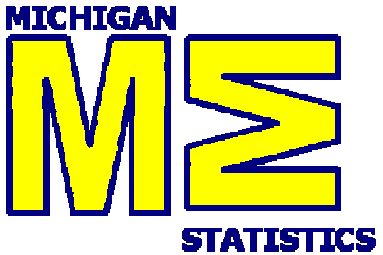
Computer science, adaptive treatment strategies, and SMART

S.A. Murphy



Outline

- Why Adaptive Treatment Strategies?
  - "new" treatment design
- What are SMART experimental designs?
  - "new" clinical trial design
- Trial Design Principles and Analysis
- An Example of a SMART
- Where is the Computer Science?!
  - Secondary data analysis of trial data

Adaptive Treatment Strategies are individually tailored treatments, with treatment type and dosage changing according to patient outcomes. Operationalize clinical practice.

- Brooner et al. (2002, 2007) Treatment of Opioid Addiction
- McKay (2009) Treatment of Substance Use Disorders
- Marlowe et al. (2008) Adaptive Drug Court Program
- Rush et al. (2003) Treatment of Depression

Why Adaptive Treatment Strategies?

- High heterogeneity in response to any one treatment
  - What works for one person may not work for another
  - What works now for a person may not work later
- Improvement often marred by relapse
- Lack of adherence or excessive burden is common
- Intervals during which more intense treatment is required alternate with intervals in which less treatment is sufficient

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

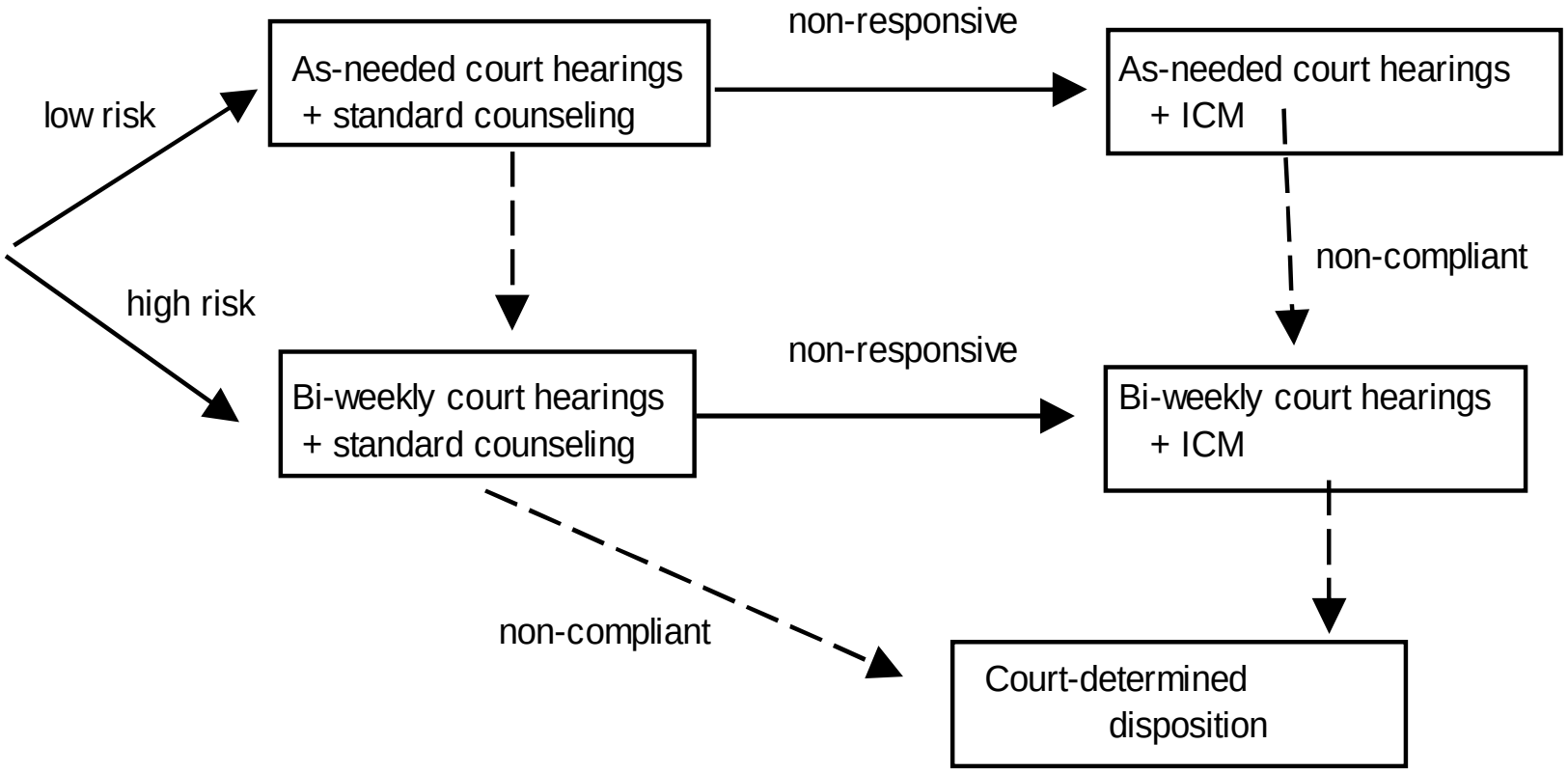
- Treatment incurs side effects and substantial burden, particularly over longer time periods.
- Problems with adherence:
  - Variations of treatment or different delivery mechanisms may increase adherence
  - Excessive treatment may lead to non-adherence
- Treatment is costly (Would like to devote additional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

More is not always better!

