SMART Study Designs for Developing Interventions

S.A. Murphy

CAPS 10/24/13

50 min. SMART Study Designs for Developing Adaptive Interventions

Outline

- Adaptive Interventions
- SMART Designs
- Trial Design Principles and Analysis
- Exploring Individualization using the "Adaptive Interventions for Children with ADHD" study (W. Pelham, PI).

2

Other names are dynamic treatment regimes, treatment algorithms, stepped care models, expert systems, adaptive treatment strategy, treatment protocols. Structured treatment interruptions in the treatment of AIDS were an early form of an adaptive intervention

Adaptive Interventions are individually tailored sequences of interventions, 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, 2012) Drug Court

•Rush et al. (2003) Treatment of Depression

3

Provide a paradigm whereby we can seek to improve clinical practice which by its nature is adaptive.

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

Most clinical scientists develop the decision rules using trial and error;
developmental and behavioral theories; clinical experience

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

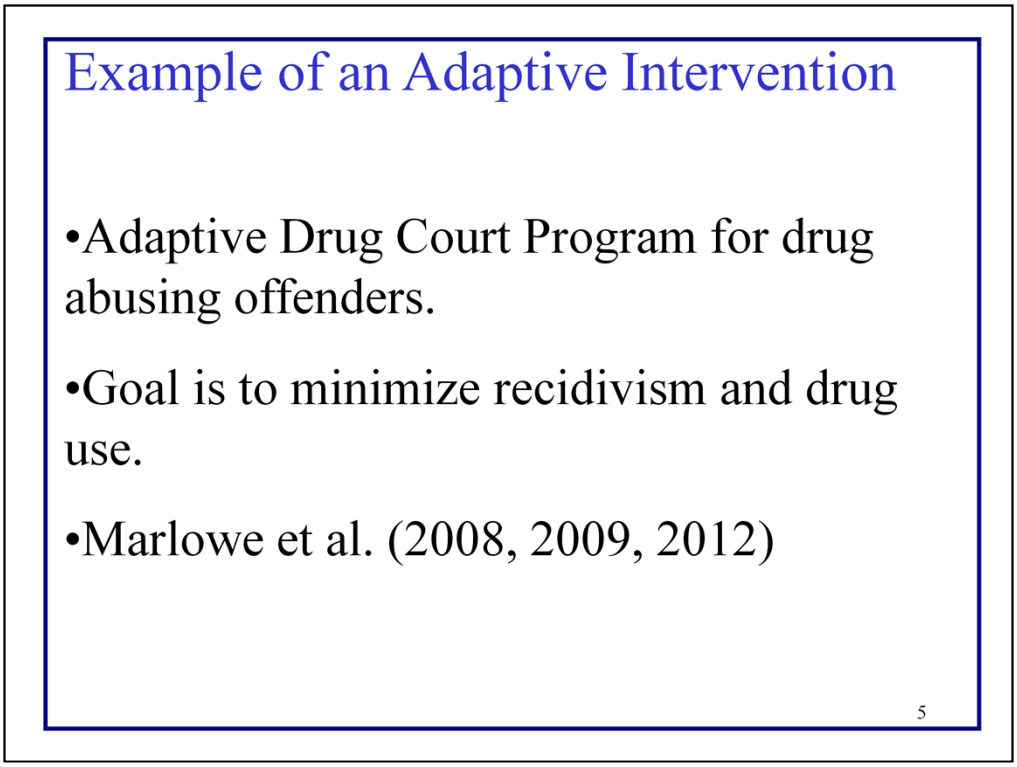
The decision rules used by Brooner et all, Marlowe et al., and McKay are quite detailed, and based on explicit actions by patient, whereas in contrast the Rush et al study (Texas Medication Algorithm Project) appears to be more losely structured; the clinician uses clinical judgment to decide if depression 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.

Why Adaptive Interventions?

- 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 (and relapse is common)
- Lack of adherence or excessive burden is common

4

These are all reasons why we need to plan ahead because we are likely to need to use a sequence of treatments



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.

