Primary Aims Using Data Arising from a SMART (Part I)

Module 4—Day 2

Getting SMART About Developing Individualized Adaptive Health Interventions

Methods Work, Chicago, Illinois, June 11-12
Daniel Almirall & 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 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
Before we begin...SAS preparation

1. Download SAS files: You received a thumb-drive
   – These files are also available for download from:
     http://www-personal.umich.edu/~dalmiral/atsworkshops.html
2. Create a folder on your notebook computer and place all of the files in that folder.
3. Inside the folder “SAS Code,” open the file “sas_code_modules_4_5_and_6_ADHD.doc”
4. Copy-paste code up to Line 20 into SAS
5. Change path on Line 20 to new folder
6. Run code (by clicking on running man). Check.

We will do step 5 together. We will also check that it worked together.
If you did this successfully, SAS should now recognize the ADHD data sets. To check: In SAS, open the Explorer tab in the left pane. Double-click on Libraries. Double-click on Libdat. You should see two data sets inside this folder: ADHD_simu... and AUTISM_simu...

In addition: You were also sent (via email) the installer files and instructions for PROC QLEARN. If you were not able to install them ahead of time, please do so before 3PM. Instructions for installing it are included in the files and on the website listed above. During lunch time I will be around to answer questions if you find trouble installing it. PROC QLEARN will be used in Module 6.
We begin by reviewing the ADHD smart study

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
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.

- **Medication**
  - Responders
  - Non-Responders

- **Behavioral Intervention**
  - Responders
  - Non-Responders

**R**

- Continue Medication
- Increase Medication Dose
- Add Behavioral Intervention
- Continue Behavioral Intervention
- Increase Behavioral Intervention
- Add Medication

O1 — A1 — O2 / R Status — A2 — Y
Response/non-response until Month 8 is the primary tailoring variable

- **Medication**
  - **Responders**
  - **Non-Responders**

- **Behavioral Intervention**
  - **Responders**
  - **Non-Responders**

**R**

- **Continue Medication**
- **Increase Medication Dose**
- **Add Behavioral Intervention**
- **Continue Behavioral Intervention**
- **Increase Behavioral Intervention**
- **Add Medication**

O1 ——— A1 ———— O2 / R Status ———— A2 ——— Y
There are a total of 6 “stage 2” treatments that any one participant may receive.
There are two “stage 2” treatment options being compared for non-responders.

- **Medication**
  - Responders
  - Non-Responders

- **Behavioral Intervention**
  - Responders
  - Non-Responders

There are two main options for non-responders:
- Continue Medication
- Increase Medication Dose
- Add Behavioral Intervention
- Continue Behavioral Intervention
- Increase Behavioral Intervention
- Add Medication

Tactical treatment options
We call responder/non-responder status the “embedded tailoring variable” since 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 treatment strategies (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)

<table>
<thead>
<tr>
<th>ID</th>
<th>O11</th>
<th>O12</th>
<th>O13</th>
<th>A1</th>
<th>R</th>
<th>A2</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1.18</td>
<td>0</td>
<td>-1</td>
<td>MED</td>
<td>1</td>
<td>.</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>-0.567</td>
<td>0</td>
<td>-1</td>
<td>0</td>
<td>1</td>
<td>INTSFY</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0.553</td>
<td>1</td>
<td>1</td>
<td>BMOD</td>
<td>0</td>
<td>-1 ADD</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>-0.013</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>-1</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>-0.571</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>-0.684</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>-1</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>1.169</td>
<td>0</td>
<td>-1</td>
<td>1</td>
<td>.</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0.369</td>
<td>1</td>
<td>-1</td>
<td>0</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>

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.
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 = intesified 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.
After you submit this code, you will see some data descriptives and a subset of the data inside the SAS output window.
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).
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.
As we go through the SAS code to analyze the simulated ADHD data set, we encourage you to follow along and actually run SAS code snippets (i.e., highlight the snippet we are discussing in the slides and hit F8). This will permit you to compare the output on your computer screen with the results shown on the slides. This will also help familiarize you with the SAS code and prepare you for the practicum using a separate data set (Autism SMART).
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!
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

<table>
<thead>
<tr>
<th>Label</th>
<th>Estimate</th>
<th>Lower</th>
<th>Upper</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Y under BMOD</td>
<td>3.0459</td>
<td>2.7859</td>
<td>3.3059</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Mean Y under MED</td>
<td>2.8608</td>
<td>2.6008</td>
<td>3.1208</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Between groups diff</td>
<td>0.1851</td>
<td>-0.1849</td>
<td>0.5551</td>
<td>0.3269</td>
</tr>
</tbody>
</table>