Example of an Adaptive Treatment Strategy

- Adaptive Drug Court Program for drug abusing offenders.
- Goal is to minimize recidivism and drug use.
- Marlowe et al. (2008)

Adaptive Drug Court Program



The Big Questions

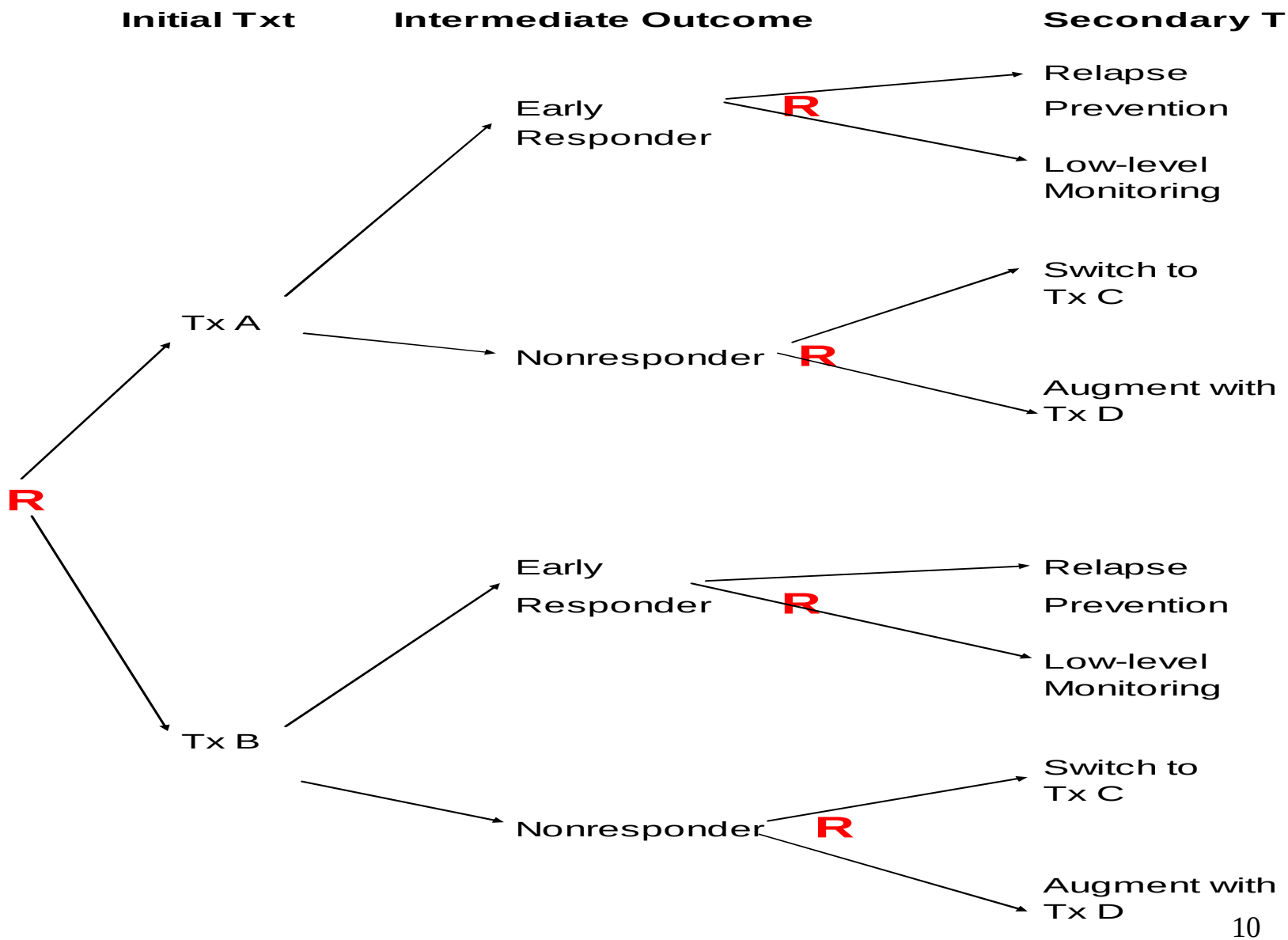
- What is the best sequencing of treatments?
- What is the best timings of alterations in treatments?
- What information do we use to make these decisions? (how do we individualize the sequence of treatments?)

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

These are multi-stage trials. Each participant moves through stages of treatment. Each stage corresponds to a critical decision. A randomization takes place at each critical decision.

