includes an interaction term between second stage treatment and adherence.

Summary & Discussion

- We have a sample size formula that specifies the sample size necessary to detect an embedded adaptive intervention that results in a mean outcome δ standard deviations better than the other embedded adaptive interventions with 90% probability.
- We also have sample size formula that specify the sample size for time-to-event studies.

See

<http://methodology.psu.edu>

43

Questions?

More information

S.A. Murphy, K.G. Lynch, J.R. McKay, D. Oslin, T. TenHave (2007). Developing Adaptive Interventions in Substance Abuse Research. *Drug and Alcohol Dependence*, 88(2):S24-S30

L.M. Collins, S.A. Murphy, V. Strecher (2007). The Multiphase Optimization Strategy (MOST) and the Sequential Multiple Assignment Randomized Trial (SMART): New Methods for More Potent e-Health Interventions. *American Journal of Preventive Medicine*, 32(5S):S112-118

A.I. Oetting, J.A. Levy, R.D. Weiss, S.A. Murphy(2011), Statistical Methodology for a SMART Design in the Development of Adaptive Interventions,, *Causality and Psychopathology: Finding the Determinants of Disorders and their Cures*, (P.E. Shrout, K.M. Keyes, K. Ornstein, Eds.) Arlington VA: American Psychiatric Publishing, Inc, pgs. 179-205

I. Nahum-Shani, M. Qian, D. Almirall, W.. Pelham, B. Gnagy, G. Fabiano, J. Waxmonsky, J. Yu and S.A. Murphy (2012). Experimental Design and Primary Data Analysis Methods for Comparing Adaptive Interventions. *To appear in Psychological Methods*

Very technical:

S.A. Murphy and D. Bingham (2009). Screening Experiments for Developing Dynamic Treatment Regimes. *JASA*. 184:391-408.

Practice Exercise

Exercise 1: Re-vise/Re-visit the list of critical questions you made at the end of Module 1. This is the list of critical questions that need to be addressed to develop a high-quality AI.

Exercise 2: Begin drafting a picture of a SMART design in your area to address these questions. You will have time to work on this again after Module 3.

SMART Case Studies

Module 3

Getting SMART about Developing Adaptive Interventions: Individualizing Sequences of Treatment

SBM 2014; Philadelphia, PA

Daniel Almirall, Billie Nahum-Shani, Susan A. Murphy



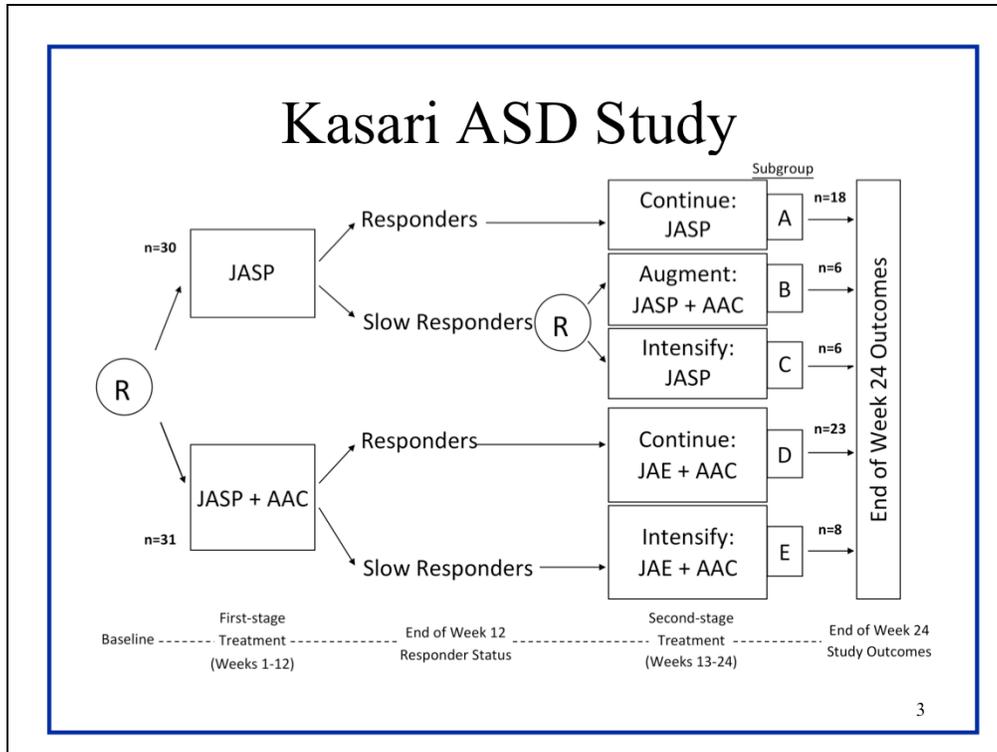
40 minutes

- Give examples of SMARTs that are completed or in the field
 - o ASD, child ADHD, women who are pregnant and abuse substances, adult alcohol use, depression
- Discuss the variety of rationales underlying the SMARTs, types of critical decisions; range of treatment modalities, differences in primary aims
- Compare balanced versus unbalanced SMART designs

Outline

- Adaptive ASD Developmental and Augmented Intervention (Kasari, PI)
- Adaptive Pharmacological and Behavioral Treatments for Children with ADHD Trial (Pelham, PI)
- Adaptive Reinforcement-Based Treatment for Pregnant Drug Abusers (Jones, PI)
- Extending Treatment Effectiveness of Naltrexone (Oslin, PI)
- Comparison of SMARTs

ASD=autism spectrum disorder



This trial was just wrapped up.

See

<http://www.semel.ucla.edu/ASD/research/project/ccnia-developmental-augmented-intervention-facilitating-expressive-language>

CCNIA=characterizing cognition in nonverbal individuals with ASD

N=96 6 month trial

But difficulties in recruiting autistic children made it difficult to meet this goal. They only achieved N=61.

JASP in the picture is actually short for JASP+EMT = Joint attention social play + enhanced mileau training. See next slides.

Kasari ASD SMART

- Population & Rationale:
 - Non-verbal children with ASD who have not made satisfactory progress by age 5 even though they have received traditional intensive interventions
 - These children experience poor outcomes yet represent 25-30% of children with ASD.
 - Planning for a “rescue” if the first treatment does not go well is crucial.

ASD: ASD spectrum disorder

6 month study

Kasari ASD SMART

- Critical Decisions:
 - Which treatment to provide first? Which treatment to provide non-responders?
- Treatments:
 - JASP, AAC, JASP+, (JASP+AAC)+

In the slides, JASP is shorthand for JASP+EMT

The joint attention social engagement (JASP) intervention was combined with two interventions, enhanced milieu teaching (EMT) and augmentative and alternative communication (AAC). JASP ([Adamson et al. 2004](#); [Kasari et al. 2006, 2008](#)) was developed to facilitate a state of supported or coordinated joint engagement between the child and a social partner. Both EMT and AAC were developed to facilitate expressive language in young children with developmental disabilities. EMT ([Hancock & Kaiser 2006](#)) is a naturalistic language intervention that promotes functional use of new language forms in the context of everyday interactions with parents and other social partners. The AAC intervention utilizes a developmentally chosen augmentative communication device ([Cafiero 2005](#)) to facilitate communicative exchanges within play routines and daily activities. Both EMT and AAC were adapted for 5- to 8-year-old children and integrated with JAE to form two interventions, JASP + EMT and JASP + EMT + AAC. More intensive versions of both JASP + EMT and JASP + AAC included additional sessions provided by a skilled child therapist and additional training with the parent to promote parent and child generalization. Overall, four intervention options are considered: JASP + EMT, JASP + EMT + AAC, intensified JASP + EMT, and intensified JASP + EMT + AAC.

Kasari ASD SMART

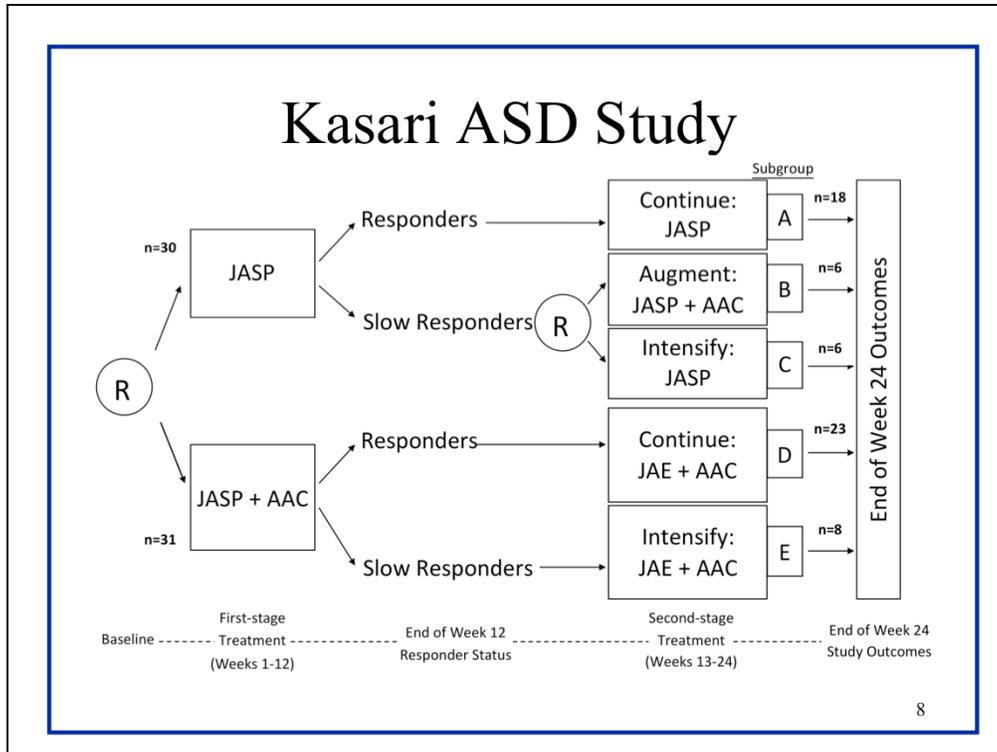
- Embedded Tailoring Variables: (*a*) total social communicative utterances, (*b*) percentage communicative utterances, (*c*) number different word roots, (*d*) mean length of utterance in words, (*e*) number of utterances where the function is to comment (rather than request), (*f*) words per minute, and (*g*) unique word combinations (included only if the child's target talk consists of more than two words).

Kasari ASD SMART

- How are the embedded tailoring variables used?
 - Two diffs for each of the 7 variables is calculated: one based on a language sample, the other based on a session sample. This yields 14 measures.
- How and when is (non) response assessed?
 - At 12 weeks
 - The criterion for response is 25% or more improvement on at least 50% of the 14 measures.

for each assessment, the first variable was calculated as the difference in the average assessment between the first two intervention sessions and the last two intervention sessions during the first stage of the intervention; the second variable was calculated as the difference between the assessment at the screening visit and the month-three visit. The above measures are collected via videotapes of the child and therapist sessions.

Preliminary studies indicated that these interventions should show changes within a 3 month period; this time frame is consistent with recommendations by the National Research Council



6 month trial. Follow up outcomes collected at wk36

This trial was just wrapped up.

See

<http://www.semel.ucla.edu/ASD/research/project/ccnia-developmental-augmented-intervention-facilitating-expressive-language>

CCNIA=characterizing cognition in nonverbal individuals with ASD

N=96 6 month trial

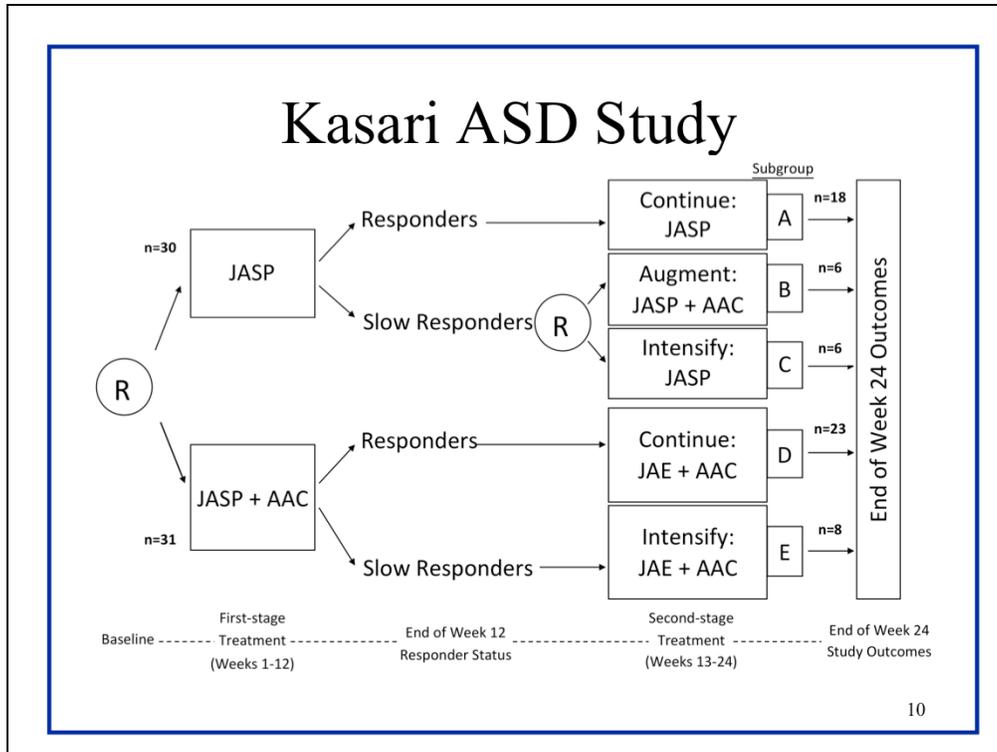
But difficulties in recruiting autistic children made it difficult to meet this goal. They only achieved N=61.

JASP in the picture is actually short for JASP+EMT = Joint attention social play + enhanced mileau training. See next slides.

Kasari ASD SMART

3 Embedded Adaptive Treatment Strategies

- 1) Start with JASP; if non-responder JASP +AAC, else JASP
- 2) Start with JASP; if non-responder JASP⁺, else JASP
- 3) Start with JASP+AAC; if non-responder (JASP+AAC)⁺, else JASP+AAC



6 month trial. Follow up outcomes collected at wk36

This trial was just wrapped up.

See

<http://www.semel.ucla.edu/ASD/research/project/ccnia-developmental-augmented-intervention-facilitating-expressive-language>

CCNIA=characterizing cognition in nonverbal individuals with ASD

N=96 6 month trial

But difficulties in recruiting autistic children made it difficult to meet this goal. They only achieved N=61.

JASP in the picture is actually short for JASP+EMT = Joint attention social play + enhanced mileau training. See next slides.

Kasari ASD SMART

- Primary Analysis
 - To compare the slopes in outcome measures of communication and language over 6 months for the two initial treatments: JASP vs JASP+AAC
- Secondary Analyses
 - Investigate moderation by baseline variables, investigate if other variables might be used to tailor treatment.

