Getting SMART about Adaptive Interventions in Education: A Conceptual Introduction

Daniel Almirall\textsuperscript{1,2}  Xi Lu (Lucy)\textsuperscript{1,2,4}
Inbal (Billie) Nahum-Shani\textsuperscript{1,2} Linda M. Collins\textsuperscript{2,3}
Susan A. Murphy\textsuperscript{1,2,4}

\textsuperscript{1}Institute for Social Research, University of Michigan
\textsuperscript{2}The Methodology Center, Penn State University
\textsuperscript{3}Department of Statistics, University of Michigan
\textsuperscript{4}Department of Statistics, Pennsylvania State University

Institute of Education Sciences - Dec-3-2014
Outline

Adaptive Interventions
  What? Why?

Evaluating versus Building an Adaptive Intervention?

Sequential Multiple Assignment Randomized Trial (SMART)
  What are SMARTs?

SMART Design Principles
  Keep it Simple
  Choosing Primary and Secondary Hypotheses

Take Home Points

SMART Case Studies
Adaptive Interventions

Evaluating versus Building an Adaptive Intervention?
Sequential Multiple Assignment Randomized Trial (SMART)
SMART Design Principles
Take Home Points
SMART Case Studies

Adaptive Interventions

What? Why?
Definition: An Adaptive Intervention is

- a sequence of individually tailored decision rules
- that specify whether, how, or when
- and based on which measures
- to alter the dosage (duration, frequency or amount), type, or delivery of treatment(s)
- at critical decision points.

Adaptive Interventions (AIs) help guide the type of sequential treatment decision making that is typical of (and often needed in!) educational or clinical practice.
Concrete Example of an Adaptive Intervention

Child ADHD in Schools, Ages 6-12

- Responder status measured by school-teacher.
- Goal is to min. symptoms / max. school performance.
What makes up a Adaptive Intervention?

1. Critical decision points: based on time or other measures
2. Treatment options at each stage
3. Tailoring variables: to decide how to adapt treatment
4. Decision rules: inputs tailoring variable, outputs treatments

aka: dynamic treatment regimens, adaptive txt strategies, treatment algorithms, medication algorithms, stepped care, txt policies, multi-stage strategies
Why are Adaptive Interventions Necessary?

In Clinical Practice...

- Nature of chronic disorders/phenomena (substance use, mental health, autism, diabetes, cancer, HIV/AIDS)
  - Waxing and waning course (multiple relapse, recurrence)
  - Life events, comorbidities, non-adherence may arise

- Disorders for which there is no widely effective treatment.

- Disorders for which there are widely effective treatments, but they are costly or burdensome.

- **Bottom line**: High heterogeneity in response to treatment
  - Within person (over time) and between person

All require sequences of treatment decisions!
PRACTICUM:
Why are Adaptive Interventions Necessary in Educational Practice or Educational Policy Settings?

- Think about educational settings in which there is high heterogeneity in response to treatment?

- Think about educational settings in which sequential decisions concerning an intervention or change in policy are made, potentially repeatedly?
Ok, so adaptive interventions are great, but...
...there are so many unanswered questions.

Now let's talk research...
GENERATING HYPOTHESES vs BUILDING vs EVALUATING ADAPTIVE INTERVENTIONS?
3 Different Research Questions/Aims
= 3 Different Research Designs

- **Aim 1**: When *generating hypotheses* about an Adaptive Intervention: e.g., Does augmenting txt (as observed in a previous trial) for non-responders correlate with better outcomes?

- **Aim 2**: When *building* an Adaptive Intervention: e.g., What are the best tailoring variables and/or decision rules?

- **Aim 3**: When *evaluating* a particular Adaptive Intervention: e.g., Does the AI have a (statistically powered) clinically significant effect compared to suitable control?
# 3 Different Research Questions/Aims

= 3 Different Research Designs

**Ex. Q1:** Does augmenting txt for non-responders (as observed in a previous trial) correlate with better outcomes?

**Ex. Q2:** What are the best tailoring variables or decision rules?

**Ex. Q3:** Does an already-developed adaptive intervention have a statistically and clinically signif. effect as compared to control intervention?