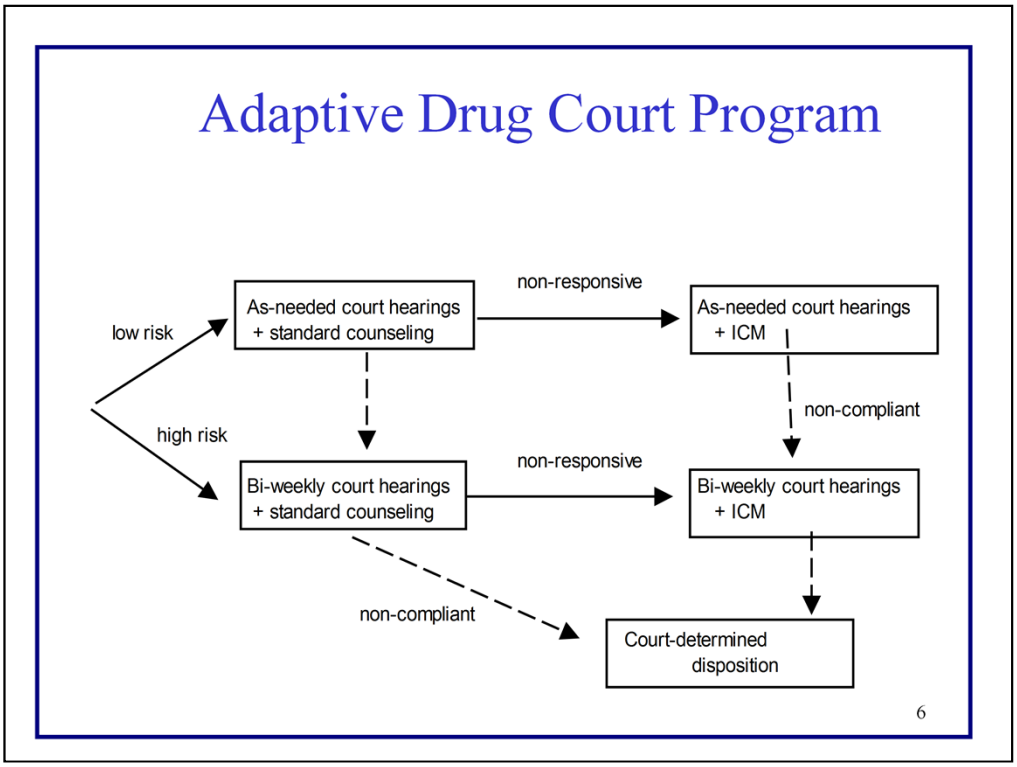
Curr Psychiatry Rep. 2009 Oct;11(5):370-6.

Adaptive Programming Improves Outcomes in Drug Court : An Experimental Trial

Criminal Justice and Behavior 2012 39: 514 Douglas B. Marlowe, David S. Festinger, Karen L. Dugosh,

Kathleen M. Benasutti, Gloria Fox and Jason R. Croft

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 formal drug abuse treatment otherwise low risk

These are assessed monthly:::
Noncompliance: is(1) falls below threshold for attendance in 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.

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

Some Critical Decisions

- 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?)

7

This is related to clinical management of chronic disorders.

Take a broad view of what constitutes therapies: changing intensity, switching medication, augmenting medication, behavioral contingencies, monitoring schedules, motivational therapy, support networks. Also how to combine therapies?

Outline

- Adaptive Interventions
- SMART Designs
- Trial Design Principles and Analysis
- Exploring Individualization using the "Adaptive Interventions for Children with ADHD" study (W. Pelham, PI).