Goal of trial is to inform the construction of adaptive treatment strategies.

Sequential Multiple Assignment Randomization



Examples of "SMART" designs:

- CATIE (2001) Treatment of Psychosis in Schizophrenia
- STAR*D (2003) Treatment of Depression
- Pelham (primary analysis) Treatment of ADHD
- Oslin (primary analysis) Treatment of Alcohol Dependence
- Jones (in field) Treatment for Pregnant Women who are Drug Dependent
- Kasari (in field) Treatment of Children with Autism

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 (adherence, etc.); information that might be used to individualize treatment.

SMART Design Principles

•Choose primary hypotheses that are both scientifically important and aid in developing the adaptive treatment strategy.

- •Power trial to address these hypotheses.

•Choose secondary hypotheses that further develop the adaptive treatment strategy and use the randomization to eliminate confounding.

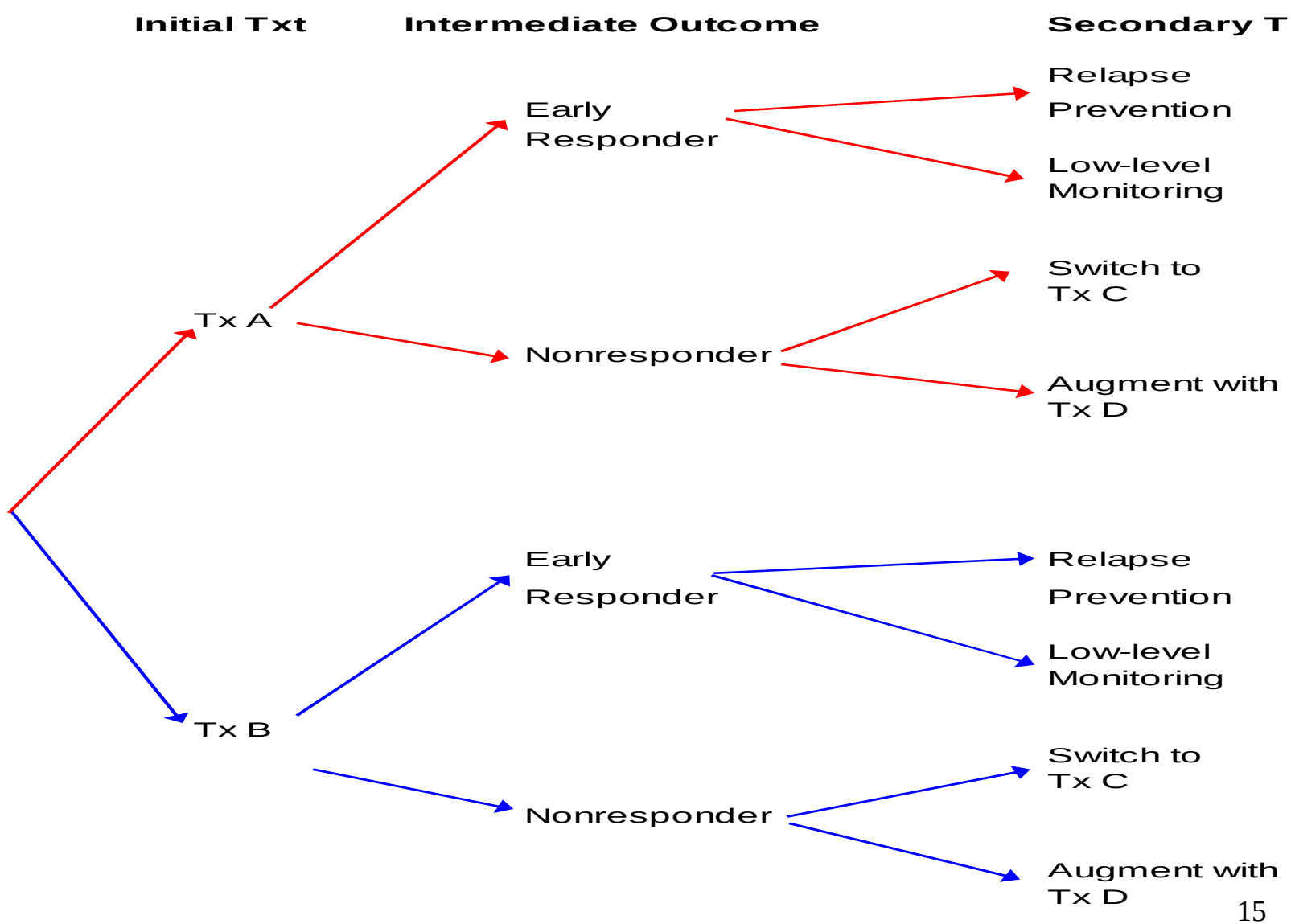
- •Trial is not necessarily powered to address these hypotheses.