<table>
<thead>
<tr>
<th>Question</th>
<th>Aim</th>
<th>Observational Studies</th>
<th>Experimental Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Hypothesis Gen.</td>
<td>YES</td>
<td>≈</td>
</tr>
<tr>
<td>2</td>
<td>Building</td>
<td>≈</td>
<td>YES ≈</td>
</tr>
<tr>
<td>3</td>
<td>Evaluating</td>
<td>≈</td>
<td>≈ YES</td>
</tr>
</tbody>
</table>
QUESTION:
Where could SMARTs fit within IES’s Goals Structure?

1. Exploration
2. Development
3. Efficacy and Replication
4. Effectiveness
5. Measurement
SEQUENTIAL MULTIPLE ASSIGNMENT RANDOMIZED TRIALS (SMARTs)
What is a Sequential Multiple Assignment Randomized Trial (SMART)?

- Multi-stage trials; same participants throughout
- Each stage corresponds to a critical decision point
- At each stage, subjects randomized to set of treatment options
- **The goal of a SMART is to inform the development of adaptive interventions.**

I will give you an example SMART, but first…
Background for an Example SMART

ADHD Treatment in Children Ages 6-12

- Both medication (MED) and behavioral modification (BMOD) have been shown to be efficacious
- However, there is much debate on whether first-line intervention should be pharmacological or behavioral, especially in younger children
- Further, there is a need for a "rescue treatment" if the first treatment does not go well because 20-50% of children do not substantially improve on BMOD or MED
- So important questions for clinical practice include “What treatment do we begin with: BMOD or MED?” ”Among non-responders, what second treatment is best?”
Concrete Example of a SMART: Child ADHD

PI: William Pelham, PhD, Florida International University, IES-Funded Grant

\( N = 153 \), 8 month study, Monthly non-response (\( ITB < 75\% \) and \( IRS > 1 \) domain)
One of Four Adaptive Interventions Within the SMART
Embedded Adaptive Interventions in this SMART

- **AI 1**
  - MED
    - Non-Responders: Intensify: Higher Dose MED
    - Responders: Continue MED

- **AI 2**
  - MED
    - Non-Responders: Augment: MED + BMOD
    - Responders: Continue MED

- **AI 3**
  - BMOD
    - Non-Responders: Intensify: Increase BMOD
    - Responders: Continue BMOD

- **AI 4**
  - BMOD
    - Non-Responders: Augment: MED + BMOD
    - Responders: Continue BMOD
SMART DESIGN PRINCIPLES
SMART Design Principles

- KISS Principle: Keep It Simple, Straightforward
- Power for simple important primary hypotheses
- Take Appropriate steps to develop a more deeply-individualized (optimized) Adaptive Intervention
Keep It Simple, Straightforward

Overarching Principle

At each stage, or critical decision point,...

- Restrict class of treatment options *only* by ethical, feasibility, or strong scientific considerations

- If you do restrict randomizations, use low dimensional summary to restrict subsequent treatments
  - Use binary responder status
  - Should be easy to use in actual clinical practice

- Collect additional, auxiliary time-varying measures
  - To develop a more deeply-tailored Adaptive Intervention
  - Think time-varying effect moderators
SMART Design: Primary Aims

Choose a **simple primary aim/question** that aids development of an adaptive intervention.

Statistical methods used here aim to reduce uncertainty so the investigator can come away with a solid answer.

Sample size for the SMART chosen based on the hypothesis test associated with this aim (e.g., use standard $\alpha = 5\%$).
Primary Aim Example 1
What is the effect of starting with BMOD vs MED on longitudinal outcomes?

<table>
<thead>
<tr>
<th>ES</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.8</td>
<td>34</td>
</tr>
<tr>
<td>0.5</td>
<td>83</td>
</tr>
<tr>
<td>0.2</td>
<td>505</td>
</tr>
</tbody>
</table>

ρ = 0.60
α = 0.05
β = 0.20
SMART Design: Secondary Aims

Choose **secondary aims/questions** that further develop the Adaptive Intervention and take advantage of sequential randomization to eliminate confounding.

Statistical methods used here aim to generate hypotheses, e.g., generate good hypotheses about additional tailoring variables or moderators.

Here, investigators will tolerate hypothesis tests with higher Type-I error, e.g., $\alpha = 10\%$. 
Secondary Aim Example 1
Among non-responders, is it better to INTENSIFY vs AUGMENT?
On various occasions, I have seen this be the Primary Aim.
Secondary Aim Example 2
Is there a difference between two of the embedded adaptive interventions?
This could also be a Primary Aim.