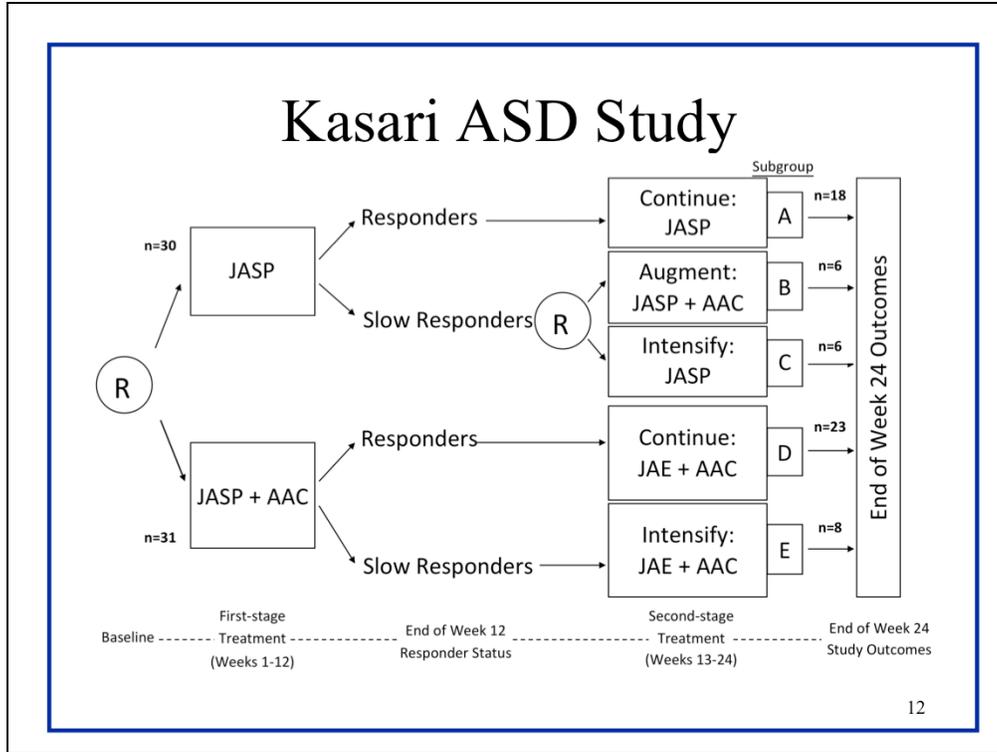
Primary Analyses involve:

Primary Y is socially communicative utterances (SCU).

Other outcomes were Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4) (given at 0, 6, 9 months): This test for receptive vocabulary development and is appropriate for children aged 2.6 years and older. and Verbal Motor Production Assessment for Children (VMPAC) (given at 0, 6, 9 months) The VMPAC is designed to examine oral and speech-motor control in children. The items are arranged from basic to complex and assess three main areas: Global motor control, focal oromotor control and sequencing.

Secondary Analyses involve:

The baseline variables included severity of repetitive/ compulsive behaviors, degree of apraxia, and developmental variables (based on cognitive and language test results). In particular, the research team hypothesized that children with greater severity of apraxia would do better on beginning with JASP + AAC than beginning with JASP because the communication device would better provide a means to communicate.



6 month trial. Follow up outcomes collected at wk36

This trial was just wrapped up.

See

<http://www.semel.ucla.edu/ASD/research/project/ccnia-developmental-augmented-intervention-facilitating-expressive-language>

CCNIA=characterizing cognition in nonverbal individuals with ASD

N=96 6 month trial

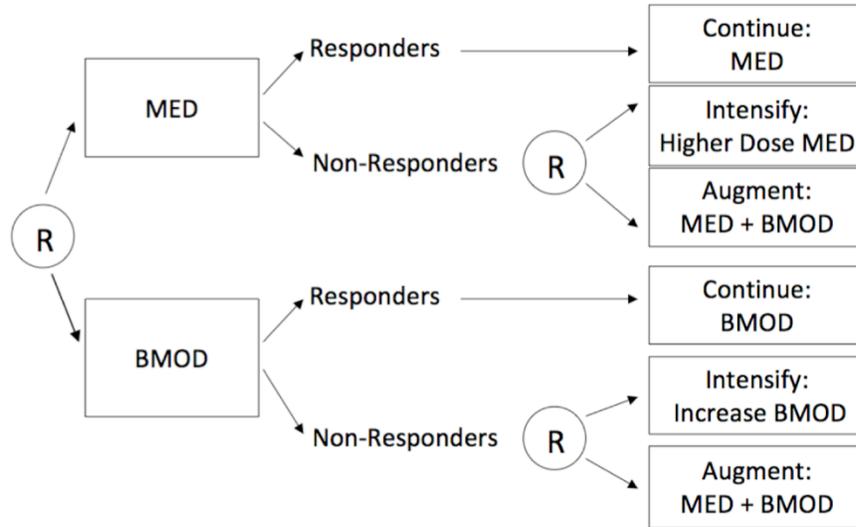
But difficulties in recruiting autistic children made it difficult to meet this goal. They only achieved N=61.

JASP in the picture is actually short for JASP+EMT = Joint attention social play + enhanced mileau training. See next slides.

Outline

- Adaptive ASD Developmental and Augmented Intervention (Kasari, PI)
- Adaptive Pharmacological and Behavioral Treatments for Children with ADHD Trial (Pelham, PI)
- Adaptive Reinforcement-Based Treatment for Pregnant Drug Abusers (Jones, PI)
- Extending Treatment Effectiveness of Naltrexone (Oslin, PI)
- Comparison of SMARTs

Pelham ADHD Study (in schools)



1-school year study (approx. 8 months) N=153

Note non-response is assessed monthly beginning at month 2 (8 weeks)

Pelham ADHD SMART

Population & Rationale:

- Children with ADHD, ages 6-12
- Much debate on whether the first-line intervention should be pharmacological or behavioral.
- Planning for a “rescue” if the first treatment does not go well is crucial because 20-50% do not substantially improve on first treatment.

for example, a task force of the American Psychological Association recommends psychosocial first ([Brown et al. 2007](#)), whereas the guidelines of the [American Academy of Child and Adolescent Psychiatry \(2007\)](#) recommend using medication first.

Pelham ADHD SMART

- Critical Decisions:
 - Which treatment to provide first? Which treatment to provide non-responders?
- Treatments:
 - Med, Med⁺⁺, BMOD, BMOD⁺⁺

Med is ritalin

The interventions include differing doses of methylphenidate (a psychostimulant drug) and differing intensities of behavioral modification (consisting of a school-based component with the teacher, a Saturday treatment component involving social skills development, and a parent-training component targeted at helping parents to identify problematic behaviors with the relevant child-functioning domains). The higher-dose option for methylphenidate includes late-afternoon doses, if needed. The higher-intensity option for the behavioral modification includes more intensive training in social skills in the school-based component and, if needed, both additional individual parent training sessions that target specific behavior management issues and practice sessions with children.

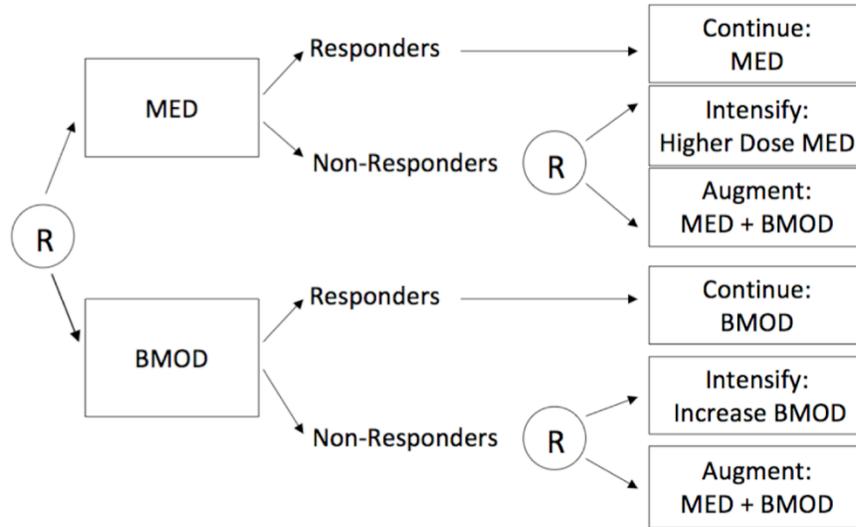
Pelham ADHD SMART

- Embedded Tailoring Variables: (a) Teacher reported Impairment Scale (IRS), (b) Teacher reported individualized list of target behaviors (ITB)
- How and when is (non) response assessed?
 - At 8 weeks and every 4 weeks thereafter
 - The criterion for non-response is an average performance of less than 75% on the ITB and a rating of impairment in at least one domain on the IRS.

The Impairment Rating Scale (IRS) ([Fabiano et al. 2006](#)) and an individualized list of target behaviors (ITB) (e.g., [Pelham et al. 1992](#)). The IRS provides a comprehensive index of a child's impairment in various domains such as peer relationships, classroom behavior, family functioning, and academic achievement. The ITB was used to assess improvement on child-specific behavior goals.

Investigators felt that 8 weeks was needed in order to obtain a reasonable assessment of children's response to treatment and to give clinicians time to implement the school-based interventions and conduct parent training

Pelham ADHD Study (in schools)



1-school year study (approx. 8 months) N=153

Note non-response is assessed monthly beginning at month 2 (8 weeks)

Pelham ADHD SMART

4 Embedded Adaptive Treatment Strategies

- 1) Start with BMOD; if non-responder BMOD⁺⁺, else BMOD
- 2) Start with BMOD; if non-responder BMOD +Med, else BMOD
- 3) Start with Med; if non-responder Med⁺⁺, else Med
- 4) Start with Med; if non-responder BMOD +Med, else Med.

Pelham ADHD SMART

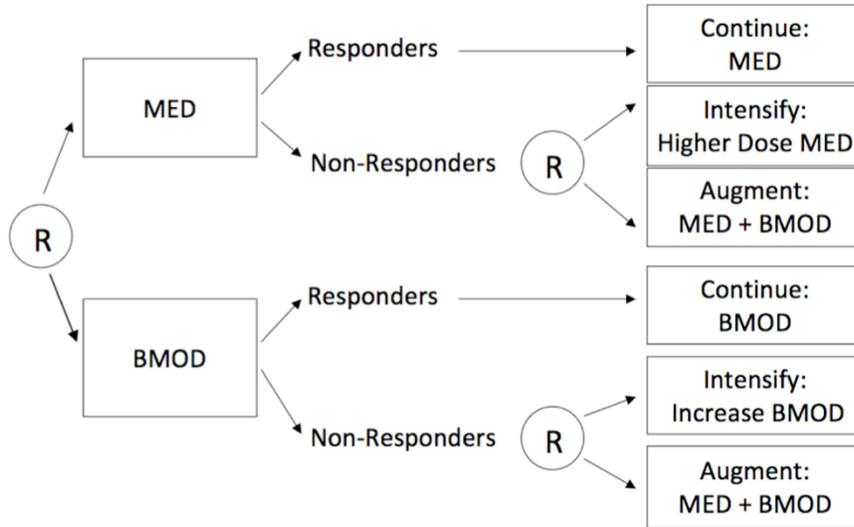
4 Embedded Adaptive Treatment Strategies

****conceptualized in terms of *tactics*****

- 1) Start with BMOD; if non-responder intensify, else continue same
- 2) Start with BMOD; if non-responder augment with other treatment, else continue same
- 3) Start with Med; if non-responder intensify, else continue same
- 4) Start with Med; if non-responder augment with other treatment, else continue same.

Conceptualize second stage in terms of tactics as opposed to the treatments.....

Pelham ADHD Study (in schools)



1-school year study (approx. 8 months) N=153

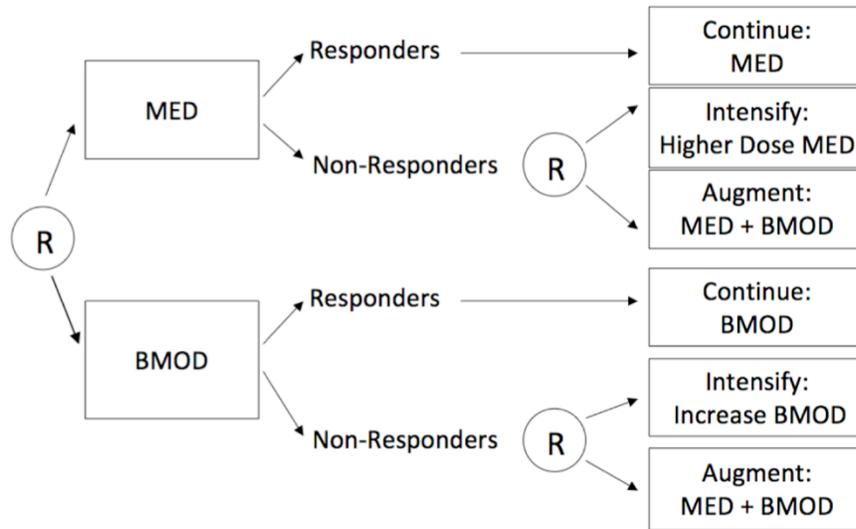
Note non-response is assessed monthly beginning at month 2 (8 weeks)

Pelham ADHD SMART

- **Primary Analysis**
 - To compare the change in teacher ratings of child behavior across 8 months for the two treatments: Med first strategies vs BMOD first strategies
- **Secondary Analyses**
 - Investigate moderation of the effect of initial treatment/secondary treatment/adaptive treatment strategies by baseline variables; investigate if other variables might be used to tailor treatment.

Potential baseline moderator was whether the child had received medication for ADHD in prior year.