SMART Designing Principles: Primary Hypothesis

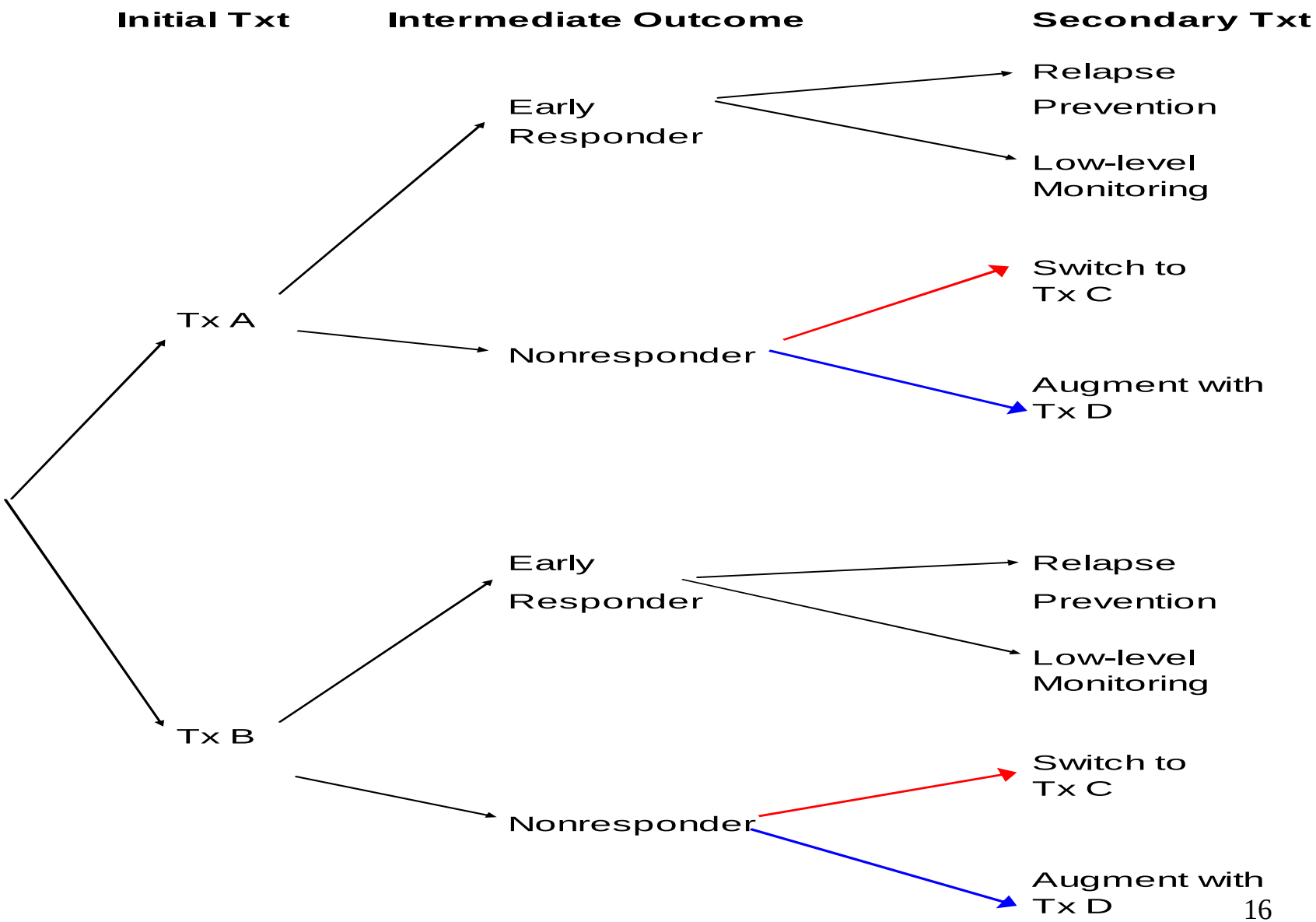
•EXAMPLE 1: (*sample size is highly constrained*): Hypothesize that controlling for the secondary treatments, the initial treatment A results in lower symptoms than the initial treatment B.

•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.

EXAMPLE 1


Initial Txt
Intermediate Outcome
Secondary Txt
Tx A
Early Responder
Relapse Prevention
Low-level Monitoring
Nonresponder
Switch to Tx C
Augment with Tx D
Tx B
Early Responder
Relapse Prevention
Low-level Monitoring
Nonresponder
Switch to Tx C
Augment with Tx D

EXAMPLE 2


Initial Txt
Intermediate Outcome
Secondary Txt
Tx A
Early Responder
Relapse Prevention
Low-level Monitoring
Nonresponder
Switch to Tx C
Augment with Tx D
Tx B
Early Responder
Relapse Prevention
Low-level Monitoring
Nonresponder
Switch to Tx C
Augment with Tx D

SMART Designing Principles: Sample Size Formula

- EXAMPLE 1: (sample size is highly constrained): 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: (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. *Sample size formula is same as a two group comparison of non-responders.*