8

SMART Studies

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

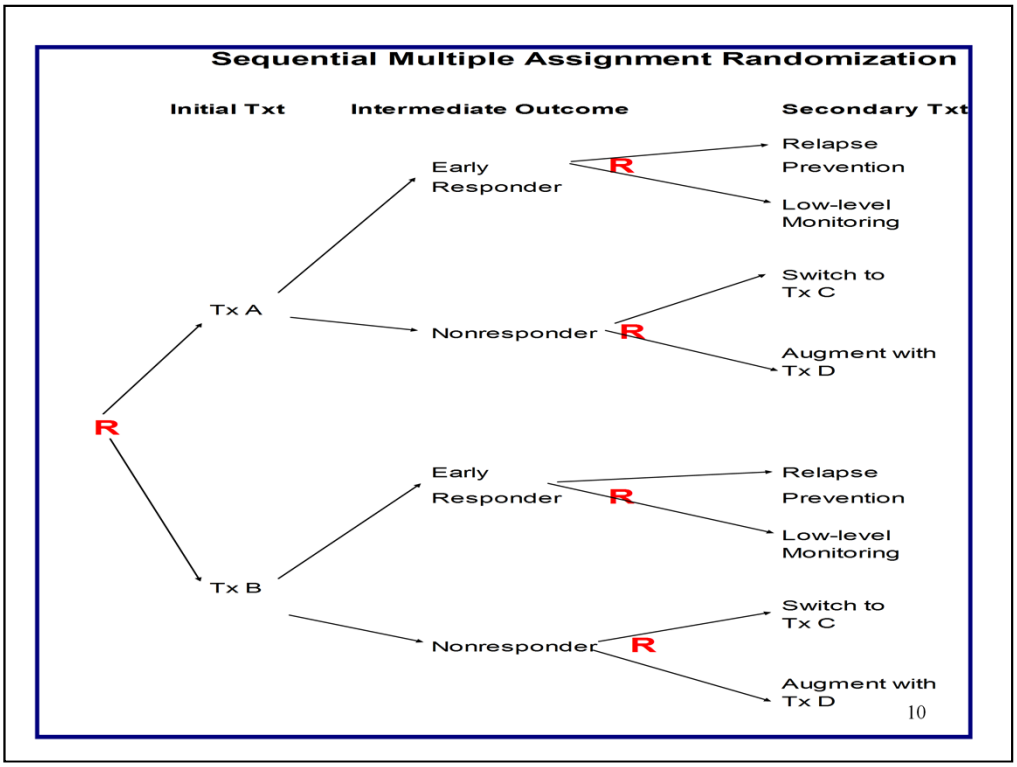
These are multi-stage clinical trials; each participant proceeds through stages of treatment.

Each stage begins with a critical decision and a randomization to treatment takes place at each critical decision.

Goal of trial is to inform the construction of an adaptive intervention.

9

In statistics people may call these multistage trials (the randomization at each stage is assumed) or sequentially randomized trials



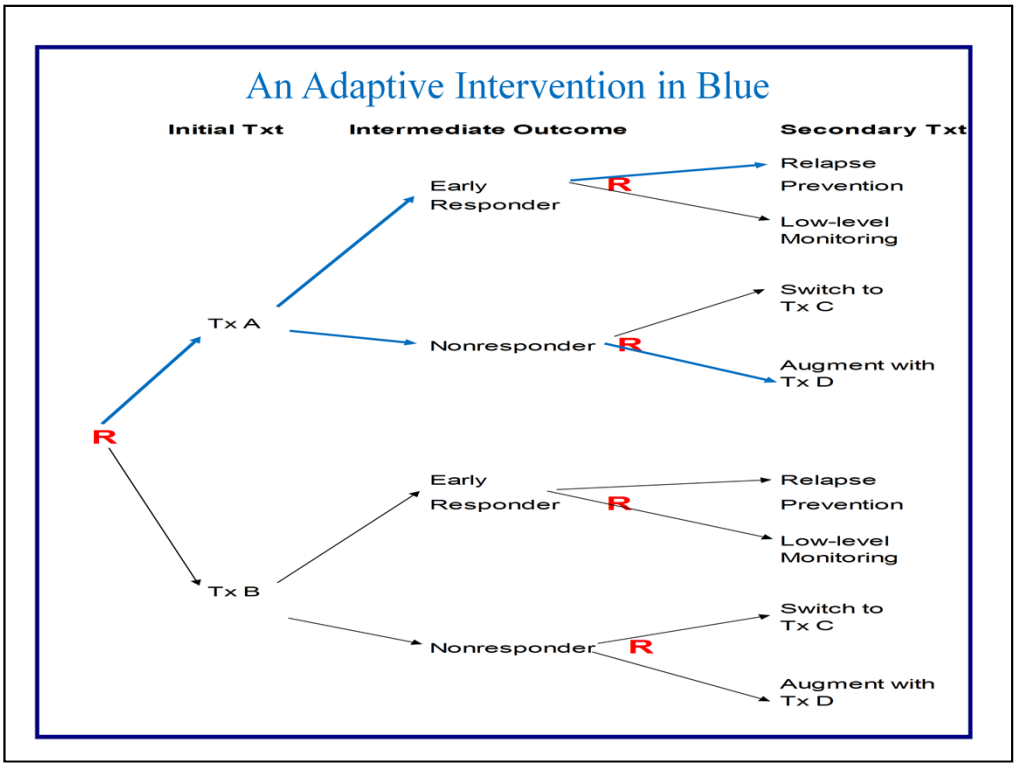
Hypothetical trial: Outcome is not shown but is on far right. The second randomization can take place up front (if you do not want to stratify or block by stage 1 outcomes such as adherence).