Secondary Aim Example 3
Build a more deeply tailored adaptive intervention (go beyond the 4 embedded adaptive interventions). Rarely, would this be a Primary Aim.

X1 = demographics, genetics, sub-diagnoses, co-morbidities, etc...

X2 = adherence, time to NR, allegiance with therapist/psychiatrist
TAKE HOME POINTS
Take Home the Following

- SMARTs are not Adaptive Trial Designs (Confusing!!)
- Adaptive Interventions individualize treatment up-front and throughout; they are guides for clinical practice
- SMARTs are used to build better Adaptive Interventions
  - Next study: RCT of SMART-optimized AI vs control
- SMARTs do not have to be complicated; Don’t do this! :)
- SMARTs do not necessarily require larger sample sizes
SMART CASE STUDIES
(the most fun part of the conceptual overview!)
Autism SMART \((N = 61, \text{ a pilot})\)

PI: Kasari (UCLA). (ages 5-8; planned \(N = 98\) but recruitment difficult, despite multi-site. Wk12 response rates much higher than anticipated.)
Longitudinal Analysis of the Autism SMART

\[ Y_t = \text{Socially communicative utterances over 36 weeks} \]

<table>
<thead>
<tr>
<th>AI</th>
<th>Estimate</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>(AAC, AAC+)</td>
<td>51.4</td>
<td>[45.6, 57.3]</td>
</tr>
<tr>
<td>(JASP, AAC)</td>
<td>40.7</td>
<td>[34.5, 46.8]</td>
</tr>
<tr>
<td>(JASP, JASP+)</td>
<td>39.3</td>
<td>[32.6, 46.0]</td>
</tr>
</tbody>
</table>

**AI Estimate**
Child ADHD SMART
PI: William Pelham, PhD, Florida International University
\( N = 153, \) 8 month study, Monthly non-response (\( ITB < 75\% \) and \( IRS > 1 \) domain)
Longitudinal Analysis of the ADHD SMART

\[ Y_t = \text{Classroom performance over 8 months (school year)} \]

![Graph showing classroom performance over 8 months with different AI conditions and their corresponding colors: (MED, MED+) - Purple, (MED, MED+BMD) - Blue, (BMD,BMD+MED) - Green, (BMD,BMD+) - Red.](image)
Interventions for Minimally Verbal Children with Autism

PI: Kasari (UCLA), Kaiser (Vanderbilt), Smith (Rochester), Lord (Cornell), Almirall (Mich)

JASP (joint attention and social play)
- Non-Responders (Parent training no feasible)
  - JASP + DTT
- Responders (Blended txt unnecessary)
  - Continue JASP
  - JASP + Parent Training

DTT (discrete trials training)
- Non-Responders (Parent training not feasible)
  - JASP + DTT
- Responders (Blended txt unnecessary)
  - Continue DTT
  - DTT + Parent Training
Interventions for Minimally Verbal Children with Autism
PI: Kasari(UCLA), Kaiser(Vanderbilt), Smith(Rochester), Lord(Cornell), Almirall(Mich)

JASP (joint attention and social play)
- Non-Responders (Parent training no feasible)
  - Responders (Blended txt unnecessary)

DTT (discrete trials training)
- Non-Responders (Parent training not feasible)
  - Responders (Blended txt unnecessary)

JASP + DTT
- Continue JASP
- JASP + Parent Training
  - Continue JASP
  - JASP + DTT
  - Continue DTT
  - DTT + Parent Training
Adaptive Implementation Intervention in Mental Health
PI: Kilbourne; Co-I: Almirall (Aim is to improve the uptake of a psychosocial intervention for mood disorders)

60 (75% of) community-based outpatient clinics (sites) that have not responded to 6mo of REP

Augment for 6mo: REP + EF

Responder Sites

Continued Non-Responding Sites

Discontinue

Augment 6mo: REP + EF + IF

Responders Sites

Continued Non-Responding Sites

Discontinue

Continue 6mo: REP + EF + IF

Continue 6mo: REP + EF + IF
Treatment for Alcohol Dependence
PI: Oslin, University of Pennsylvania

Early Trigger for NR: 2+ HDD

Late Trigger for NR: 5+ HDD

8 Week Response

Non-Response

8 Week Response

Non-Response

Naltrexone

TDM + Naltrexone

CBI

CBI + Naltrexone

Naltrexone

TDM + Naltrexone

CBI

CBI + Naltrexone
Thank you! Questions?