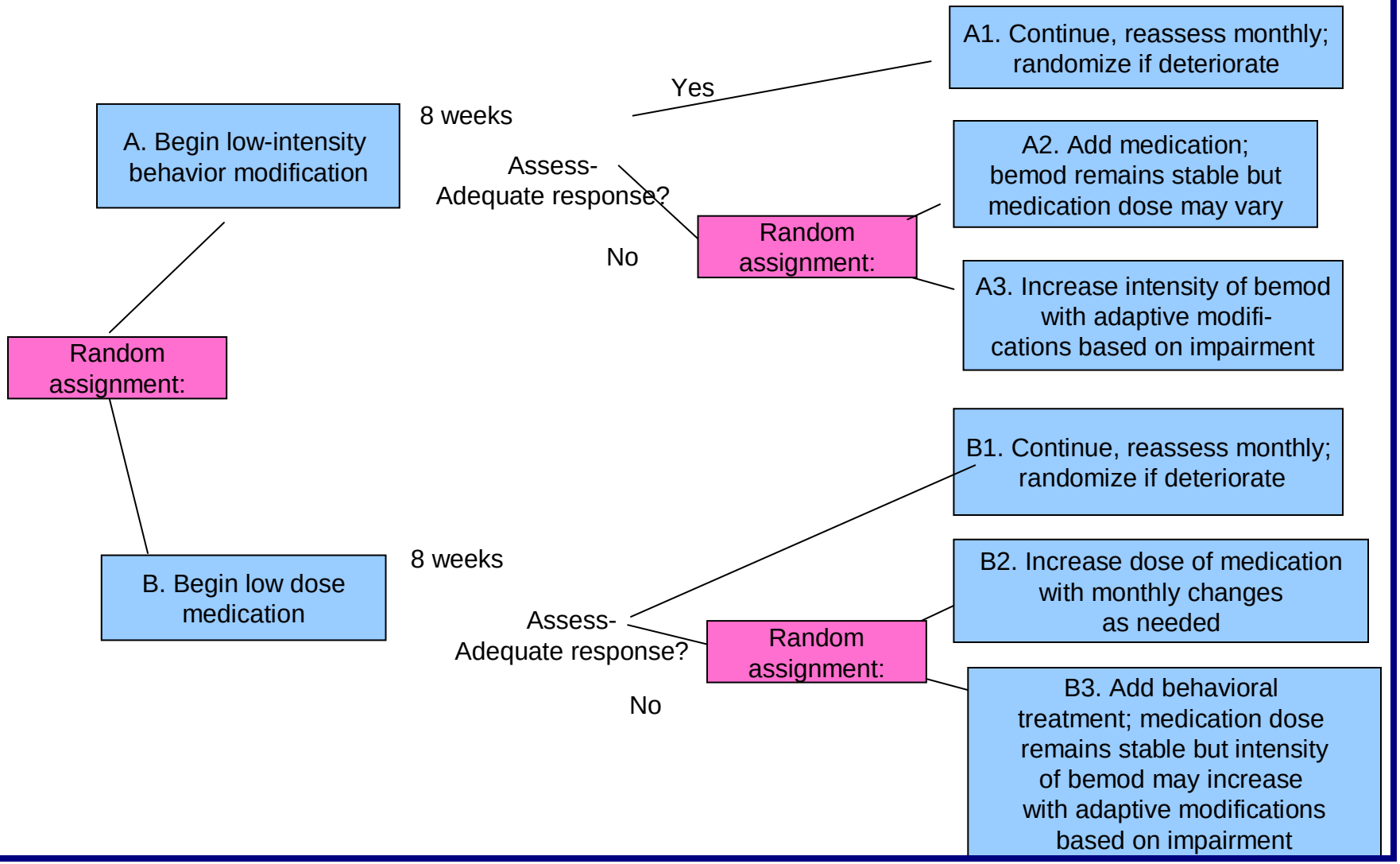
Sample Sizes

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

α = .05, power =1 – β=.85

Example: Pelham ADHD Study



Where is the Computer Science?! Secondary Data Analysis using Q-Learning

Q-Learning is an extension of regression to sequential treatments.

- This regression results in a proposal for an optimal adaptive treatment strategy.
- A subsequent trial would evaluate the proposed adaptive treatment strategy.

Adaptive Treatments for Children with ADHD

- Stage 1 data: (X_1, A_1, R_1)
  - $R_1 = 1$ if responder; $= 0$ if non-responder
  - $A_1 = 1$ if BMOD, $A_1 = -1$ if MED
  - X_1 includes baseline school performance, (Y_0) and a measure of acceptability of medication (S_1)
    - $S_1 = 1$ if medication known to be acceptable; $= 0$, if this is unknown.
- Stage 1 involves all children

Adaptive Treatments for Children with ADHD

- Stage 2 data: (X_2, A_2, Y)
  - Y = end of year school performance
  - A_2=*1* if Enhance, A_2=-*1* if Augment
  - X_2 includes the month of non-response, (M_2) and a measure of adherence in stage 1 (S_2)
    - S_2 =*1* if adherent in stage 1; =*0*, if non-adherent
- Stage 2 involves only children who do not respond in Stage 1 (R_1=*0*).