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 adaptive intervention is indicated in blue

Alternate Approach to Constructing an Adaptive Intervention

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

Particularly attractive since potential initial treatments 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.

A consequence is that comparing two initial therapies based on a proximal outcome may produce different results from the comparison of two initial therapies when followed by a maintenance therapy and comparing more distal outcomes. Additionally, restricting comparisons to longer term outcomes, a comparison of two initial therapies followed by usual care or no therapy may yield different results from the comparison of two initial therapies when followed by one of several maintenance therapies.

We can expect that in optimized adaptive interventions, the best subsequent therapy will build on the gains achieved by prior therapies and thus these delayed effects should be common.

We want big positive delayed effects. We want profound positive cross-over effects!!!

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 nonresponders 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 responsivity to future treatment; see Thall et al. (2007), Bembom and van der Laan (2007) for an example of the latter in cancer research.

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 enhancement of an 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.

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:

- When evaluating and comparing initial treatments, *in a sequence of treatments*, we need to take into account, e.g. control, the effects of the secondary treatments thus SMART
- Standard single-stage randomized trials may yield information about different populations from SMART trials.

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

Examples of "SMART" designs:

- Pelham (2012) Treatment of ADHD
- Oslin (primary analysis) Treatment of Alcohol Dependence
- Kasari (primary analysis, in field) Treatment of Children with Autism
- McKay (in field) Treatment of Alcohol and Cocaine Dependence

http://methodology.psu.edu/ra/adap-treat-strat/projects

18

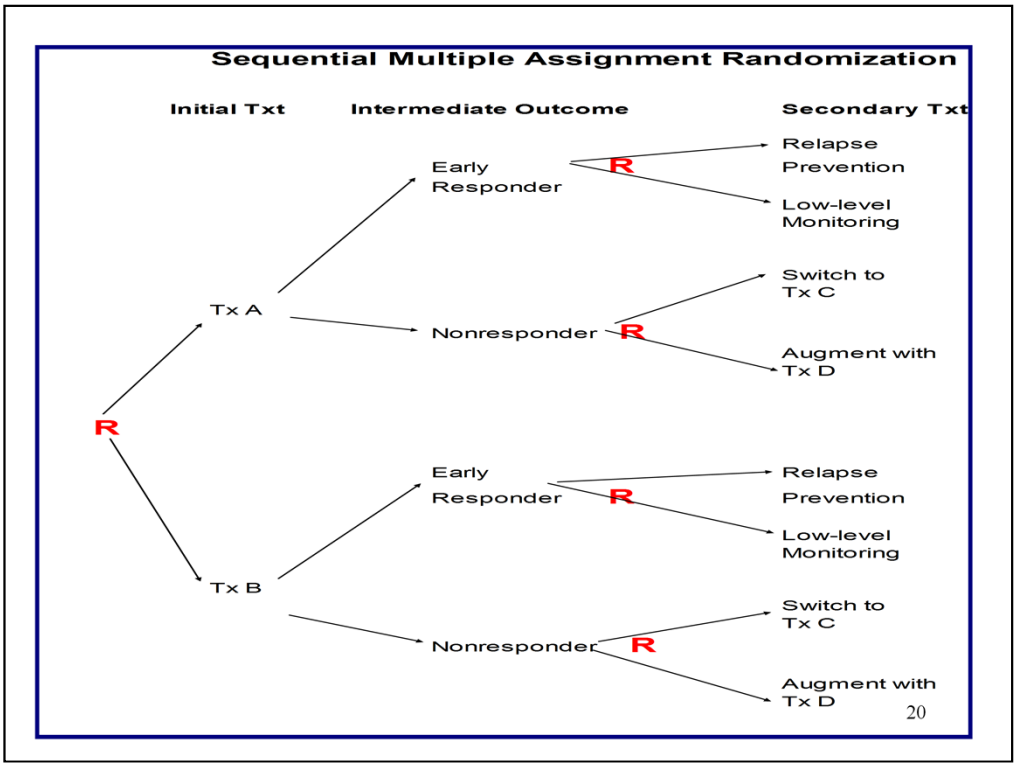
These are intervention development trials. These trials are not confirmatory in the sense of confirming that one adaptive intervention is best.