Pelham ADHD Study (in schools)



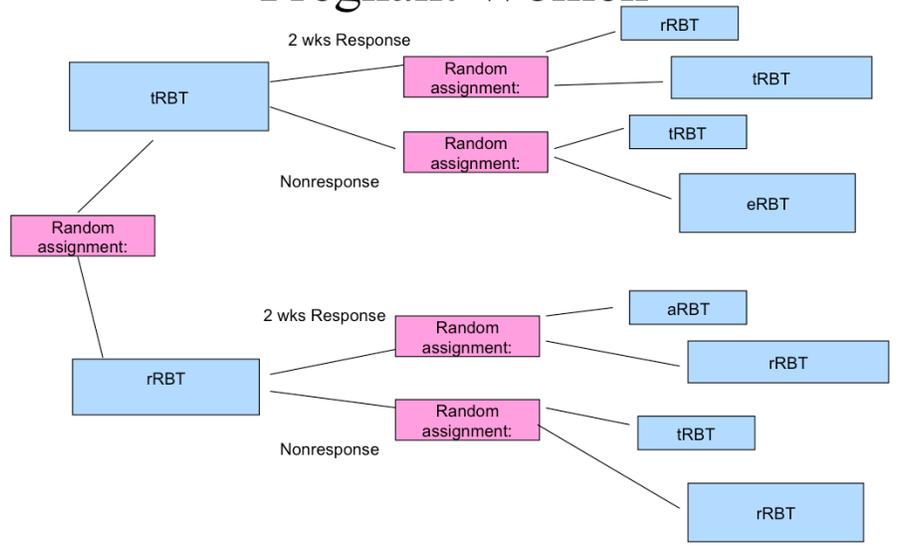
1-school year study (approx. 8 months) N=153

Note non-response is assessed monthly beginning at month 2 (8 weeks)

Outline

- Adaptive ASD Developmental and Augmented Intervention (Kasari, PI)
- Adaptive Pharmacological and Behavioral Treatments for Children with ADHD Trial (Pelham, PI)
- Adaptive Reinforcement-Based Treatment for Pregnant Drug Abusers (Jones, PI)
- Extending Treatment Effectiveness of Naltrexone (Oslin, PI)
- Comparison of SMARTs

Jones' Study for Drug-Abusing Pregnant Women



This study is in the field $n=300$
 RBT==reinforcement based tx

These differ in intensity and scope (in increasing order below)
 aRBT is abbreviated RBT
 rRBT is reduced RBT
 tRBT is traditional
 eRBT is enhanced

Jones Drug Abuse SMART

Population & Rationale:

- Drug Abusing (Opioid/Cocaine) Pregnant Women
- Reinforcement based treatment (RBT) is an efficacious intervention but,
 - RBT is costly to administer and time-consuming on the part of the participant,
 - About 40% of participants do not respond as well as desired

The women must have completed a eight-day residential detoxification stay

Jones Drug Abuse SMART

Critical Decisions:

- (a) Whether the frontline version of RBT can be reduced in intensity and scope;
- (b) whether a woman who does not respond quickly should continue on the same version or be moved to a more-intensive, larger-scope version of RBT; and
- (c) whether the intensity and scope of RBT can be reduced if a woman responds quickly.

Jones Drug Abuse SMART

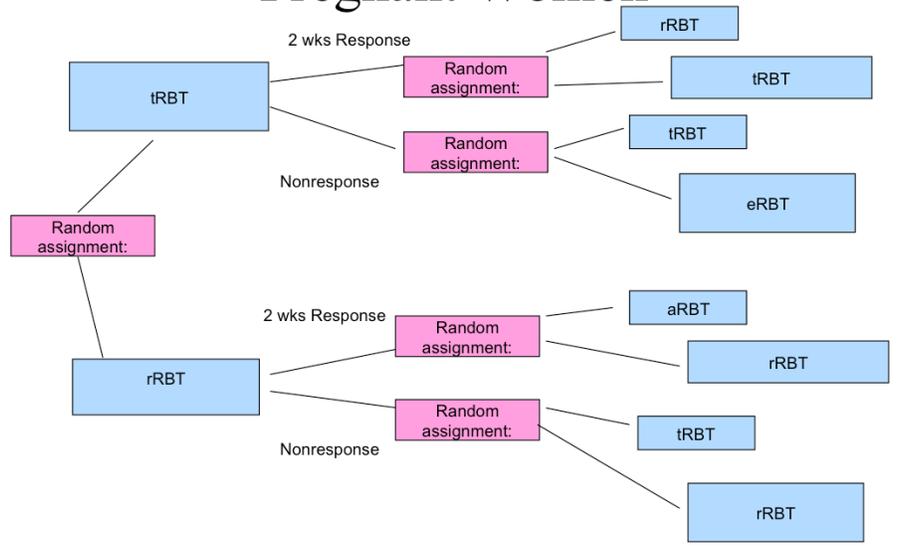
- Treatments:
 - aRBT < rRBT < tRBT < eRBT (increasing order in intensity/scope)
- Embedded Tailoring Variables:
 - a) self-reported drug use, b) results of urine tests, and c) attendance on intervention days

Jones Drug Abuse SMART

- How and when is (non) response assessed?
 - At 2 weeks
 - The criterion for nonresponse is missing an intervention day with no excuse, or a positive opioid/cocaine urine specimen, or self-report use of either drug.

Prior studies documented that the most vulnerable period for treatment drop-out is during the first two weeks of outpatient care and that very early drug use lapse or relapse is a predictor of poor treatment response

Jones' Study for Drug-Abusing Pregnant Women



Jones Drug Abuse SMART

8 Embedded Adaptive Treatment Strategies

- 1) Always tRBT
- 2) Start with tRBT; if non-responder tRBT, if responder rRBT
- 3) Start with tRBT; if non-responder eRBT, if responder tRBT
- 4) Start with tRBT; if non-responder eRBT, if responder rRBT

Jones Drug Abuse SMART

8 Embedded Adaptive Treatment Strategies

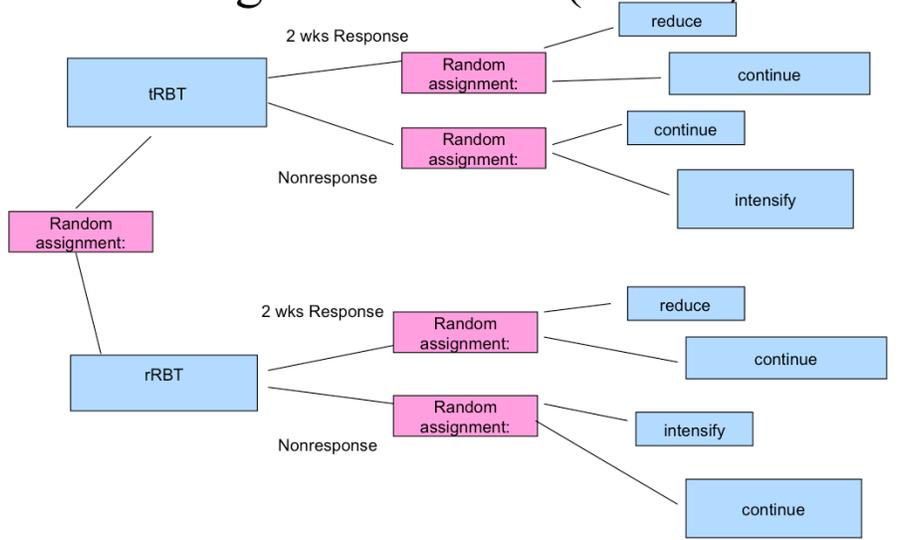
- 5) Always rRBT
- 6) Start with rRBT; if non-responder tRBT, if responder rRBT
- 7) Start with rRBT; if non-responder rRBT, if responder aRBT
- 8) Start with rRBT; if non-responder tRBT, if responder aRBT

Jones Drug Abuse SMART

- **Primary Analysis**
 - To compare program completion (delivery of child while in treatment) of the always tRBT arm versus the always rRBT arm (two non-adaptive strategies!)
- **Secondary Analyses**
 - Investigate moderation by baseline variables, investigate if other variables might be used to tailor treatment.

Secondary aims involve assessing the usefulness of candidate tailoring variables, such as the amount of illegal activity (e.g., prostitution).

Jones' Study for Drug-Abusing Pregnant Women (tactics)

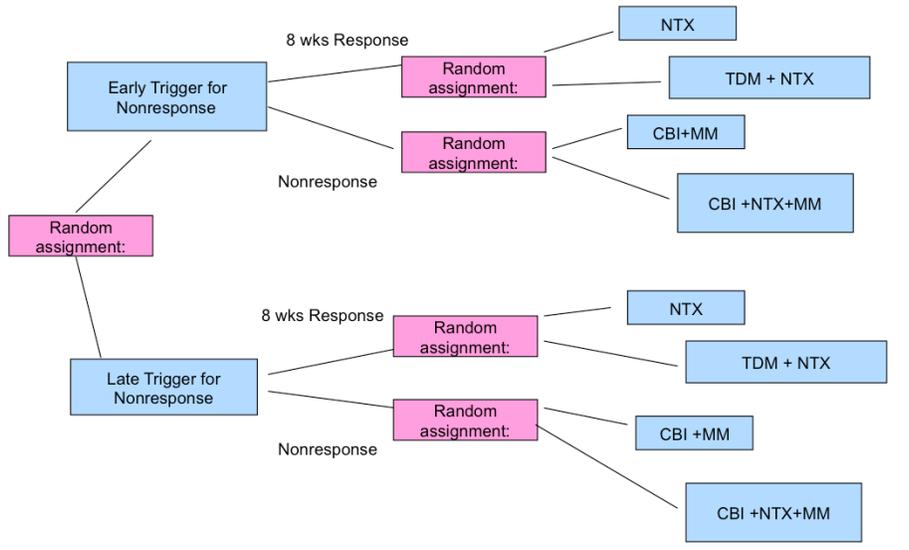


Other secondary analyses?

Outline

- Adaptive ASD Developmental and Augmented Intervention (Kasari, PI)
- Adaptive Pharmacological and Behavioral Treatments for Children with ADHD Trial (Pelham, PI)
- Adaptive Reinforcement-Based Treatment for Pregnant Drug Abusers (Jones, PI)
- Extending Treatment Effectiveness of Naltrexone (Oslin, PI)
- Comparison of SMARTs

Oslin Alcoholism SMART



All Alcohol dependent subjects begin on Naltrexone, an opioid receptor antagonist + medical management (NTX+MM)

N=302

Oslin Alcoholism SMART

Population & Rationale:

- Alcohol Dependent Adults who completed an Intensive Outpatient Program
- Naltrexone (NTX, an opiate antagonist) is efficacious but clinical use is limited.
 - Around 1/3 of patients relapse while on NTX.
 - Would like to inform longer term management based on NTX
 - Non-adherence is common

Oslin wrote in his justification: Despite the efficacy of naltrexone (NTX) for prevention of relapse to alcoholism as established by the majority of randomized clinical trials, as many as a third of subjects relapse while taking NTX. These studies have raised a second generation of questions regarding the best long-term management of subjects who are non responders: do these subjects require some type of augmented therapy or stepped care approach (more intensive psychotherapy, a second medication, etc.), should they be switched to a different therapy altogether and if so is there any benefit to remaining on NTX, or do they need further exposure to NTX to demonstrate a response? In considering testable hypotheses for non-responders we relied on our existing data and experience with other common chronic diseases such as depression, hypertension and arthritis. For instance in depression management, after treatment non-response with one medication it is usually assumed that a second medication or psychotherapy will be tried. However, there is considerable debate over whether the first medication should be continued or discontinued, as there may have been partial response to the first medication or potential synergistic effects with the second treatment. We are proposing to mirror this type of design by testing the benefits of remaining on NTX after adding a combination of motivational enhancement therapy and cognitive behavioral therapy (Combined Behavioral Intervention -CBI) to Medical Management (MM). Given the economic costs related to long term NTX treatment, we see this question as critical in developing long term treatment strategies that involve the use of NTX. The economic impact of this issue was highlighted by Ilstrup in a commentary on ineffective treatments . Given that a significant proportion of non-response to NTX may be due to non-adherence, a secondary aim of this project is to examine the role of medication adherence as a mediating factor in treatment improvement among those randomized to NTX.

Oslin Alcoholism SMART

Critical Decisions:

- (a) What extent of drinking behavior best reflects nonresponse to NTX?
- (b) What type of treatment would be useful for participants who do not respond adequately to NTX?
- (c) What type of treatment would be useful in reducing the chance of relapse among participants who respond adequately to NTX?

Oslin Alcoholism SMART

- Treatments:
 - NTX, MM, CBI, TDM
- Embedded Tailoring Variable:
 - Weekly self report of heavy drinking days.

These are NTX, medical management (MM), combined behavioral intervention (CBI), and telephone disease management (TDM). MM is a face-to-face, basic, minimal clinical support for the use of effective pharmacotherapy and reduction in drinking ([Pettinati et al. 2004](#), [2005](#)). CBI is a multicomponent intervention that includes components targeting adherence to pharmacotherapy and enhancement of participant motivation for change. This intervention includes family involvement when possible and emphasizes the utilization of the participant's social/community context to reinforce abstinence ([Longabaugh et al. 2005](#), [Miller et al. 2003](#)). TDM includes the same content as MM, but it is delivered via telephone.

Heavy drinking days (≥ 5 drinks/day for males; ≥ 4 for females)

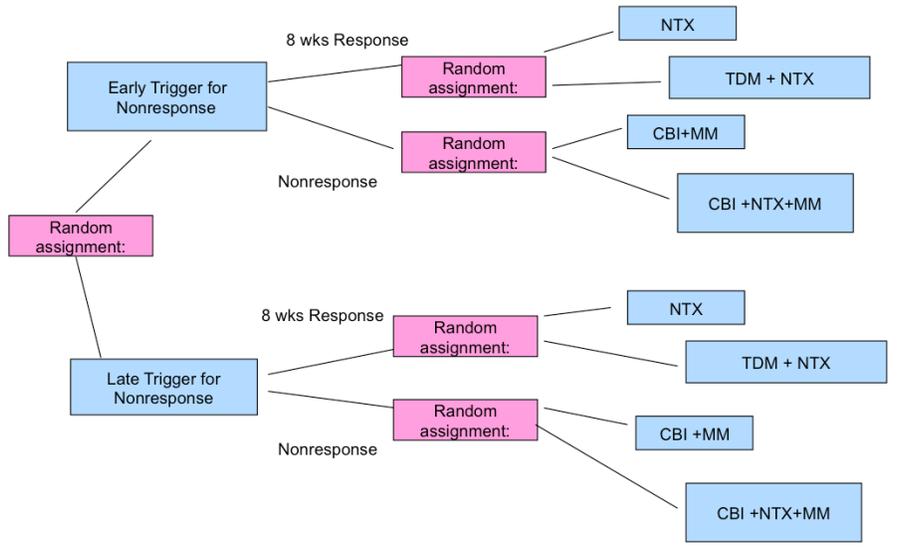
Oslin Alcoholism SMART

How and when is (non) response assessed?

- Initially, each week, for 8 weeks
 - Early trigger criterion for non-response: 2 or more heavy drinking days since beginning NTX
 - Late trigger criterion for non-response : 5 or more heavy drinking days since beginning NTX
- If, after 8 weeks, the non-response criterion is not met then the participant is a responder.

This criterion was supported by preliminary data generated from a prior NTX study conducted. This study gave alcohol dependent subjects for 100mg/day or placebo with a less structured form of medical monitoring called BRENDA for 32 weeks. Results indicated that subjects who had taken the NTX (not placebo) and had 2 to 5 days of heavy drinking in the first 60 days were not likely to reduce their drinking if they just continued NTX and medical management.