(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).
Try it yourself in SAS

• Using the file:
  “sas_code_modules_4_5_and_6_ADHD.doc”

• **Copy the SAS code on Page 2**

• Paste into SAS Enhanced Editor window

• Press F8 or click the Submit button (the little running man)
### Primary Question 1 Results

#### Contrast Estimate Results

<table>
<thead>
<tr>
<th>Label</th>
<th>Estimate</th>
<th>Lower</th>
<th>Upper</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Y under BMOD</td>
<td>3.0459</td>
<td>2.7859</td>
<td>3.3059</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Mean Y under MED</td>
<td>2.8608</td>
<td>2.6008</td>
<td>3.1208</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Between groups diff</td>
<td>0.1851</td>
<td>-0.1849</td>
<td>0.5551</td>
<td>0.3269</td>
</tr>
</tbody>
</table>

*(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

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

<table>
<thead>
<tr>
<th>Frequency</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>R = 0</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>R = 1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>A1 = -1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MED</td>
<td>47</td>
<td>28</td>
</tr>
<tr>
<td><strong>A1 = 1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMOD</td>
<td>52</td>
<td>23</td>
</tr>
</tbody>
</table>

- **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.
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....]
This is not a comparison of adaptive treatment strategies, 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 treatment strategies, 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.
On average, the tactic of ADDING is better and it is statistically significant, p-value < 0.01.
Try it yourself in SAS

• Go to the file:
  “sas_code_modules_4_5_and_6_ADHD.doc”
• **Copy the SAS code on Page 4**
• Paste into SAS Enhanced Editor window
• Press F8 or click the Submit button (the little running man)
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
This primary aim is a comparison of 2 adaptive treatment strategies 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...

---

**Medication**
- **Responders**
- **Non-Responders**

**Behavioral Intervention**
- **Responders**
- **Non-Responders**

---

**R**

---

**Continue Medication**
- Increase Medication Dose

**Add Behavioral Intervention**
- Continue Behavioral Intervention
- Increase Behavioral Intervention
- Add Medication

---

**O1 ——— A1 ———— O2 / R Status ———— A2 ——— Y**
...versus the mean outcome had the population followed the (BMOD, MED) ATS

O1 ——— A1 ——— O2 / R Status ——— A2 ——— Y
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 x 0.5 = 0.25 = 1/4 chance of having had followed the red ATS

This picture is just heuristic. There are actually R*N/2 in the top red cell and (1-R)*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.
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 (\text{MED,BMOD}) ATS

Analysis Of GEE Parameter Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>SError</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>2.9153</td>
<td>0.1084</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Z1</td>
<td>-0.0504</td>
<td>0.1084</td>
<td>0.6417</td>
</tr>
</tbody>
</table>

Contrast Estimate Results

<table>
<thead>
<tr>
<th>95% Conf Limits</th>
<th>Estimate Lower</th>
<th>Upper</th>
<th>SError</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Y under the red ATS</td>
<td>2.0649</td>
<td>2.5305</td>
<td>3.1992</td>
</tr>
</tbody>
</table>
Try it yourself in SAS

• Go to the file:
  “sas_code_modules_4_5_and_6_ADHD.doc”
• **Copy the SAS code on Page 5**
• Paste into SAS Enhanced Editor window
• Press F8 or click the Submit button (the little running man)
Results: Estimate of mean outcome had population followed \( (\text{MED,BMOD}) \) ATS

Analysis Of GEE Parameter Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>SError</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>2.9153</td>
<td>0.1084</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Z1</td>
<td>-0.0504</td>
<td>0.1084</td>
<td>0.6417</td>
</tr>
</tbody>
</table>

Contrast Estimate Results

<table>
<thead>
<tr>
<th>95% Conf Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimate Lower</td>
</tr>
<tr>
<td>Upper</td>
</tr>
<tr>
<td>SError</td>
</tr>
</tbody>
</table>

Mean Y under the red ATS

This analysis is with simulated data.
Citations


  – Technical Report available at the Methodology Center, PSU
Practicum

_Autism Exercises:_ In the next slide, we will briefly go over the Autism SMART study. We will also familiarize you with the “AUTISM exercise analyses starter SAS file.sas” which you will use to do the practicum.

We will go through the practicum together by filling in the ??? in the SAS starter file! The solutions have been provided to you in print so you can type in the answers. The solution is also available on my website.
We are now going to practice all of our new data analysis skills using a new data set based on an AUTISM SMART that is still currently in the field.

You have a handout with this design printed on it. Keep this handout while we go through the practicum.