Other trials in cancer.

Outline

- Adaptive Interventions
- SMART Designs
- Trial Design Principles and Analysis
- Exploring Individualization using the "Adaptive Interventions for Children with ADHD" study (W. Pelham, PI).

19



Hypothetical trial: Outcome is not shown but is on far right. The second randomization can take place up front (if you do not want to stratify or block by stage 1 outcomes such as adherence).

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.

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 subsequent treatment.

21

Note we considered different txt's for the responders as compared to the nonresponders.

In mental illness studies feasibility considerations may force us to use preference in this low dimensional summary.

SMART Design Principles

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

- •Power trial to address these hypotheses.

•Conduct secondary analyses that further develop the adaptive intervention (take advantage of the randomization in eliminating confounding).

22

SMART Designing Principles: Primary Hypothesis

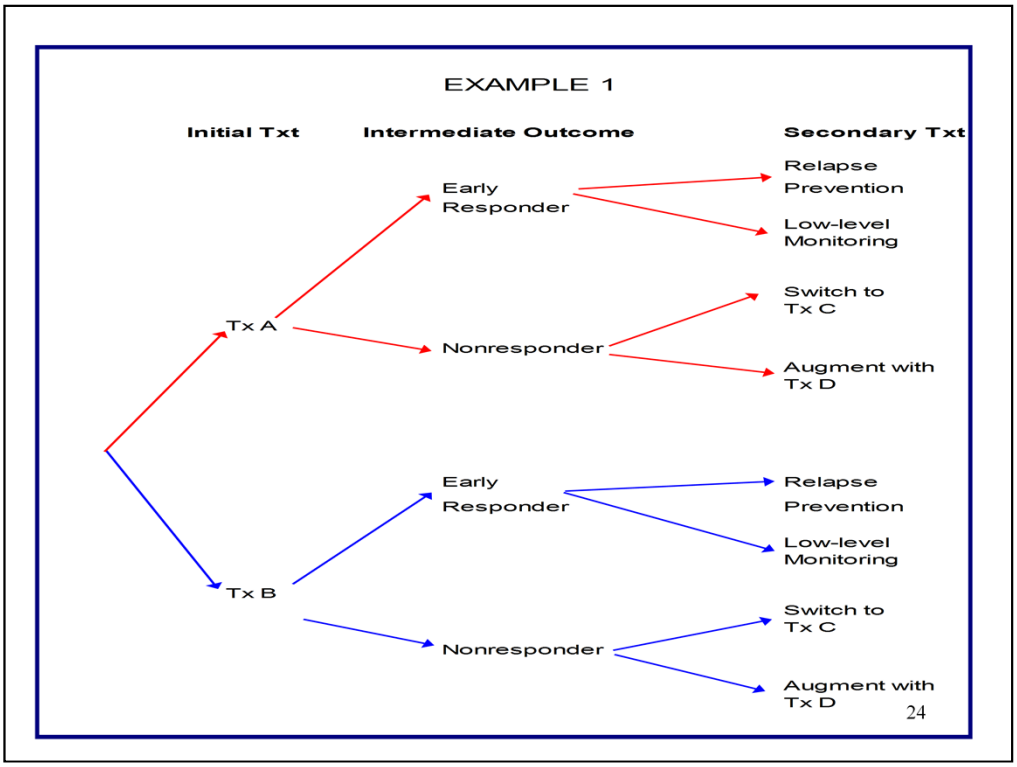
•EXAMPLE 1: (*sample size is highly constrained*): Hypothesize that adaptive interventions beginning with treatment A result in lower symptoms than adaptive interventions beginning with treatment B.