Oslin Alcoholism SMART



All Alcohol dependent subjects begin on Naltrexone, an opioid receptor antagonist + medical management (NTX+MM)

N=302

Oslin Alcoholism SMART

8 Embedded Adaptive Treatment Strategies

- 1) Start with NTX+MM; if 2 HDD occurs prior to 8 weeks, augment to CBI+NTX+MM, else at 8 weeks continue on NTX
- 2) Start with NTX+MM; if 2 HDD occurs prior to 8 weeks, switch to CBI +MM, else at 8 weeks continue on NTX
- 3) Start with NTX+MM; if 2 HDD occurs prior to 8 weeks, augment to CBI+NTX+MM, else at 8 weeks continue on NTX and add TDM

HDD: heavy drinking days (≥ 5 drinks/day for males; ≥ 4 for females)

Oslin Alcoholism SMART

8 Embedded Adaptive Treatment Strategies

- 4) Start with NTX+MM; if 2 HDD occurs prior to 8 weeks, switch to CBI+MM, else at 8 weeks continue on NTX and add TDM
- 5) ..
- 6) ..
- 7) ..
- 8) ..

HDD: heavy drinking days (≥ 5 drinks/day for males; ≥ 4 for females)

Oslin Alcoholism SMART

- **Primary Analysis**
 - Focus on non-responders to NTX+MM.
Compare drinking outcomes (e.g. percent days abstinent) on CBI+NTX+MM versus to CBI+MM.
- **Secondary Analyses**
 - Test effectiveness of TDM for responders; test two criteria for non-response; assess moderation (psychosocial distress, severity of alcohol dependence, adherence in first stage)

Note the primary aim. Quite different from other case studies.

Outline

- Adaptive ASD Developmental and Augmented Intervention (Kasari, PI)
- Adaptive Pharmacological and Behavioral Treatments for Children with ADHD Trial (Pelham, PI)
- Adaptive Reinforcement-Based Treatment for Pregnant Drug Abusers (Jones, PI)
- Extending Treatment Effectiveness of Naltrexone (Oslin, PI)
- Comparison of SMARTs

Comparison of SMART Studies

Dimensions

- 1) Which participants are multiply randomized?
- 2) When are participants re-randomized?
- 3) The types of the critical decisions
- 4) What are the primary research questions?

Comparison of SMART Studies

Which participants are multiply randomized?

- A subset of non-responders: ASD (only nonresponders to JASP)
- All non-responders: ADHD, Drug Abusing Pregnant Women, Alcohol Dependence
- All responders: Drug Abusing Pregnant Women, Alcohol Dependence

47

The larger the number of categories of people re-randomized, the larger the number of embedded adaptive treatment strategies.

Comparison of SMART Studies

When are participants randomized for the second time?

- At one fixed point in time only
 - ASD (month 3), Drug Abusing Pregnant Women (week 2), Alcohol Dependence (responders at week 8)
- At any one of several fixed times
 - ADHD (at month 2 and each month thereafter), Alcohol Dependence (non-responders at week 2 and weekly until week 8)

48

Also in both the ADHD and the Alcohol Dependence SMARTS as soon as non-response detected, the participant is re-randomized.

Comparison of SMART Studies

What kinds of critical decisions are investigated?

- Which treatment first and which second?
 - ASD, ADHD, Drug Abusing Pregnant Women
- How soon to give up on initial treatment and which treatment to provide second?
 - Alcohol Dependence

Comparison of SMART Studies

What are the primary research questions?

- Comparison of stage 1 treatments, controlling, by design, for stage 2 treatments.
 - ASD, ADHD
- Comparison of stage 2 treatments, controlling, by design, for stage 1 treatment
 - Alcohol Dependence (non-responders)
- Comparison of two embedded treatment strategies.
 - Drug Abusing Pregnant Women

50

These are the comparisons that are used to size the SMART

Questions?

More information:

H. Lei, I. Nahum-Shani, K. Lynch, D. Oslin and S.A. Murphy. A SMART Design for Building Individualized Treatment Sequences, *The Annual Review of Clinical Psychology* (2012), Vol. 8: 21-48

Practice Exercise

Exercise 1: Re-visit/Re-visit the list of critical questions you made at the end of Module 1. This is the list of critical questions that need to be addressed to develop a high-quality AI.

Exercise 2: Re-visit the draft SMART design you started at the end of Module 2. Make changes, if necessary.

Primary Aims Using Data Arising from a SMART

Module 4

Getting SMART about Developing Adaptive Interventions:
Individualizing Sequences of Treatment

SBM 2014; Philadelphia, PA

Daniel Almirall, Billie Nahum-Shani, Susan A. Murphy



60 minutes

This is the 4th module of a 6-module Seminar on experimental designs for building optimal adaptive health interventions.

By now, you know what an ATS is. You have discussed why they are important in terms of managing chronic disorders (indeed, an ATS formalizes the type of clinical practice taking place today). And, you have been introduced to the SMART clinical trial design, the rationale for SMARTs, and some important SMART design principles.

In this module, we are going to discuss data analysis methods used to address 2 of the typical primary research aims posed in SMART trials. We are also going to warm up and begin to describe a 3rd primary aim (which we finish describing in Part II, Module 5)

Primary Aims Part I, Outline

- Review the *Adaptive Interventions for Children with ADHD Study* design (a SMART design)
- Will learn how to analyze two typical primary research questions in a SMART design
 - PI(a): Main effect of initial (first-stage) treatment?
 - PI(b): Comparing second-stage tactics?
- Will prepare for a third primary aim analysis by
 - PI(c): Learning to estimate the mean outcome under each of the embedded ATS (separately) using an easy-to-use weighting approach

Outline

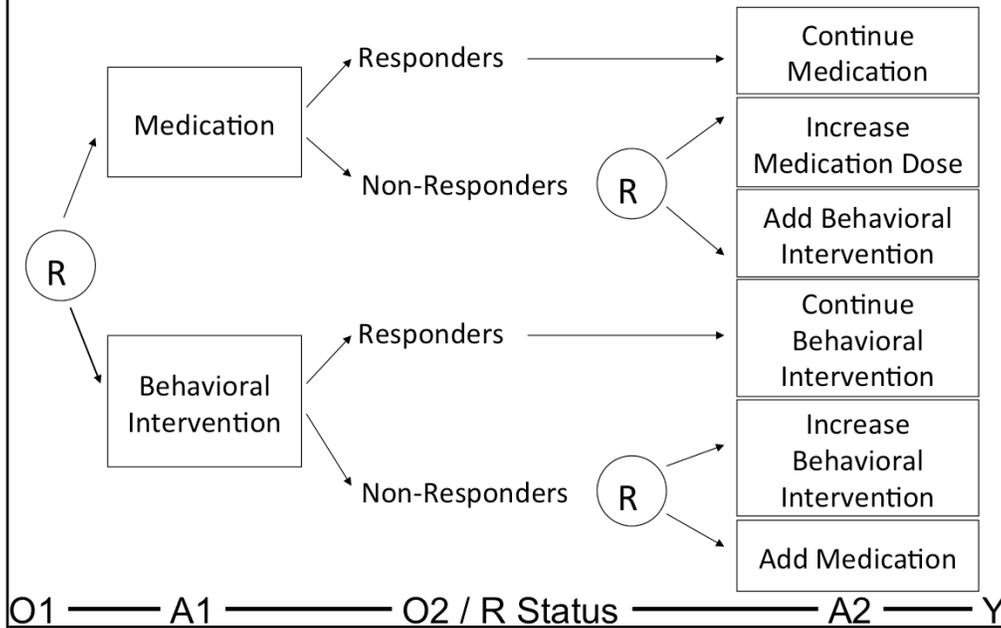
Primary Aims Part I, Outline

- Review the *Adaptive Interventions for Children with ADHD Study* design (a SMART design)
- Will learn how to analyze two typical primary research questions in a SMART design
 - PI(a): Main effect of initial (first-stage) treatment?
 - PI(b): Comparing second-stage tactics? treatments?
- Will prepare for a third primary aim analysis by
 - PI(c): Learning to estimate the mean outcome under each of the embedded ATS (separately) using an easy-to-use weighting approach

We begin by reviewing the ADHD smart study

Review the ADHD SMART Design

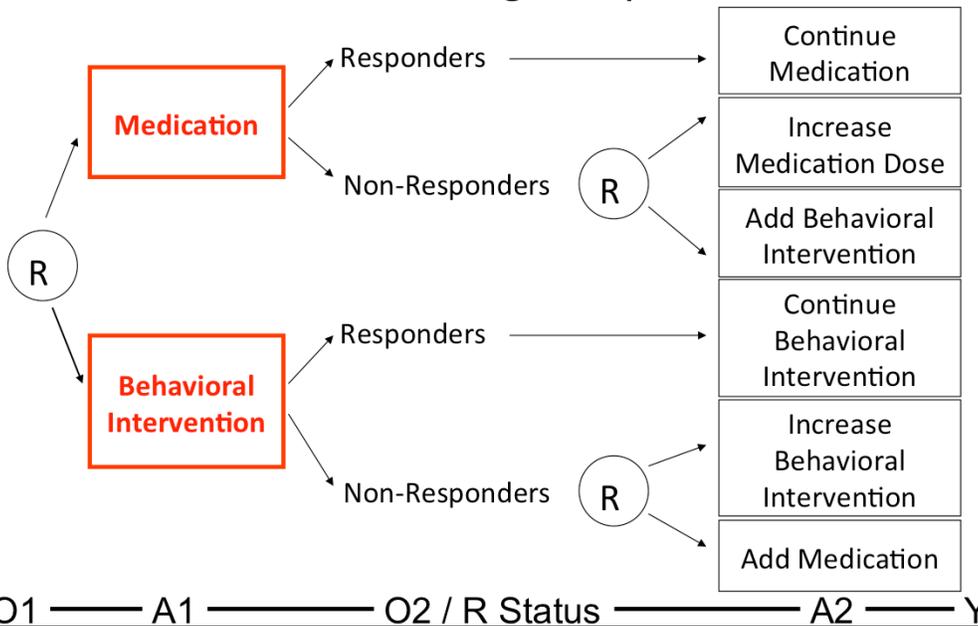
For the next 3 modules, keep handy the handout that describes this design.



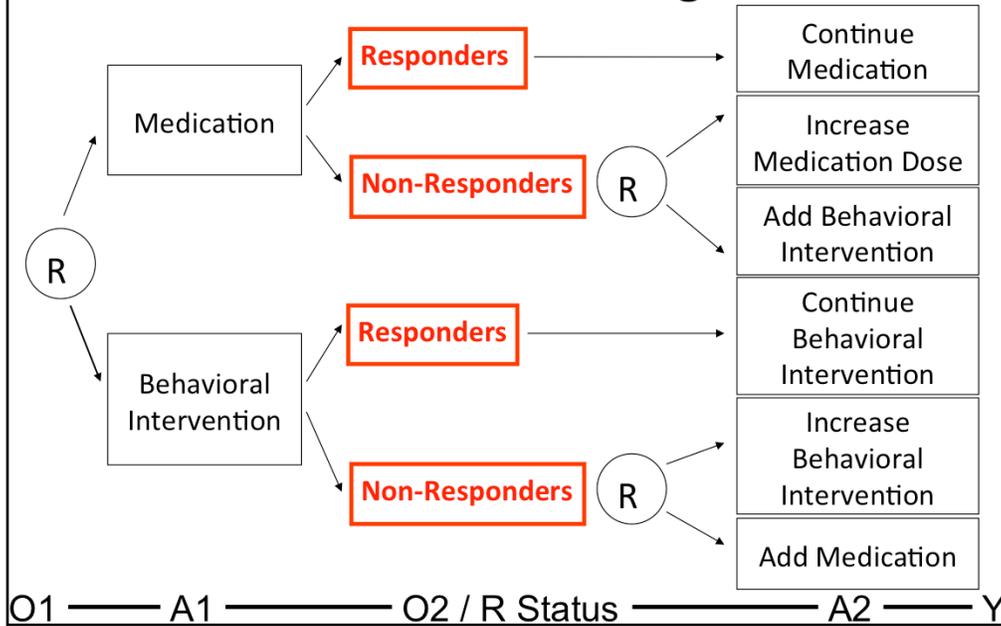
You were introduced to this SMART in modules 1, 2 and 3.

Let's review some of the characterizing features of this SMART design.

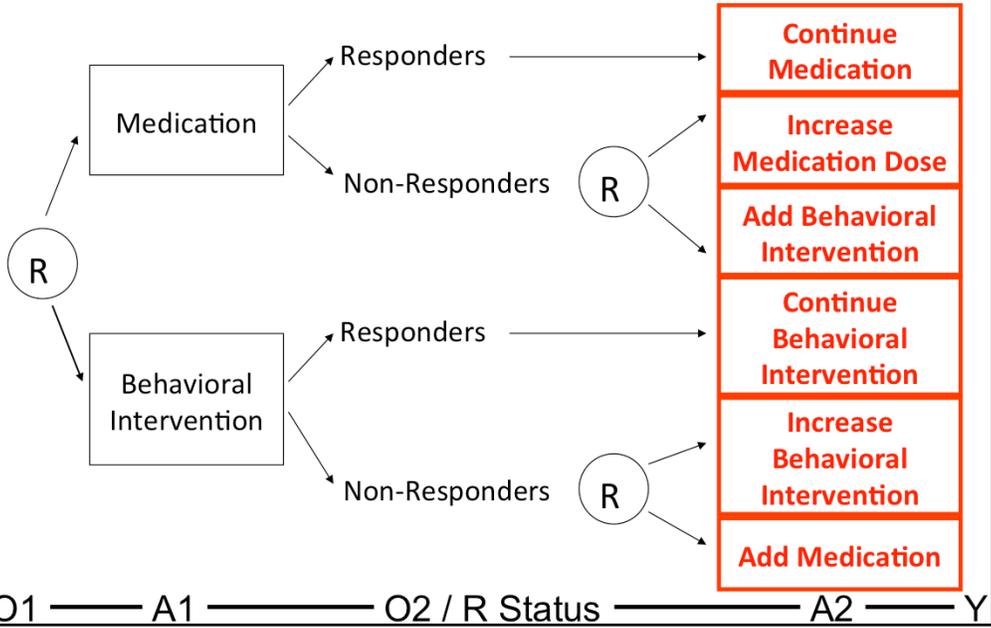
There are two “stage 1” treatment options that are being compared