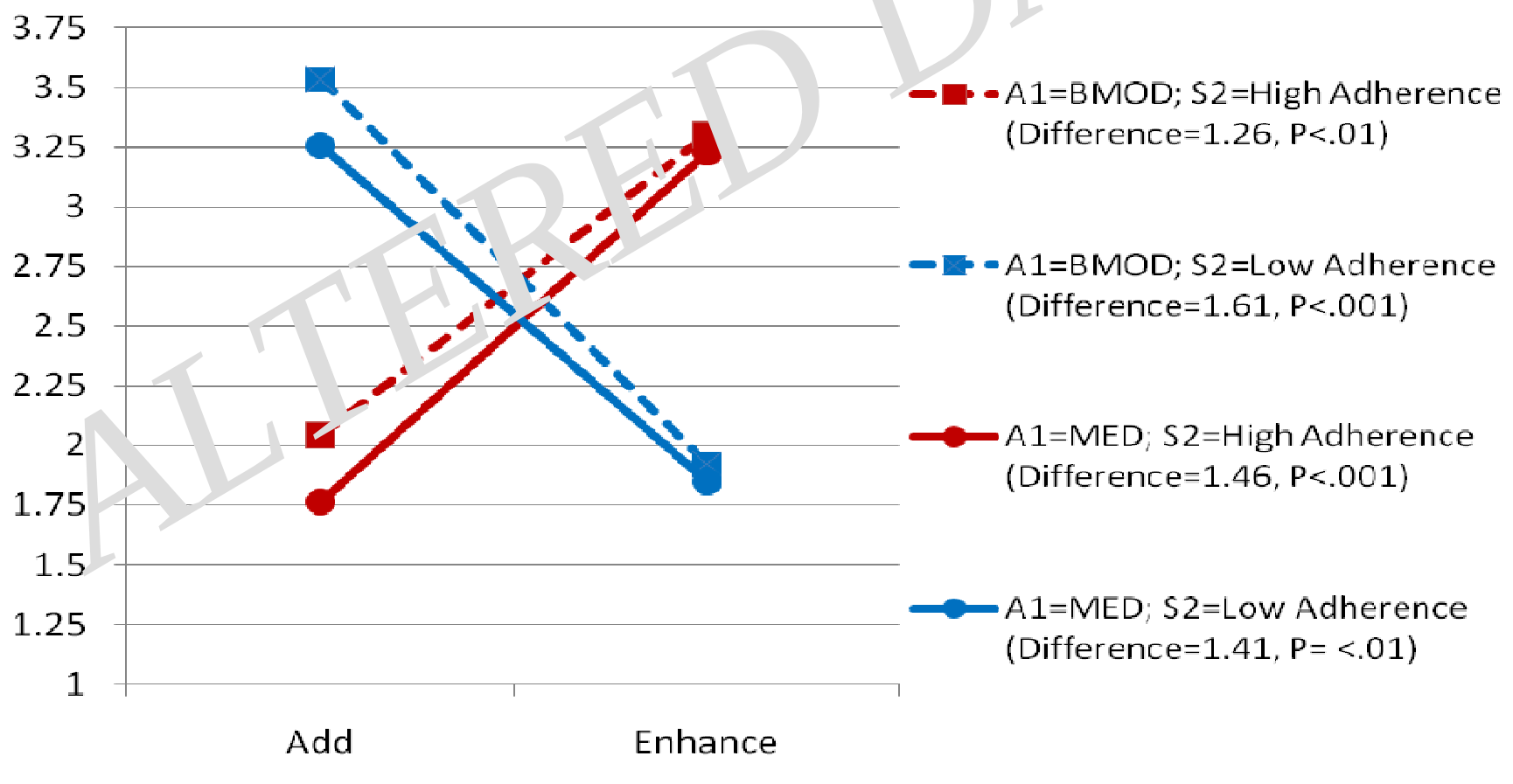
Q-Learning for SMART Studies

- Conduct the regressions in backwards order! E.g. Stage 2 first, then Stage 1.
- Why?
  - Stage 1 dependent variable must be controlled for Stage 2 treatment.
  - Stage 1 dependent variable is a predictor of *Y* under optimal treatment in stage 2.
  - Stage 2 analysis is used to construct Stage 1 dependent variable

Stage 2 Regression for Non-responding Children

- Dependent Variable: Y (end of school year performance)
- Treatment: $A_2=1$ if Enhance, $A_2=-1$ if Augment
- Interactions with Treatment, A_2: stage 1 treatment (A_1) and adherence (S_2)
- Controls: baseline school performance, (Y_0) and acceptability of medication (S_1), month of non-response (M_2), stage 1 treatment (A_1)