•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

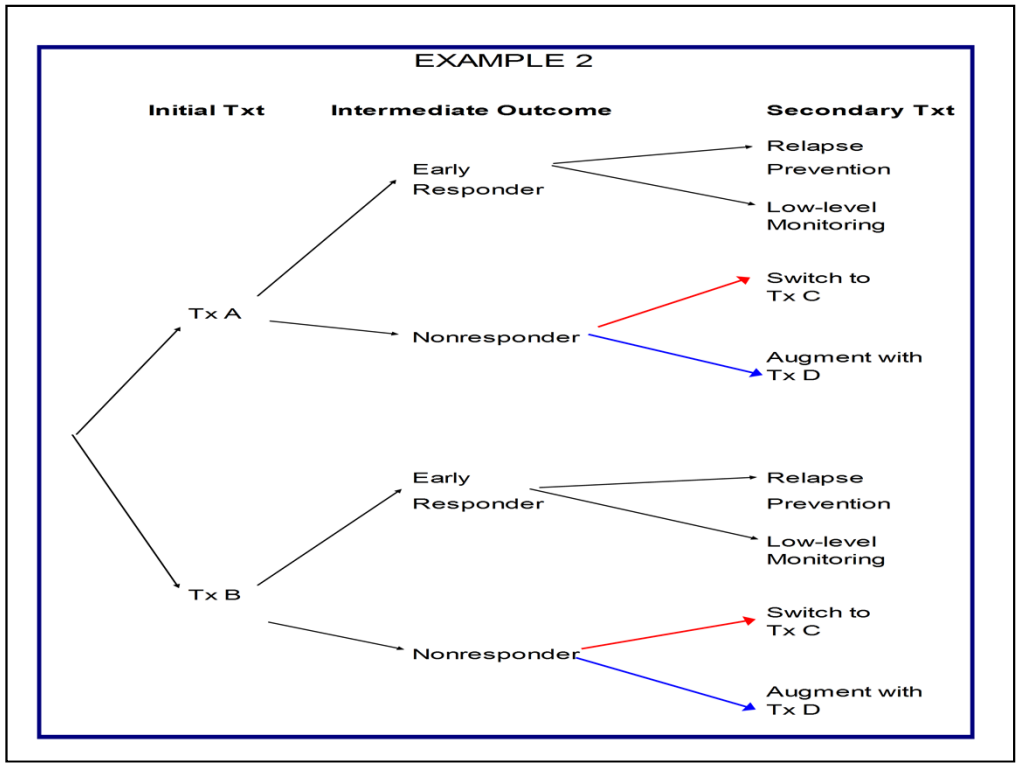
The second would be appropriate if you initially wanted to run a trial for non-responders and are now considering SMART

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: 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.*

These are main effects a la' ANOVA

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

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)

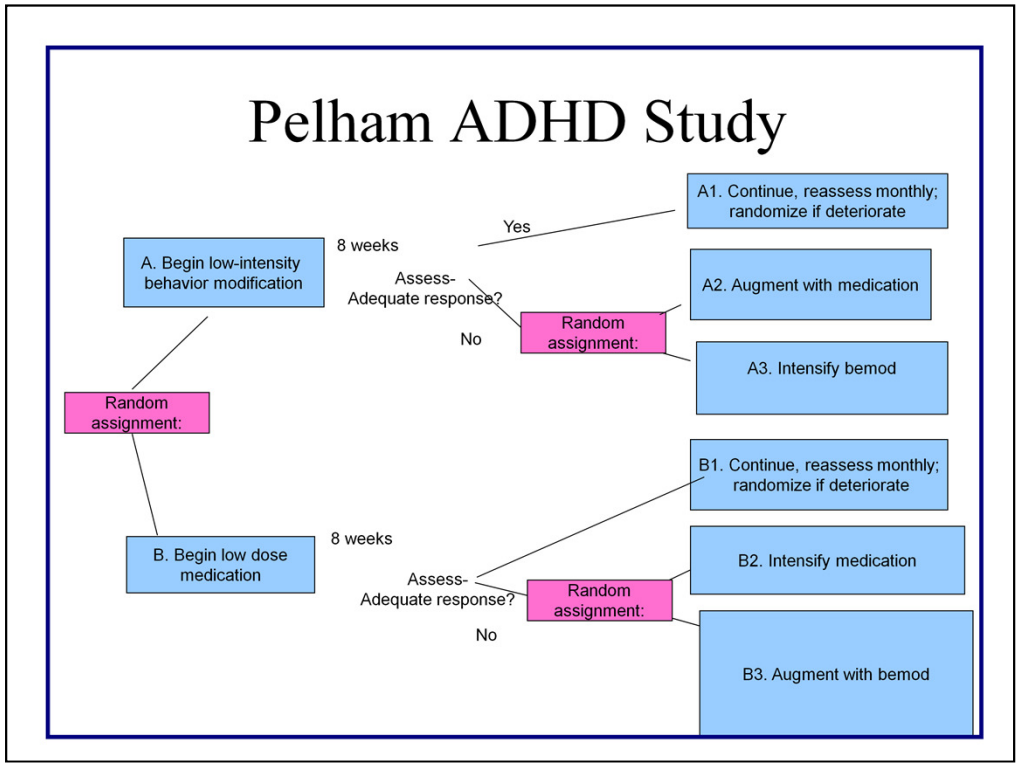
Throughout working assumptions are equal variances and normality