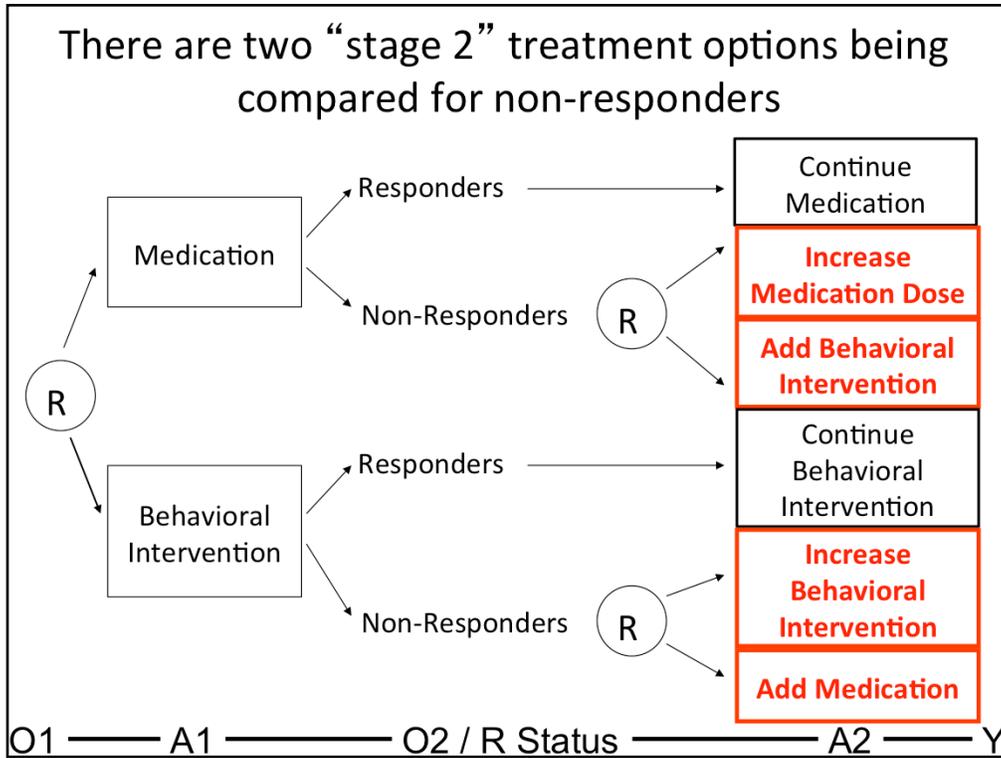


Response/non-response until Month 8 is the embedded tailoring variable



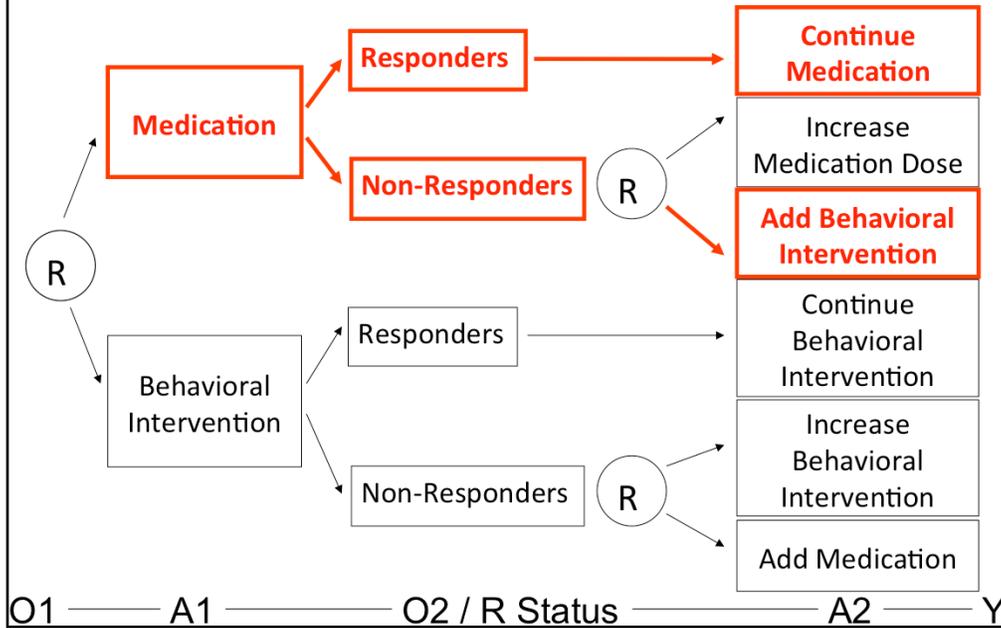
There are a total of 6 “stage 2” treatments that any one participant may receive





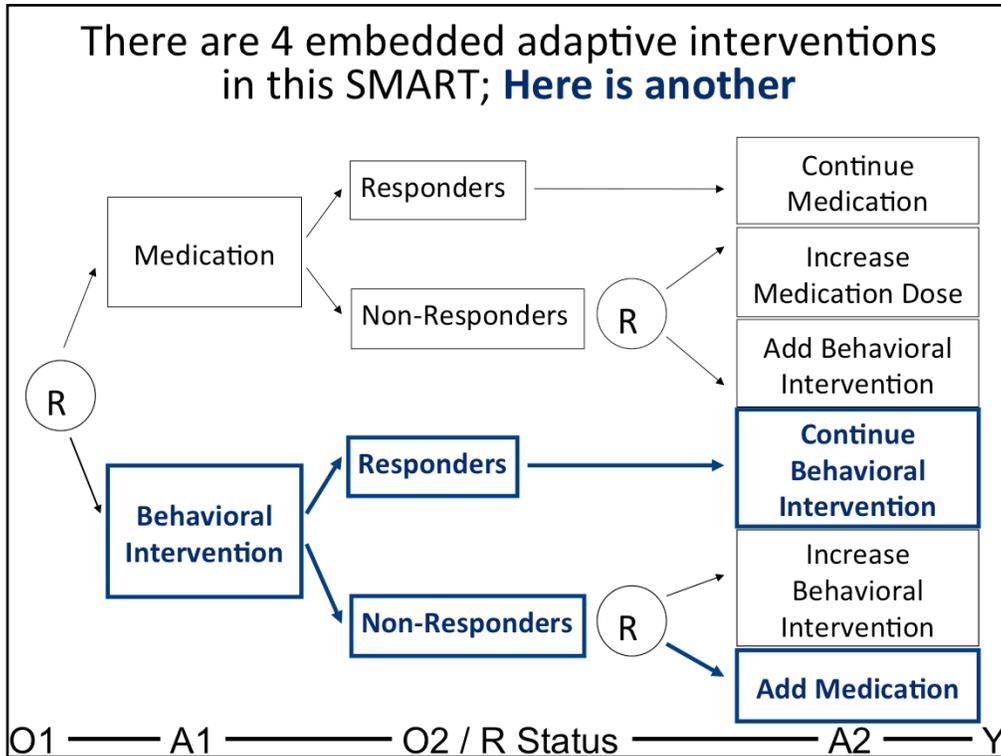
Tactical treatment options

There are 4 embedded adaptive interventions in this SMART; **Here is one**



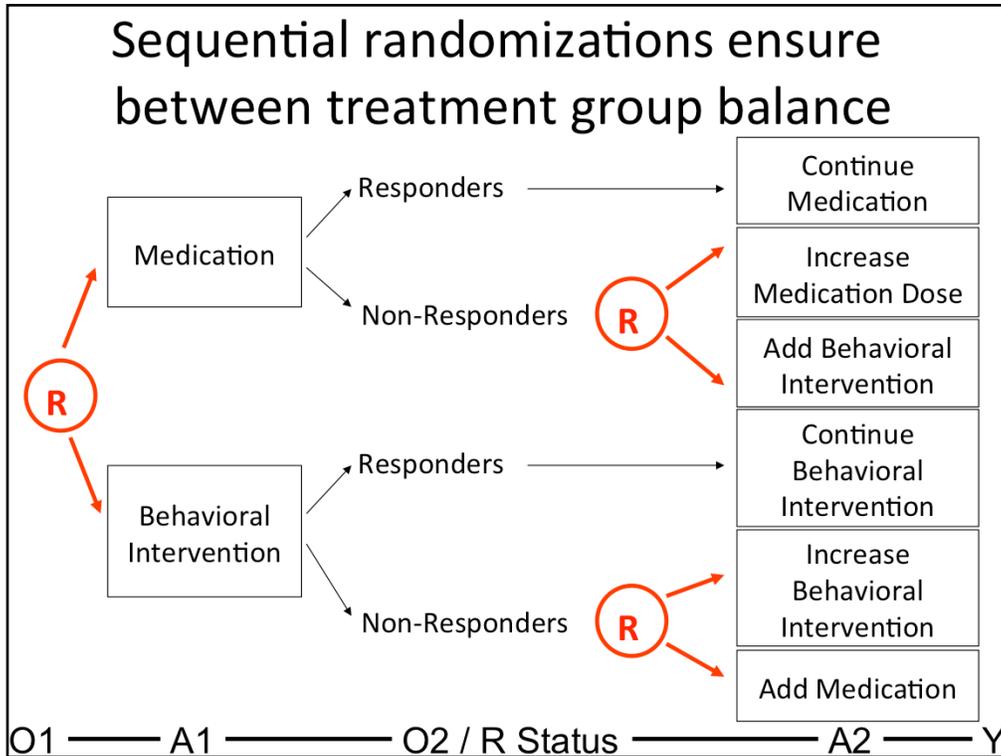
We call responder/non-responder status the “embedded tailoring variable” sine it is part of the study design in the sense that it restricts future treatment assignments.

Later, in Module 6, we describe analyses that permit investigators to study additional non-embedded (but potentially very useful) tailoring variables collected as part of O2.



This is another adaptive treatment strategy .

Note, randomizations have nothing to do with the definition of the each of the 4 adaptive intervention (i.e., adaptive interventions) . Important to distinguish between the interventions (the strategies) and the experiment (the randomization). This leads to the next characterizing feature of this design...[next slide] which is the sequential randomizations.



The sequential randomizations ensure unbiased comparisons between assigned treatments both initially (at the first line) and in the future (at the second stage) among non-responders.

A subset of the data arising from this SMART may look like this (part 1)

	ODD Dx	Baseline ADHD Score	Prior Med ?	First Line Txt	Resp/ Non-resp	Second Line Txt	School Perfm
ID	O11	O12	O13	A1	R	A2	Y
1	1	1.18	0	-1 MED	1	.	3
2	0	-0.567	0	-1	0	1 INTSFY	4
3	0	0.553	1	1 BMOD	0	-1 ADD	4
4	0	-0.013	0	1	0	-1	4
5	0	-0.571	1	1	0	1	2
6	0	-0.684	1	1	0	-1	4
7	0	1.169	0	-1	1	.	3
8	0	0.369	1	-1	0	1	3

This is simulated data.

Refer to your handouts to understand what these variables are. Go slow here and explain...

Y is the end-of-study outcome, measured after initial and second line treatments. Here Y is continuous end of study outcome measuring school performance, on 1 to 5 scale.

O11 O12 and O13 are baseline covariates. In the simulated data online

O11 = ODD Dx,

O12 = pre-txt ADHD scores,

O13 = Whether or not child had taken medication prior to enrolling in the trial

O14 (not shown in this slide) = race = white=1 or nonwhite=0

A1 = 1 = behavioral modification initially

A1 = -1 = medication initially

A2 = 1 = intensified the initial intervention

A2 = -1 = added the other intervention to the initial one

R = 1 = response

R = 0 = non-response

In addition to R, there can be other covariates measured after A1 but before A2, such as

O21 = Time in weeks until non-response (only measured for those with R=0)

O22 = Adherence to first-line treatment = YES(1) or NO(0).

Note that A2 is not applicable/missing by design if R = 1 = response because all participants who respond continue getting their initial treatment

In the data A2 can be either missing '.' for this subjects, or it can be some other number 99. That data will not get used.

A subset of the data arising from this SMART may look like this (part 2)

	Race	First Line Txt	Resp/Non-resp	Time until NR (mnths)	Adherence	Second Line Txt	School Perfm
ID	O14	A1	R	O21	O22	A2	Y
1	1	-1 MED	1	.	0	.	3
2	0	-1	0	6	0	1 INTSFY	4
3	0	1 BMOD	0	1	1	-1 ADD	4
4	0	1	0	7	0	-1	4
5	0	1	0	5	1	1	2
6	0	1	0	3	1	-1	4
7	0	-1	1	.	0	.	3

This is simulated data.

Refer to your handouts to understand what these variables are. Go slow here and explain...

Y is the end-of-study outcome, measured after initial and second line treatments. Here Y is continuous end of study outcome measuring school performance, on 1 to 5 scale.

O14 = race = white=1 or nonwhite=0

O21 = month until non-response

O22= adherence

A1 = 1 = behavioral modification initially

A1 = -1 = medication initially

A2 = 1 = intensified the initial intervention

A2 = -1 = added the other intervention to the initial one

R = 1 = response

R = 0 = non-response

In addition to R, there can be other covariates measured after A1 but before A2, such as

O21 = Time in weeks until non-response (only measured for those with R=0)

O22 = Adherence to first-line treatment = YES(1) or NO(0).

Note that A2 is not applicable/missing by design if R = 1 = response because all participants who respond continue getting their initial treatment

In the data A2 can be either missing '.' for this subjects, or it can be some other number 99. That data will not get used.

Primary Aims Part I, Outline

- Review the *Adaptive Interventions for Children with ADHD Study* design (a SMART design)
- Will learn how to analyze two typical primary research questions in a SMART design
 - PI(a): Main effect of initial (first-stage) treatment?
 - PI(b): Comparing second-stage tactics? treatments?
- Will prepare for a third primary aim analysis by
 - PI(c): Learning to estimate the mean outcome under each of the embedded ATS (separately) using an easy-to-use weighting approach

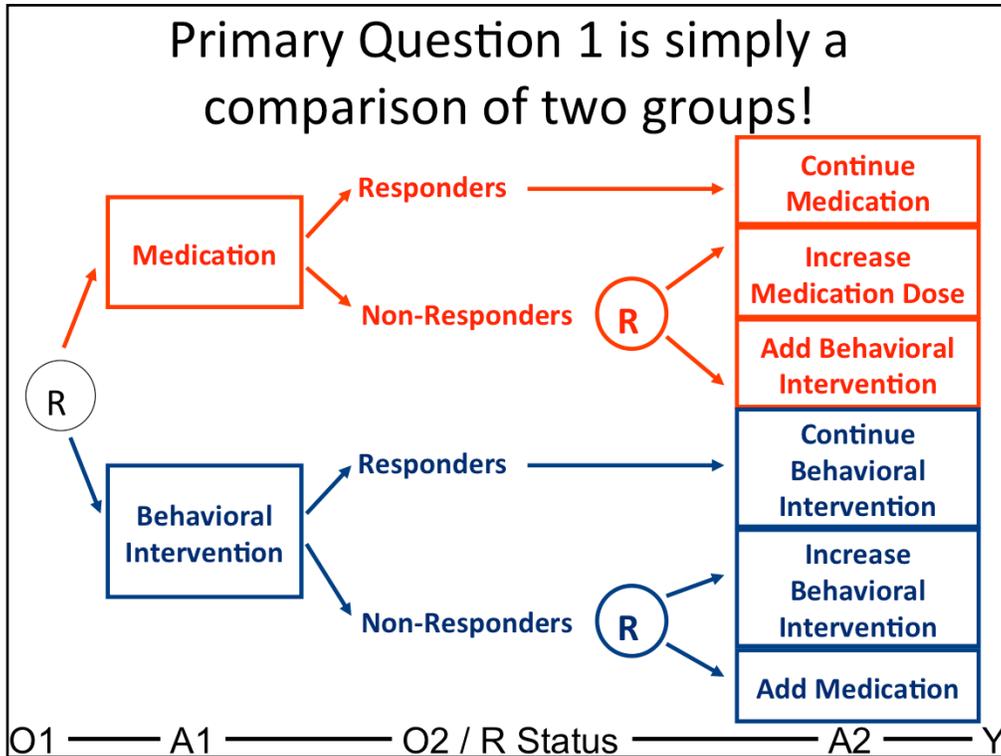
Typical Primary Aim 1: Main effect of first-line treatment? Stated 3 ways.