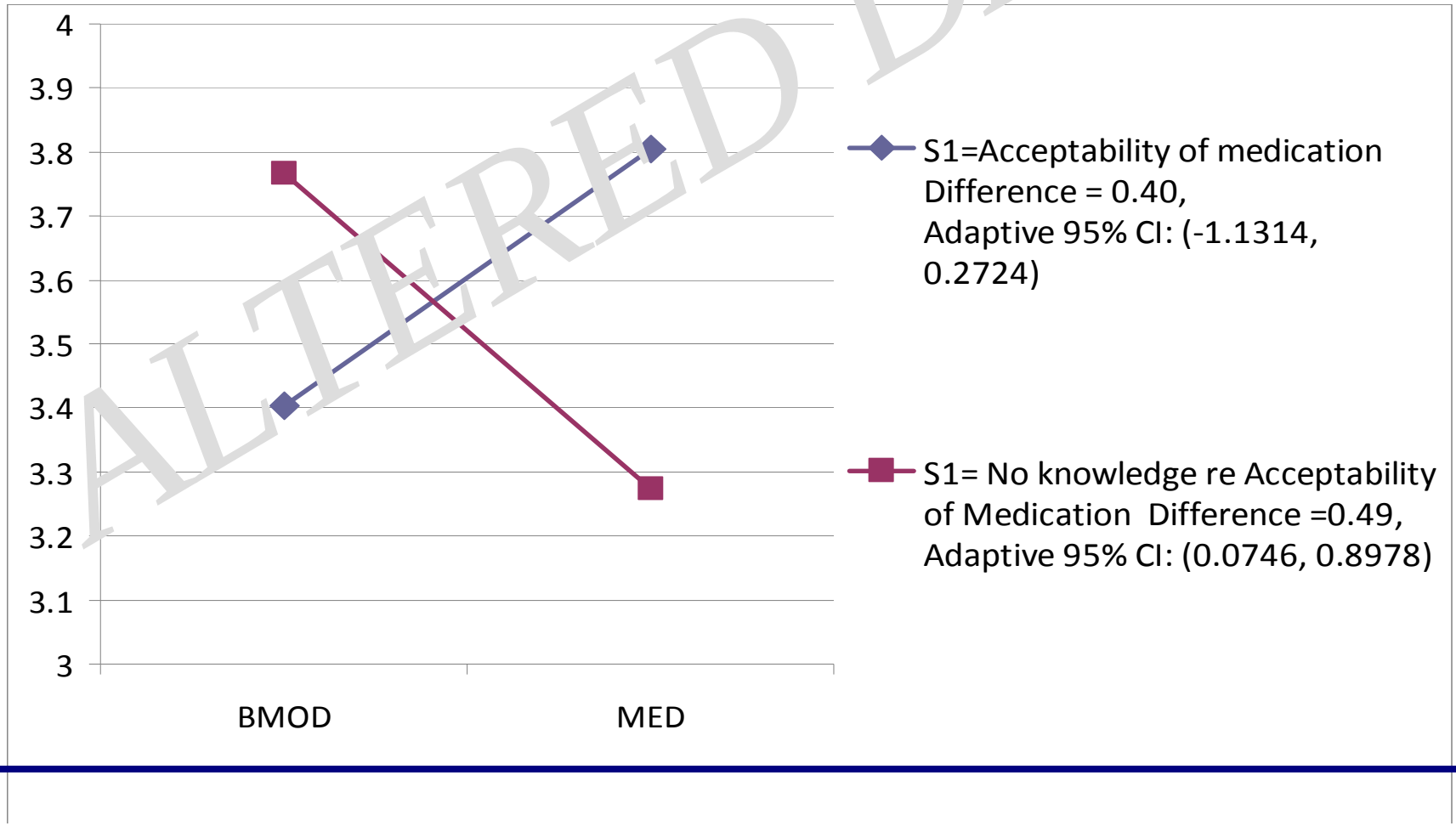
Stage 2 Regression for Non-responding Children



Stage 1 Regression for All Children

- Dependent Variable: $\hat{Y}$ (predicted end of school year performance under optimal stage 2 treatment)
- Treatment: $A_1 = 1$ if BEMOD, $A_1 = -1$ if MED
- Interactions with Treatment, A_1: acceptability of medication (S_1)
- Control: baseline school performance, (Y_0)

Stage 1 Regression for All Children



Adaptive Treatments for Children with ADHD

Adaptive Treatment Strategy Proposal:

- If there is no evidence that medication is highly acceptable begin with BMOD; otherwise select either BMOD or MED.
- If the child is nonresponsive and was non-adherent, augment present treatment; if the child is nonresponse and was adherent, select intensification of current treatment.

Discussion

- Confidence intervals for the regression coefficients in Q-Learning are coming out soon!

- We have a sample size formula that specifies the sample size necessary to detect an adaptive treatment strategy that results in a mean outcome δ standard deviations better than the other strategies with 90% probability (A. Oetting, J. Levy & R. Weiss are collaborators)

- Aside: Non-adherence is an outcome (like side effects) that indicates need to tailor treatment.