Email me with questions about this presentation:
- Daniel Almirall: dalmiral@umich.edu

Find papers on SMART:
- http://www.lsa.stat.umich.edu/~samurphy/ (Susan Murphy)
- http://methcenter.psu.edu (Linda Collins)

More papers and these slides on my website (Daniel Almirall):
- http://www-personal.umich.edu/~dalmiral/
EXTRA SLIDES
Hypothesis-generating Observational Studies
Post-hoc Analyses Useful for Building Adaptive Interventions

- Give examples of different observational study questions they can examine using data from a previous 2-arm RCT

- Standard observational study caveats apply:
  - No manipulation usually means lack of heterogeneity in txt options (beyond what is controlled by experimentation in original RCT)
  - Some RCTs use samples that are too homogeneous
  - Confounding by observed baseline and time-varying factors
  - Unobserved, unknown, unmeasured confounding by baseline and time-varying factors
Hypothesis-generating Observational Studies
Post-hoc Analyses Useful for Building Adaptive Interventions

- There exists a literature for examining the impact of time-varying treatments in observational studies
  - Marginal Structural Models (Robins, 1999; Bray, Almirall, et al., 2006) to examine the marginal impact of observed time-varying sequences of treatment
  - Structural Nested Mean Models (Robins, 1994; Almirall, et al., 2010, 2011) to examine time-varying moderators of observed time-varying sequences of treatment
  - Marginal Mean Models (Murphy, et al., 2001): to examine the impact of observed adaptive interventions
Early precursors to SMART

- CATIE (2001) Treatment of Psychosis in Patients with Alzheimer’s
- CATIE (2001) Treatment of Psychosis in Patients with Schizophrenia
- STAR*D (2003) Treatment of Depression
Other Alternatives

- Piecing Together Results from Multiple Trials
  - Choose best first-line treatment on the basis of a two-arm RCT; then choose best second-line treatment on the basis of another separate, two-arm RCT
  - Concerns: delayed therapeutic effects, and cohort effects

- Observational (Non-experimental) Comparisons of AIs
  - Using data from longitudinal randomized trials
  - May yield results that inform a SMART proposal
  - Understand current treatment sequencing practices
  - Typical problems associated with observational studies

- Expert Opinion
Why Not Use Multiple Trials to Construct an AI

Three Concerns about Using Multiple Trials as an Alternative to a SMART

1. Concern 1: Delayed Therapeutic Effects

2. Concern 2: Diagnostic Effects

3. Concern 3: Cohort Effects

All three concerns emanate from the basic idea that constructing an adaptive intervention based on a myopic, local, study-to-study point of view may not be optimal.
Why Not Use Multiple Trials to Construct an AI

Concern 1: Delayed Therapeutic Effects, or Sequential Treatment Interactions

*Positive Synergy Btwn First- and Second-line Treatments*

Tapering off medication after 12 weeks of use may not appear best initially, but may have enhanced long term effectiveness when followed by a particular augmentation, switch, or maintenance strategy.

Tapering off medication after 12 weeks may set the child up for better success with any one of the second-line treatments.
Why Not Use Multiple Trials to Construct an AI

Concern 1: Delayed Therapeutic Effects, or Sequential Treatment Interactions

*Negative Synergy Btwn First- and Second-line Treatments*

Keeping the child on medication an additional 12 weeks may produce a higher proportion of responders at first, but may also result in side effects that reduce the variety of subsequent treatments available if s/he relapses.

The burden associated with continuing medication an additional 12 weeks may be so high that non-responders will not adhere to second-line treatments.
Why Not Use Multiple Trials to Construct an AI

Concern 2: Diagnostic Effects

Tapering off medication after 12 weeks initial use may not produce a higher proportion of responders at first, but may elicit symptoms that allow you to better match subsequent treatment to the child.

The improved matching (personalizing) on subsequent treatments may result in a better response overall as compared to any sequence of treatments that offered an additional 12 weeks of medication after the initial 12 weeks.
Why Not Use Multiple Trials to Construct an AI

Concern 3: Cohort Effects

► Children enrolled in the initial and secondary trials may be different.
► Children who remain in the trial(s) may be different.
► Characteristics of adherent children may differ from study to study.
► Children that know they are undergoing adaptive interventions may have different adherence patterns.

**Bottom line:** The population of children we are making inferences about may simply be different from study-to-study.