- “What is the best first-line treatment in terms of long-term outcomes, controlling for future treatment by design?”
- “What is the effect in terms of end of study school performance of starting with MED vs starting with BMOD?”
- “Is it better on average, in terms of end of study mean school performance, to begin treatment with BMOD or with MED?”

These are different ways to talk/write about Typical Primary Question #1:

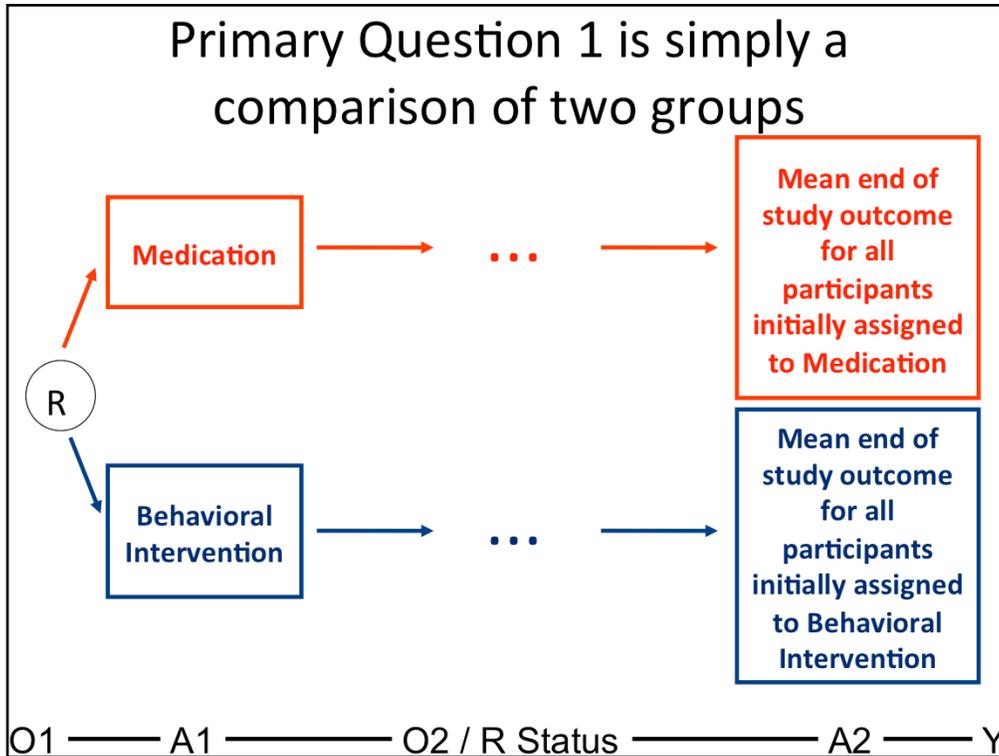
On average, how do longitudinal outcomes differ between children assigned first to medication versus children assigned first to behavioral intervention?

On average, what is the between-groups difference in change in outcomes from baseline to 8 months between children assigned first to behavioral intervention versus children assigned first to medication?



Given a continuous, end of study (e.g., 12 weeks) outcome, then a two-sample t-test is all that is needed.

This is just a comparison of two groups of study participants (the blue participants versus the red participants).



The way to think about this is to think for the moment of the 2 arm RCT and imagine that even in those studies “we do things” or “things happen” even after we offer treatment. This is no different except here it is more like “we do things” because we actually control the future treatments by design..

Given a continuous, end of study (e.g., 12 weeks) outcome, then a two-sample t-test is all that is needed.

This is just a comparison of two groups of study participants (the blue participants versus the red participants).

Before we show you how to do this in SAS, we review contrast coding

Recall that $A1 = 1$ = behavioral modification = BMOD

Whereas $A1 = -1$ = medication = MED

The Regression and Contrast Coding Logic:

$Y = b_0 + b_1 * A1 + e$ or you can fit

$Y = b_0 + b_1 * A1 + b_2 * O11c + b_3 * O12c + b_4 * O13c + b_5 * O14c + e$
($O11c$, $O12c$ and $O13c$ are mean-centered $O11$, $O12$, $O13$)

Overall Mean Y under BMOD = $b_0 + b_1 * 1$

Overall Mean Y under MED = $b_0 + b_1 * (-1)$

Between groups diff = $b_0 + b_1 - (b_0 - b_1) = 2 * b_1$

Instead of a regression, you can also run a two-sample t-test. The regression might be more efficient, and most clinical trialists recommend using the regression approach and adjusting for covariates that were used in the stratified randomization procedure.

SAS code for a 2-group mean comparison in end of study outcome

```
* mean center covariates prior to regression;
data dat1;
  set adhddat;
  o11c = o11 - 0.3533333; o12c = o12 - -0.1205948;
  o13c = o13 - 0.3133333; o14c = o14 - 0.8067777;
run;
* run regression to get between groups difference;
proc genmod data = dat1;
  model y = a1 o11c o12c o13c o14c;
  estimate 'Mean Y under BMOD' intercept 1 a1 1;
  estimate 'Mean Y under MED' intercept 1 a1 -1;
  estimate 'Between groups difference' a1 2;
run;
```

This analysis is with simulated data.

Instead of a regression, you can also run a two-sample t-test. The regression is usually more efficient, and most clinical trialists recommend using the regression approach and adjusting for covariates that were used in the stratified randomization procedure.

You could also use a linear mixed model (HLM/growth curve) or any other standard longitudinal analysis to address this aim. A longitudinal analysis is recommended because it has more power!

The SAS code corresponds to a simple regression model

```
proc genmod data = dat1;  
  model y = a1 o11c o12c o13c o14c;  
  estimate 'Mean Y under BMOD' intercept 1 a1 1 o11c 0;  
  estimate 'Mean Y under MED' intercept 1 a1 -1;  
  estimate 'Between groups difference' a1 2;  
run;
```

In SAS “estimate” statements, setting a coefficient to zero is just like leaving it blank.

The Regression Logic:

$$Y = b_0 + b_1 \cdot A_1 + b_2 \cdot O_{11c} + b_3 \cdot O_{12c} + b_4 \cdot O_{13c} + b_5 \cdot O_{14c} + e$$

$$\text{Mean Y under BMOD} = E(Y | A_1=1) = b_0 + b_1 \cdot 1$$

$$\text{Mean Y under MED} = E(Y | A_1=-1) = b_0 + b_1 \cdot (-1)$$

$$\text{Between groups diff} = E(Y | A_1=1) - E(Y | A_1=-1)$$

$$= b_0 + b_1 - (b_0 - b_1) = 2 \cdot b_1$$

Instead of a regression, you can also run a two-sample t-test. The regression might be more efficient, and most clinical trialists recommend using the regression approach and adjusting for covariates that were used in the stratified randomization procedure.

Primary Question 1 Results

Contrast Estimate Results

Label	Estimate	95% Conf Limits		P-value
		Lower	Upper	
Mean Y under BMOD	3.0459	2.7859	3.3059	<.0001
Mean Y under MED	2.8608	2.6008	3.1208	<.0001
Between groups diff	0.1851	-0.1849	0.5551	0.3269

(SE = standard err)(0.1889)

In this simulated data set/experiment, it is slightly better to begin with BMOD (vs MED) in terms of school performance at end of study, but not statistically significant (p-value = 0.33).

This analysis is with simulated data.

Primary Question 1 Results

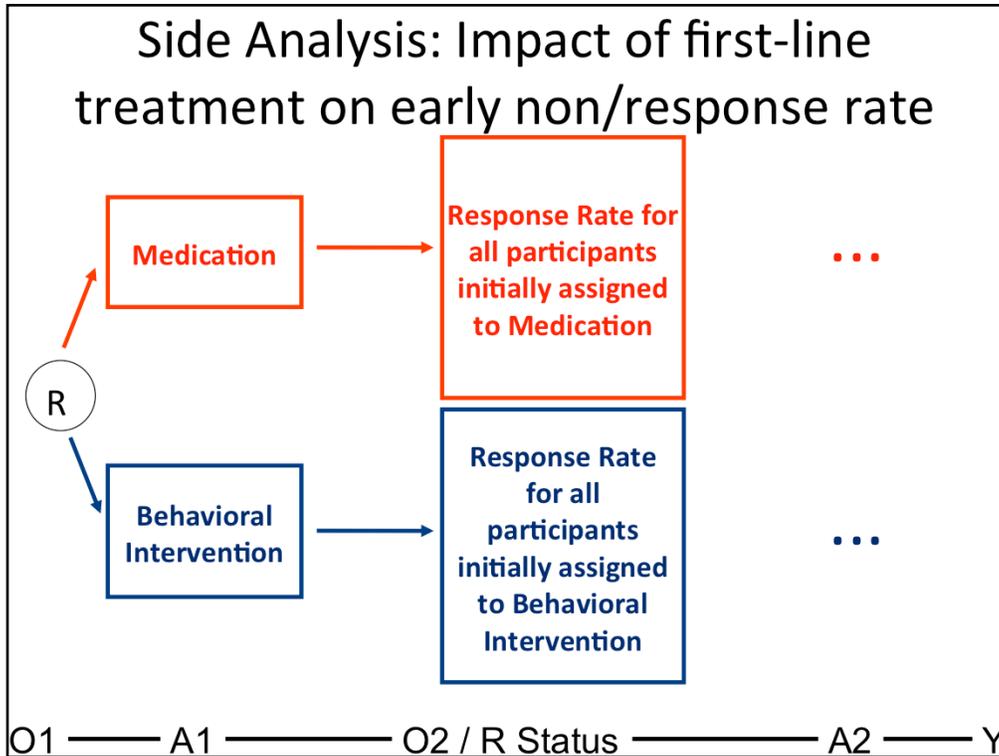
Contrast Estimate Results

Label	Estimate	95% Conf Limits		P-value
		Lower	Upper	
Mean Y under BMOD	3.0459	2.7859	3.3059	<.0001
Mean Y under MED	2.8608	2.6008	3.1208	<.0001
Between groups diff	0.1851	-0.1849	0.5551	0.3269

(SE = standard err)(0.1889)

In this simulated data set/experiment, it is slightly better to begin with BMOD (vs MED) in terms of school performance at end of study, but not statistically significant (p-value = 0.33).

This analysis is with simulated data.



This is NOT a primary aim. But useful nonetheless.

Note that this analysis is less useful in terms of building adaptive treatment strategy because this outcome does not incorporate the effects of future/second-line treatments (second-line treatments haven't been offered yet!)

Therefore, this is not a typical primary question in SMARTs. Rather, this is the “acute effect” first-line treatment (in terms of early response rate outcome). It is nonetheless interesting and you will want to examine this in your data to see what treatment would be recommended if we based our choice of best first-line treatment in terms of the early non/response outcome.

We do this here for completeness to help put the results of our data analysis in further context.

Side analysis: SAS code and results for “acute effect” of first-line treatment

```
proc freq data=dat1;
  table a1*r / chisq nocol nopercnt;
run;
```

Frequency,				
Row Pct	, R = 0,	R = 1,	Total	

A1 = -1	, 47	, 28	, 75	
MED	, 62.67	, 37.33		

A1 = 1	, 52	, 23	, 75	
BMOD	, 69.33	, 30.67		

	99	51	150	

**In terms of early non/
response rate, initial
MED is slightly
better (but NS) than
initial BMOD by 7%
(p-value = 0.39) .**

This analysis is with simulated data.

This analysis is on Page 3 of your SAS code Word document.

Primary Aims Part I, Outline

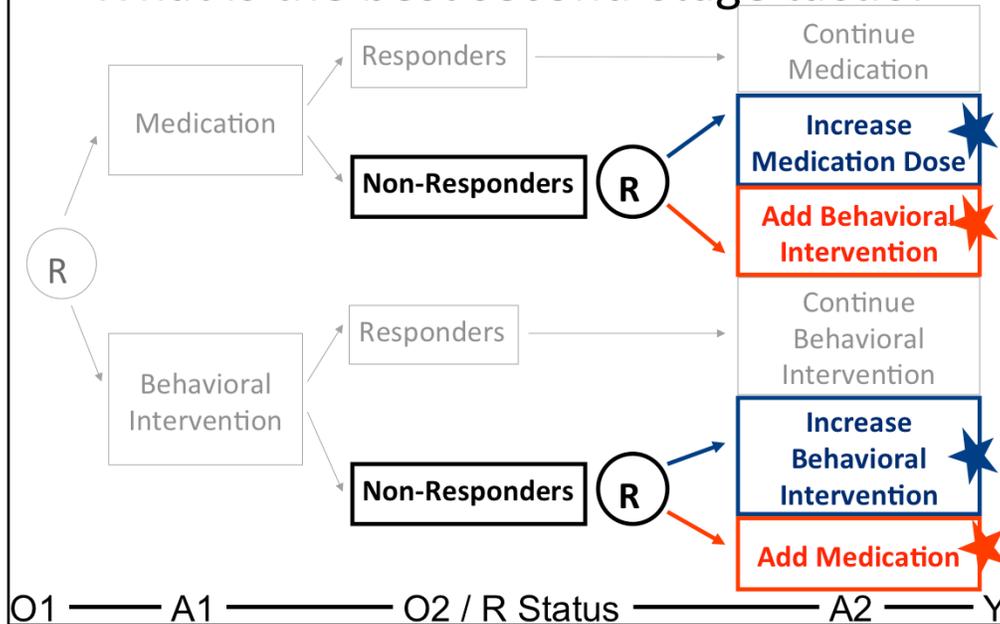
- Review the *Adaptive Interventions for Children with ADHD Study* design (a SMART design)
- Will learn how to analyze two typical primary research questions in a SMART design
 - PI(a): Main effect of initial (first-stage) treatment?
 - PI(b): Comparing second-stage tactics?
- Will prepare for a third primary aim analysis by
 - PI(c): Learning to estimate the mean outcome under each of the embedded ATS (separately) using an easy-to-use weighting approach

Typical Primary Aim 2:
What is the best second-stage tactic?

- Among children who do not respond to (either) first-line treatment, is it better to increase initial treatment or to add a different treatment to the initial treatment?
 - Regardless of history of treatment.