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

Outline

- Adaptive Interventions
- SMART Designs
- Trial Design Principles and Analysis
- Exploring Individualization using the "Adaptive Interventions for Children with ADHD" study (W. Pelham, PI).

28



The medication is Ritalin

Adaptive Pharmacological and Behavioral Treatments for Children with ADHD: Sequencing, Combining, and Escalating Doses

(1) Average performance on the teacher rated Individualized Target Behavior Evaluations – ITB-- is less than 75% AND
(2) Rating by teachers as impaired (i.e., greater than 3) on the (Impairment Rating Scale) IRS in at least one domain.

Our outcome will be a teacher rated classroom performance recorded at 8 months. N=149

I. Nahum-Shani, M. Qian, D. Almiral, 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. Psychological Methods 17(4), 457-477.

II. I. Nahum-Shani, M. Qian, D. Almiral, W.. Pelham, B. Gnagy, G. Fabiano, J. Waxmonsky, J. Yu and S.A. Murphy (2012). Q-Learning: A Data Analysis Method for Constructing Adaptive Interventions. Psychological Methods 17(4):478-94.

Exploring Greater Individualization via Q-Learning

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

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

30

Q-Learning using data on 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, whether medicated in prior year (S_1), ODD (O_1)
  - $S_1=1$ if medicated in prior year; $=0$, otherwise.
- Stage 1 involves all children

31

A1=1 if BMOD, -1 if MED

A2=1 if intensify, -1 if augment

Q-Learning using data on children with ADHD

- Stage 2 data: (X_2, A_2, Y)
  - Y = end of year school performance
  - $A_2=1$ if Intensify, $A_2=-1$ 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 $(R_1=0)$.

32

A1=1 if BMOD, -1 if MED

A2=1 if intensify, -1 if augment

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 include effects of 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 the stage 1 dependent variable—the predictor of Y, $\hat{Y}$

33

A1=1 if BMOD, -1 if MED

A2=1 if intensify, -1 if augment

Stage 2 Regression for Non-responding Children

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

34

A1=1 if BMOD, -1 if MED

A2=1 if intensify, -1 if augment

Q-Learning using data on children with ADHD

- Stage 2 regression for Y:

$$\alpha_{21} + \alpha_{22}Y_0 + \alpha_{23}S_1 + \alpha_{24}O_1 + \alpha_{25}A_1 + \alpha_{26}M_2 + \alpha_{27}S_2 + (\beta_{21} + \beta_{22}A_1 + \beta_{23}S_2)A_2$$

- **Interesting Stage 2 contrast**: Does the best stage 2 tactic (intensify versus augment) differ by whether the child/family is adherent?

35

A1=1 if BMOD, -1 if MED

S2=1 if adherent to initial txt; S2=0 if nonadherent to initial treatment.

A2=1 if intensify, -1 if augment

Q-Learning using data on children with ADHD

- Decision rule is "if child is non-responding then intensify initial treatment if $-.72 + .05A_1 + .97S_2 > 0$, otherwise augment"

Decision Rule for Non-responding Children	Initial Treatment =BMOD	Initial Treatment=MED
Adherent	Intensify	Intensify
Not Adherent	Augment	Augment

36

A1=1 if BMOD, -1 if MED

S2=1 if adherent to initial txt.