This seminar can be found at:

http://www.stat.lsa.umich.edu/~samurphy/seminars/BehInter.06.28.10.pdf

This seminar is based on work with many collaborators some of which are: L. Collins, K. Lynch, J. McKay, D. Oslin, T. Ten Have, I. Nahum-Shani & B. Pelham. Email me with questions or if you would like a copy:

samurphy@umich.edu

Adaptive Treatments for Children with ADHD

- Stage 2 regression for Y:

$$(1, Y_0, S_1, A_1, M_2, S_2)\alpha_2 + A_2(\beta_{21} + A_1\beta_{22} + S_2\beta_{23})$$

- Stage 1 outcome: $R_1 Y + (1 - R_1)\hat{Y}$

$$\hat{Y} = (1, Y_0, S_1, A_1, M_2, S_2)\hat{\alpha}_2 + |\hat{\beta}_{21} + A_1\hat{\beta}_{22} + S_2\hat{\beta}_{23}|$$

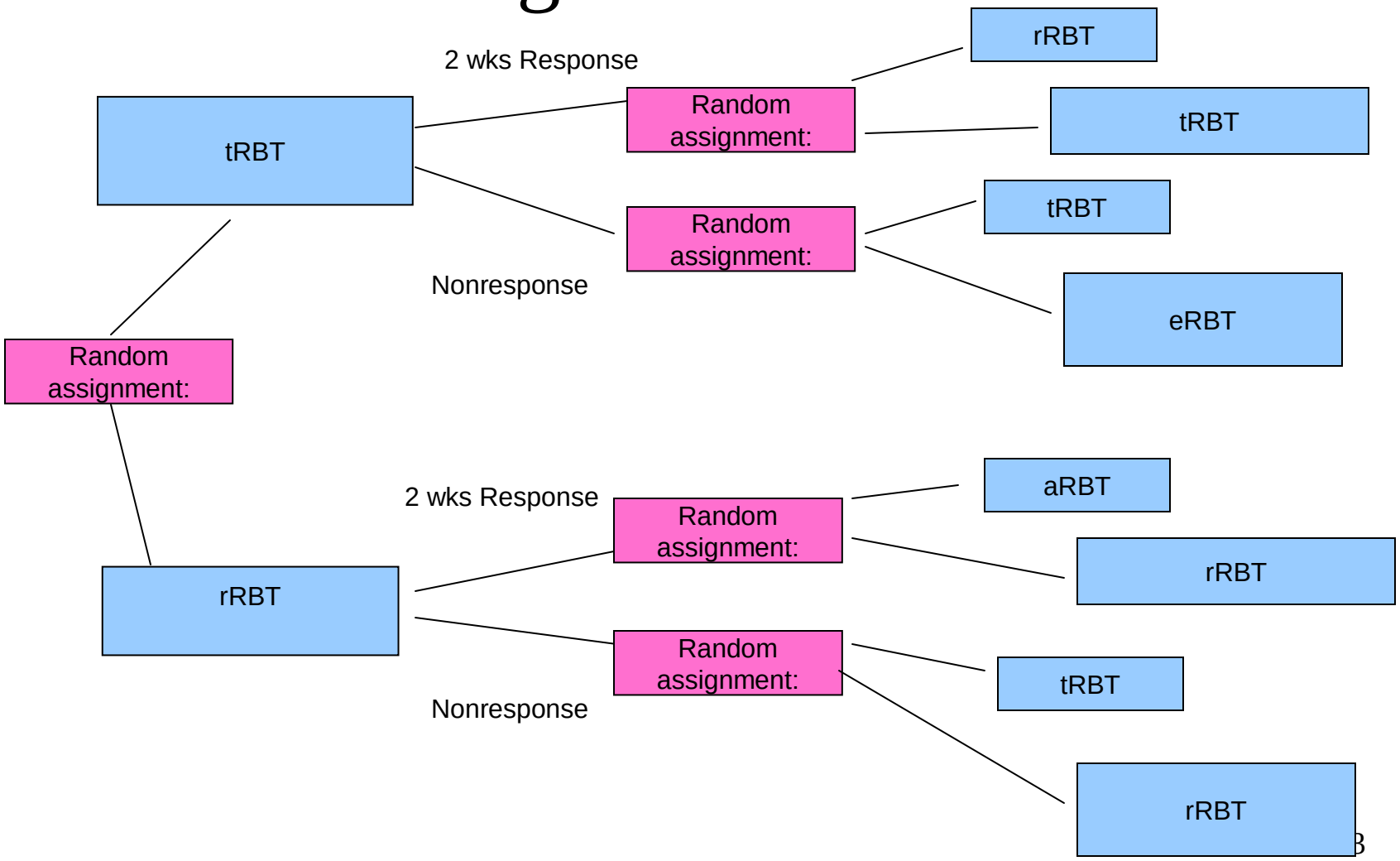
Adaptive Treatments for Children with ADHD

- Stage 1 regression for $\hat{Y}$:

$$(1, Y_0, S_1)\alpha_1 + A_1(\beta_{11} + S_1\beta_{12})$$

- Interesting stage 1 contrast: should the knowledge that medication is highly acceptable, determine the best initial treatment in the sequence?

Jones' Study for Drug-Addicted Pregnant Women



Oslin ExTENd