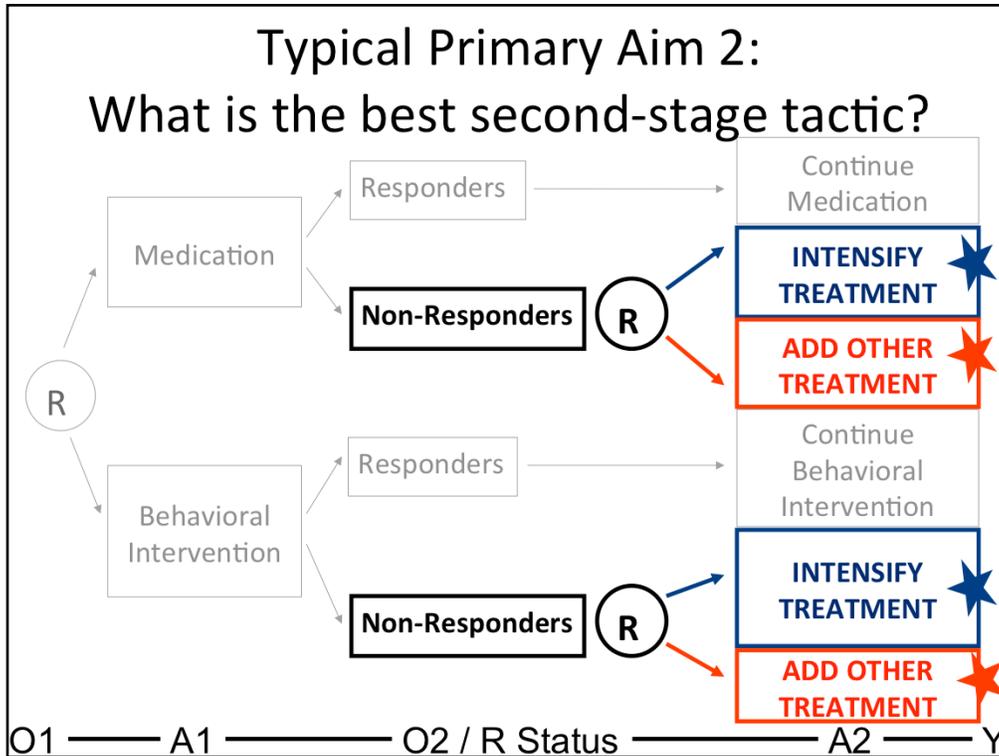
[explained with a picture on the next slide....]

Typical Primary Aim 2: What is the best second-stage tactic?



This is not a comparison of adaptive intervention, per se. Rather it informs the tactical decision often made in clinical practice of whether to add to the treatment with something new versus increase the dosage/intensity of treatment.

Note that this is a comparison of the blue star cells versus the red star cells, pooled over (averaged over) first-line. The pooling leads to more power (i.e., larger sample size for the comparison of tactics) but the pooling does not always make sense. Here it does if we think of it from a mental health services delivery point of view.



This is not a comparison of adaptive intervention, per se. Rather it informs the tactical decision often made in clinical practice of whether to add to the treatment with something new versus increase the dosage/intensity of treatment.

Note that this is a comparison of the blue star cells versus the red star cells, pooled over (averaged over) first-line. The pooling leads to more power (i.e., larger sample size for the comparison of tactics) but the pooling does not always make sense. Here it does if we think of it from a mental health services delivery point of view.

SAS code and results for Primary Aim 2: Second-stage tactic

```

* use only non-responders;
data dat3;
  set dat2; if R=0;
run;
* simple comparison to compare mean Y on add vs intensify (A2);
proc genmod data = dat3;
  model y = a2          o11c o12c o13c o14c o21c o22c;
  estimate 'Mean Y w/INTENSIFY tactic' intercept 1 a2 1;
  estimate 'Mean Y w/ADD TXT tactic'    intercept 1 a2 -1;
  estimate 'Between groups difference'          a2 2;
run;

```

Contrast Estimate Results		95% Conf Limits		
Label	Estimate	Lower	Upper	P-value
Mean Y w/INTENSIFY tactic	2.6064	2.3055	2.9072	<.0001
Mean Y w/ADD TXT tactic	3.1942	2.8904	3.4981	<.0001
Between groups difference	-0.5879	-1.0206	-0.1552	0.0077
(SE = standard error)	(0.2208)			

This analysis is with simulated data.

On average, the tactic of ADDING is better and it is statistically significant, p-value < 0.01.

SAS code and results for Primary Aim 2: Second-stage tactic

```

* use only non-responders;
data dat3;
  set dat2; if R=0;
run;
* simple comparison to compare mean Y on add vs intensify (A2);
proc genmod data = dat3;
  model y = a2          o11c o12c o13c o14c o21c o22c;
  estimate 'Mean Y w/INTENSIFY tactic' intercept 1 a2 1;
  estimate 'Mean Y w/ADD TXT tactic'    intercept 1 a2 -1;
  estimate 'Between groups difference'   a2 2;
run;

```

Contrast Estimate Results

Label	Estimate	95% Conf Limits		P-value
		Lower	Upper	
Mean Y w/INTENSIFY tactic	2.6064	2.3055	2.9072	<.0001
Mean Y w/ADD TXT tactic	3.1942	2.8904	3.4981	<.0001
Between groups difference	-0.5879	-1.0206	-0.1552	0.0077
(SE = standard error)	(0.2208)			

This analysis is with simulated data.

On average, the tactic of ADDING is better and it is statistically significant, p-value < 0.01.

Note: you won't see the line "(SE = standard error) (0.2208)". I added this line myself to the above. But you will see a column with SEs printed on your screen.

Primary Aims Part I, Outline

- Review the *Adaptive Interventions for Children with ADHD Study* design (a SMART design)
- Will learn how to analyze two typical primary research questions in a SMART design
 - PI(a): Main effect of initial (first-stage) treatment?
 - PI(b): Comparing second-stage tactics? treatments?
- Will prepare for a third primary aim analysis by
 - PI(c): Learning to estimate the mean outcome under each of the embedded ATS (separately) using an easy-to-use weighting approach

Typical Primary Aim 3:

Best of two adaptive interventions?

- In the next module, we will learn how to answer the following question: Which is the best of the following two “design-embedded” ATs?

First treat with medication, then

- **If respond, then continue treating with medication**
- **If non-response, then add behavioral intervention**

versus

First treat with behavioral intervention, then

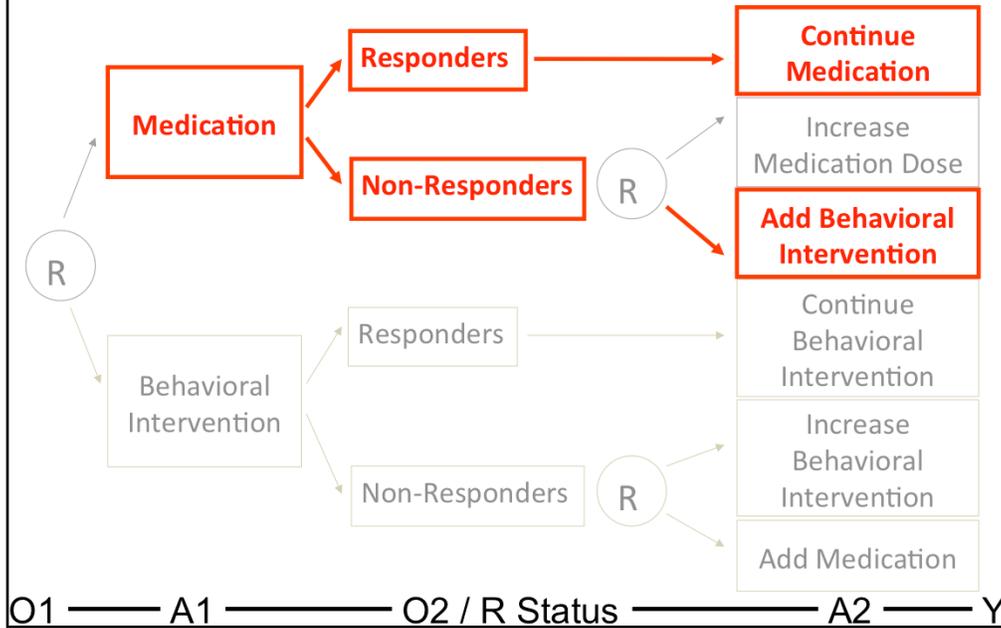
- **If response, then continue behavioral intervention**
- **If non-response, then add medication**

This primary aim is a comparison of 2 adaptive intervention that begin with *different* first line treatment.

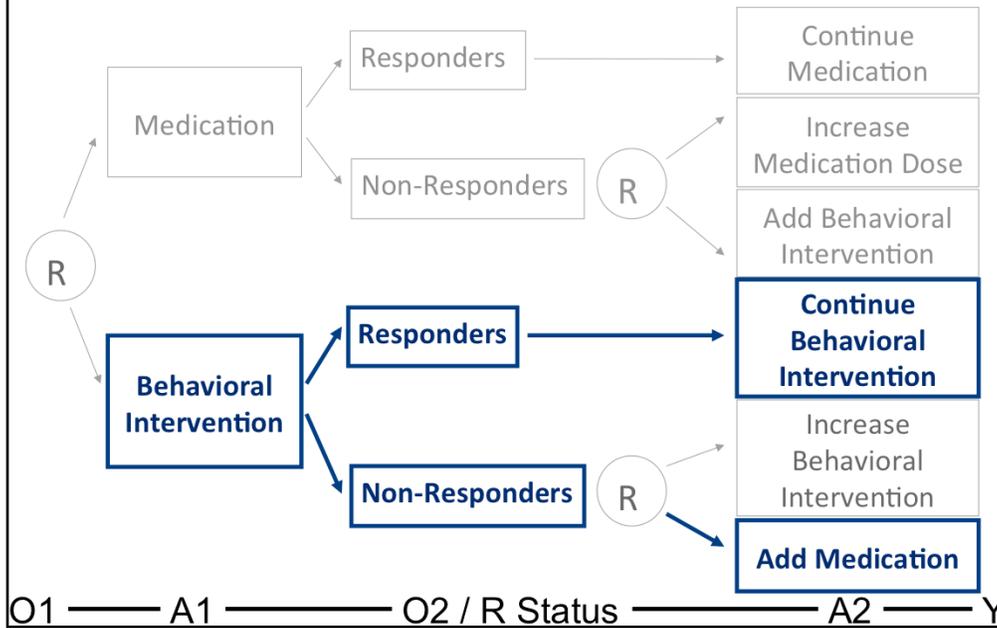
It is a comparison of two decision rules (notice the if/then).

One could also do all remaining pair-wise comparisons between the 4 embedded ATs. Here we chose 1 pair for illustration.

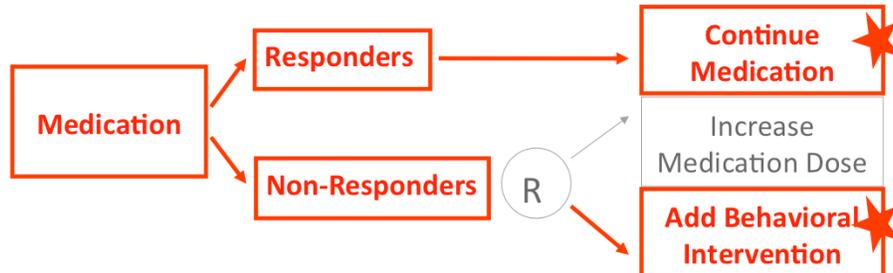
This is a comparison of mean outcome had population followed (MED, BMOD) ATS vs...



...versus the mean outcome had the population followed the (BMOD, MED) ATS



Before learning how to analyze this, we first learn how to obtain mean outcome under (MED, BMOD)



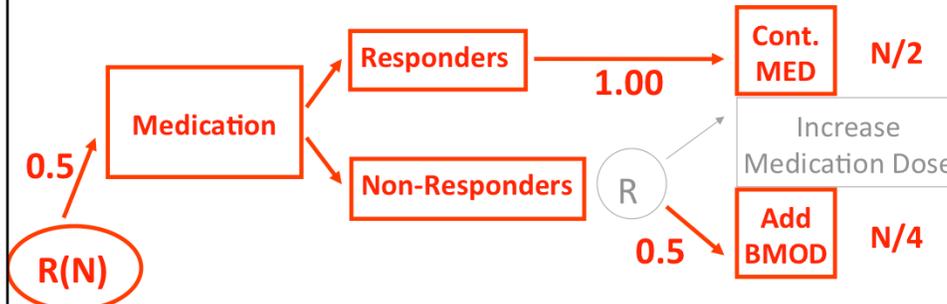
It turns out that we cannot just take the mean outcome for all subjects who ended up in the “Continue Medication” and “Add Behavioral Intervention” boxes! Why?

Why can't we just compare the mean of the sample of Ss who followed red vs mean for those who followed blue?

There is imbalance in the responders and non-responders who followed the red ATS.

For example, let's first consider estimating the mean outcome had all participants followed the red ATS . The issue is...[next slide]

There is imbalance in the non/responding participants following the red ATS...



...because, by design,

- Responders to MED had a $0.5 = 1/2$ chance of having had followed the red ATS, whereas
- Non-responders to MED only had a $0.5 \times 0.5 = 0.25 = 1/4$ chance of having had followed the red ATS

This picture is just heuristic. There are actually $R \cdot N/2$ in the top red cell and $(1-R) \cdot N/4$ in the bottom red cell.

Another way to say this: Responders are over-represented in the data BY DESIGN.

So to estimate mean school performance had all participants followed the red ATS:



- Assign $W = \text{weight} = 2$ to responders to MED
- Assign $W = \text{weight} = 4$ to non-responders to MED
- This “balances out” the responders and non-responders. Then we take W -weighted mean of sample who ended up in the 2 boxes.

So we can just take a weighted mean (with weights define as above) of the outcomes for those participants falling into the 2 red boxes above.

In the next slides we show how to do something equivalent to this using a regression approach.

SAS code to estimate mean outcome had all participants followed (MED, BMOD) ATS

```
* create indicator and assign weights;
data dat5; set dat2;
  Z1=-1;
  if A1=-1 and R=1 then Z1=1;
  if A1=-1 and R=0 and A2=-1 then Z1=1;
  W=4*R + 2*(1-R);
run;
* run W-weighted regression Y = b0 + b1*z1 + e;
* b0 + b1 will represent the mean outcome under red ATS;
proc genmod data = dat5;
  class id;
  model y = z1;
  scwgt w;
  repeated subject = id / type = ind;
  estimate 'Mean Y under red ATS' intercept 1 z1 1;
run;
```

This analysis is with simulated data.

Request robust standard errors: Why? Weights depend on responder status, which is unknown ahead of time.

Instead of a regression, you can also calculate the W-weighted mean outcomes for all participants following the red ATS.

Robust standard errors to account for the sampling error in the “estimation” of the weights. What this really means is we don’t know ahead of time how many responders and non-responders there will be, so the weights are unknown ahead of time. i.e., they are estimated. Another way to say this, is we will not know ahead of time, how many participants get a weight of 2 versus a weight of 4. The standard errors need to account for this uncertainty, and the robust standard errors help us do this.