A2=1 if intensify, -1 if augment

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: prior medication (S_1)
- Control: baseline school performance, (Y_0), baseline ODD, (O_1)

37

A1=1 if BMOD, -1 if MED

A2=1 if enhance, -1 if augment

S1= 1 if prior meds, 0 otherwise

Constructing the Dependent Variable for the Stage 1 Regression

- Stage 2 regression for Y:

$$\alpha_{21} + \alpha_{22}Y_0 + \alpha_{23}S_1 + \alpha_{24}O_1 + \alpha_{25}A_1 + \alpha_{26}M_2 + \alpha_{27}S_2 + (\beta_{21} + \beta_{22}A_1 + \beta_{23}S_2)A_2$$

- Stage 1 dependent variable:

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

$$\hat{Y} = \hat{\alpha}_{21} + \hat{\alpha}_{22}Y_0 + \hat{\alpha}_{23}S_1 + \hat{\alpha}_{24}O_1 + \hat{\alpha}_{25}A_1 + \hat{\alpha}_{26}M_2 + \hat{\alpha}_{27}S_2 + |\hat{\beta}_{21} + \hat{\beta}_{22}A_1 + \hat{\beta}_{23}S_2|$$

Q-Learning using data on children with ADHD

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

$$\alpha_{11} + \alpha_{12} Y_0 + \alpha_{13} S_1 + \alpha_{14} O_1 + (\beta_{11} + \beta_{12} S_1) A_1$$

- **Interesting Stage 1 contrast**: does the best initial treatment differ by whether a child received medication in the prior year for ADHD?

39

S1=1 if on med in prior year, =0 otherwise

Q-Learning using data on children with ADHD

- Decision rule is "Begin with BMOD if $.17 - .32S_1 > 0$, otherwise begin with MED"

Initial Decision Rule	Initial Treatment
Prior MEDS	MEDS
No Prior MEDS	BMOD

40

S1= 1 if prior meds, =0 if not.

A1=1 if BMOD, -1 if MED

1st Adaptive Intervention Proposal

IF medication was not used in the prior year
THEN begin with BMOD;
ELSE select MED.

IF the child is nonresponsive and was non-adherent, THEN augment present treatment;
ELSE IF the child is nonresponsive and was adherent, THEN intensify current treatment.

41

ADHD Example

- The adaptive intervention is quite decisive. We developed this adaptive intervention using a trial on *only 138 children*. Is there sufficient evidence in the data to warrant this level of decisiveness??????
- Would a similar trial obtain similar results?
- There are strong opinions regarding how to treat ADHD.
- One solution –use confidence intervals.

42

ADHD Example

Treatment Decision for Non-responders. Positive Treatment Effect → Intensify

	90% Confidence Interval
Adherent to BMOD	(-0.08, 0.69)
Adherent to MED	(-0.18, 0.62)
Non-adherent to BMOD	(-1.10, -0.28)
Non-adherent to MED	(-1.25, -0.29)

43

ADHD Example

Initial Treatment Decision: Positive Treatment Effect → BMOD

	90% Confidence Interval
Prior MEDS	(-0.48, 0.16)
No Prior MEDS	(-0.05, 0.39)

44

2nd Adaptive Intervention Proposal

IF medication was not used in the prior year
THEN begin with BMOD;
ELSE select either BMOD or MED.

IF the child is nonresponsive and was non-adherent, THEN augment present treatment;
ELSE IF the child is nonresponsive and was adherent, THEN select either intensification or augmentation of current treatment.

45

Discussion

- For Q-Learning Software in R and in SAS:

http://methodology.psu.edu/downloads

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

46

Adherence is not a statistical nuisance; adherence indicates need to tailor treatment.

Where are we going?.....

- Increasing use of wearable computers (e.g smart phones, etc.) to both collect real time data and provide just-in-time adaptive interventions.

- We are working on the design of studies aimed at constructing and optimizing just-in-time adaptive interventions.

47

This seminar can be found at:

http://www.stat.lsa.umich.edu/~samurphy/seminars/CAPS.10.24.13.pdf

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

samurphy@umich.edu