Results: Estimate of mean outcome had population followed (MED,BMOD) ATS

Analysis Of GEE Parameter Estimates

	Parameter Estimate	SError	P-value
Intercept	2.9153	0.1084	<.0001
Z1	-0.0504	0.1084	0.6417

Contrast Estimate Results

	95% Conf Limits			
	Estimate	Lower	Upper	SError
Mean Y under the red ATS	2.8649	2.5305	3.1992	0.1706

This analysis is with simulated data.

Results: Estimate of mean outcome had population followed (MED,BMOD) ATS

Analysis Of GEE Parameter Estimates

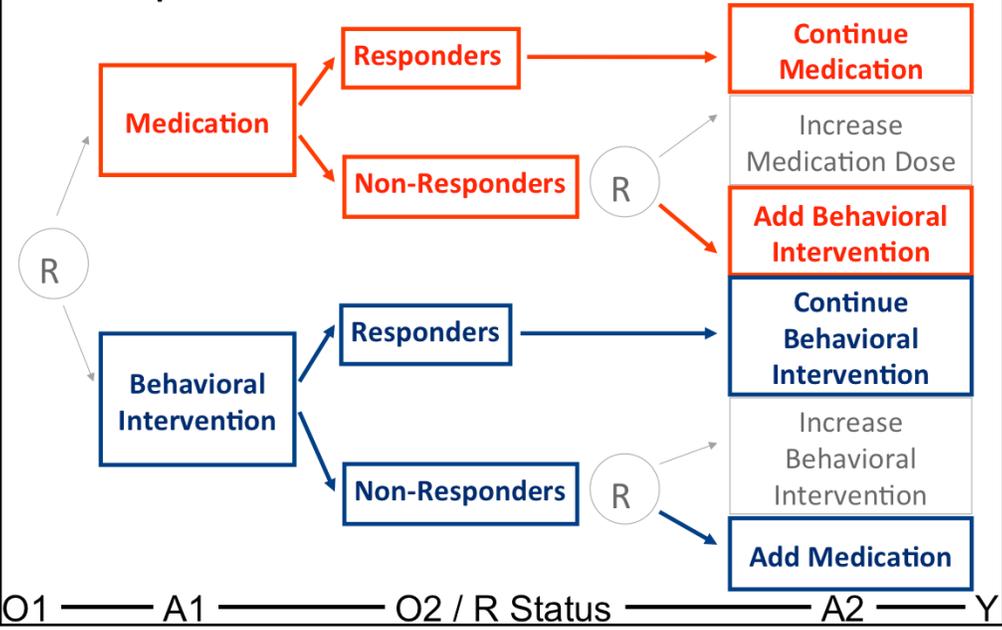
	Parameter Estimate	SError	P-value
Intercept	2.9153	0.1084	<.0001
Z1	-0.0504	0.1084	0.6417

Contrast Estimate Results

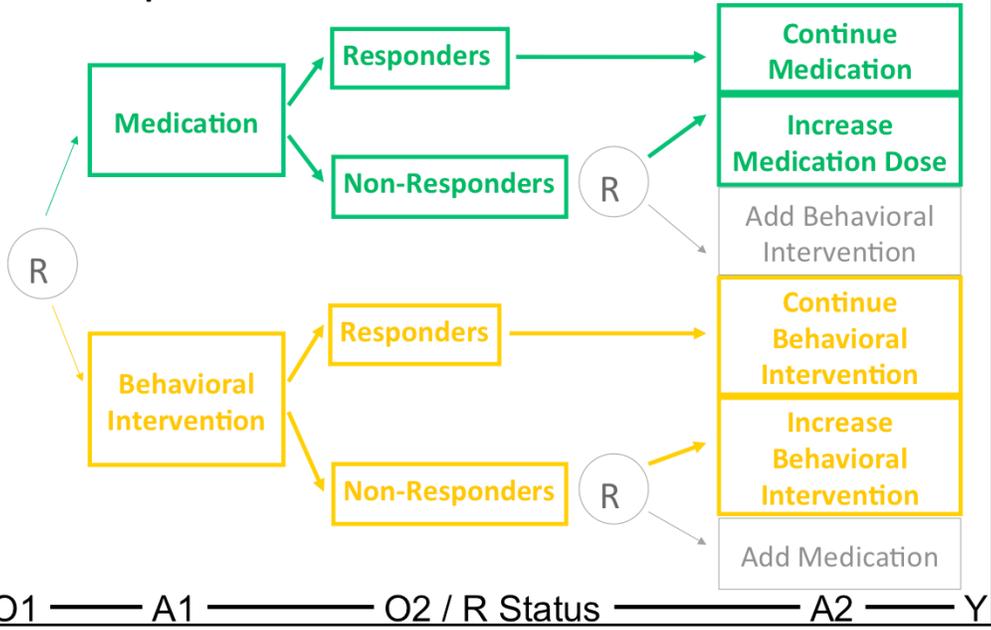
	95% Conf Limits			
	Estimate	Lower	Upper	SError
Mean Y under the red ATS	2.8649	2.5305	3.1992	0.1706

This analysis is with simulated data.

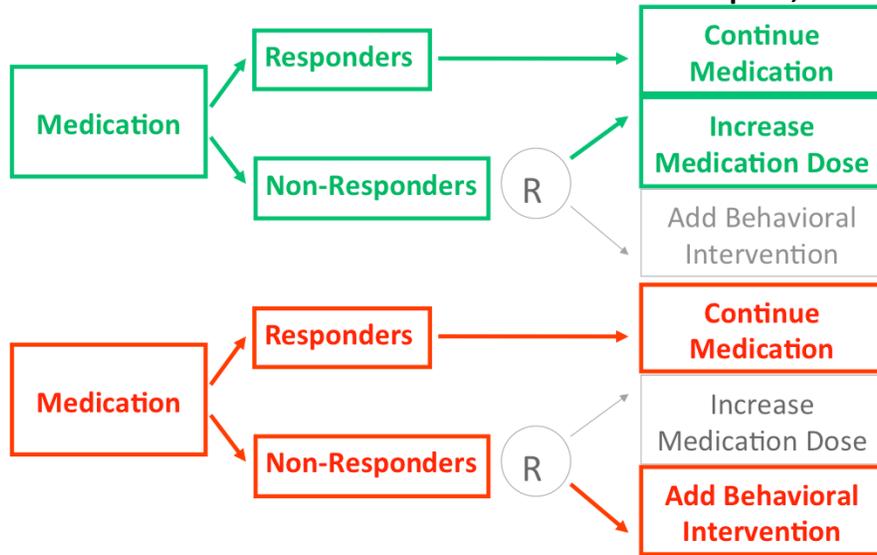
What about a regression that allows comparison of mean under all four AIs?



What about a regression that allows comparison of mean under all four AIs?



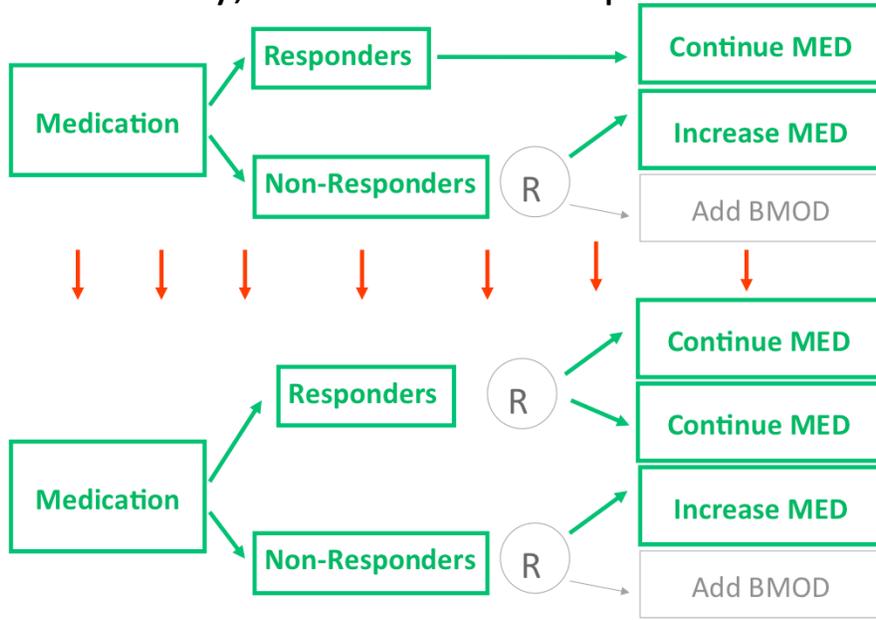
Note that all responders are consistent with 2 of the embedded AIs. For example, ...



So, since all responders are consistent with 2 of the embedded AIs, we...

- We just need to “trick” or “explain” this to SAS
- Do this by replicating responders:
 - Create 2 observations for each responder
 - We assign $\frac{1}{2}$ of them $A_2=1$, the other $\frac{1}{2}$ $A_2=-1$
 - As before, assign $W=2$ to responders and $W=4$ to non-responders
- Robust standard errors take care of the fact that we are “re-using” the responders. No cheating here!

Pictorially, what does the replication do?



SAS code for replication-and-weighting to compare means under all four AIs

```
data dat9; set dat2;
  * define weights and create responders replicates
  * (with equal "probability of getting A2");
  if R=1 then do;
    ob = 1; A2 = -1; weight = 2; output;
    ob = 2; A2 = 1; weight = 2; output;
  end;
  else if R=0 then do;
    ob = 1; weight = 4; output;
  end;
run;
```

This analysis is with simulated data.

Basically we require an extra step to replicate observations (i.e., rows in the data set) of responders, such that instead of one observation per responder, there are 2 observations per responder (one with $A2=1$ and the other with $A2=-1$).

The working intuition is that since a responder's treatment is consistent with this person having been assigned either of two AIs, then we need use each responder's data twice. The first time to estimate the mean for the first AI and the second time to estimate the mean for the second AI.

After replication-and-weighting, the SAS code for the weighted regression to estimate mean under all four AIs is easy!

```
proc genmod data = dat9;
  class id;
  model y = a1 a2 a1*a2;
  scwgt weight;
  repeated subject = id / type = ind;
  estimate 'Mean Y under red    AI' int 1 a1 -1 a2 -1 a1*a2 1;
  estimate 'Mean Y under blue   AI' int 1 a1  1 a2 -1 a1*a2 -1;
  estimate 'Mean Y under green  AI' int 1 a1 -1 a2  1 a1*a2 -1;
  estimate 'Mean Y under orange AI' int 1 a1  1 a2  1 a1*a2  1;
  estimate '    Diff:  red - blue' int 0 a1 -2 a2  0 a1*a2  0;
  estimate '    Diff: orange - blue' int 0 a1  0 a2  2 a1*a2  2;
  estimate '    Diff:  green - blue' int 0 a1 -2 a2  2 a1*a2  0;
  * etc...;
run;
```

Why only four parameters? Because there are only 4 means in total that we wish to estimate.

This analysis is with simulated data.

Results: replication-and-weighting to estimate mean outcome under all 4 AIs

Contrast Estimate Results

			95% Conf Limits		
		Estimate	Lower	Upper	P-value
Mean Y under red	AI	2.8649	2.5305	3.1992	<0.0001
Mean Y under blue	AI	3.5067	3.1643	3.8490	<0.0001
Mean Y under green	AI	2.7895	2.4644	3.1145	<0.0001
Mean Y under orange	AI	2.6533	2.2515	3.0552	<0.0001
Diff:	red - blue	-0.6418	-1.1203	-0.1633	0.0086
etc...					
etc...					

NOTE: We get the exact same results as before when we compared red vs blue, but now we can simultaneously make inference for all the comparisons.

This analysis is with simulated data.

Replication-and-weighting to estimate outcome under all 4 AIs **with more power**

```
proc genmod data = dat7;
  class id;
  model y = a1 a2 a1*a2 o12c o14c;
  scwgt weight;
  repeated subject = id / type = ind;
  estimate 'Mean Y under red    AI' int 1 a1 -1 a2 -1 a1*a2 1;
  estimate 'Mean Y under blue   AI' int 1 a1  1 a2 -1 a1*a2 -1;
  estimate 'Mean Y under green  AI' int 1 a1 -1 a2  1 a1*a2 -1;
  estimate 'Mean Y under orange AI' int 1 a1  1 a2  1 a1*a2  1;
  estimate '    Diff:  red - blue'      a1 -2 a2  0 a1*a2  2;
  estimate '    Diff: orange - blue' int 0 a1  0 a2  2 a1*a2  2;
  estimate '    Diff:  green - blue' int 0 a1 -2 a2  2 a1*a2  0;
  * etc...;
run;
```

Improve efficiency: Adjusting for baseline covariates that are associated with outcome leads to more efficient estimates (lower standard error = more power = smaller p-val).

This analysis is with simulated data.

Results: **more powerful** wtd. Regression to estimate mean outcome under all 4 AIs

Improved efficiency:
 Adjusting for baseline covariates resulted in tighter confidence intervals. Point estimates remained about the same, as expected.

		Contrast Estimate Results			
		Estimate	95% Conf Limits		P-value
			Lower	Upper	
Mean Y under red	AI	2.8801	2.5869	3.1733	<0.0001
Mean Y under blue	AI	3.3854	3.0689	3.7018	<0.0001
Mean Y under green	AI	2.8149	2.5163	3.1135	<0.0001
Mean Y under orange	AI	2.7338	2.3596	3.1081	<0.0001
Diff: red - blue		-0.5053	-0.9401	-0.0704	0.0228
etc...					
etc...					

This analysis is with simulated data.

Results: **more powerful** wtd. Regression to estimate mean outcome under all 4 AIs

Improved efficiency:
 Adjusting for baseline covariates resulted in tighter confidence intervals. Point estimates remained about the same, as expected.

		Contrast Estimate Results			
		Estimate	95% Conf Limits		P-value
			Lower	Upper	
Mean Y under red	AI	2.8801	2.5869	3.1733	<0.0001
Mean Y under blue	AI	3.3854	3.0689	3.7018	<0.0001
Mean Y under green	AI	2.8149	2.5163	3.1135	<0.0001
Mean Y under orange	AI	2.7338	2.3596	3.1081	<0.0001
Diff: red - blue		-0.5053	-0.9401	-0.0704	0.0228
etc...					
etc...					

This analysis is with simulated data.

Citations

- Murphy, S. A. (2005). An experimental design for the development of adaptive intervention. *Statistics in Medicine*, 24, 455-1481.
- Nahum-Shani, I., Qian, M., Almirall, D., Pelham, W. E., Gnagy, B., Fabiano, G., Waxmonsky, J., Yu, J., & Murphy, S. (2012, accepted). *Experimental design and primary data analysis for informing sequential decision making processes*. Forthcoming, to appear in the journal *Psychological Methods*.
 - Technical Report available at the Methodology Center, PSU