Using Data Arising from a SMART to Address Primary Aims (Part II)

Module 4
General Objectives

- A taste of how data from a SMART can be analyzed to address various scientific questions
  - How to frame scientific questions
  - Experimental cells to be compared
  - Resources you can use for data analysis
Outline

Review
- ADHD SMART study
- Weighted regression approach for estimating the mean outcome under one AI

Learn
- Use weighted regression to compare the mean outcomes for two AIs that begin with different treatments
- Use weighted-and-replicated regression to simultaneously compare all embedded AIs in a SMART
Outline

Review
- ADHD SMART study
- Weighted regression approach for estimating the mean outcome under one AI

Learn
- Use weighted regression to compare the mean outcomes for two AIs that begin with different treatments
- Use weighted-and-replicated regression to simultaneously compare all embedded AIs in a SMART
ADHD SMART
PI: Pelham

First-stage intervention | Intermediate outcome | Second-stage intervention | Experimental Conditions
---|---|---|---
MED | Response | Continue | Subgroups
| Non-Response | Augment | A

BMI0 | Response | Intensify | B
| Non-Response | | C

Beginning of school year | Week 8 | End of school year
O1 | A1 | O2 / R Status | A2 | Y
ADHD SMART

PI: Pelham

4 embedded adaptive interventions

**AI #1:**
Start with MED;
if non-responder AUGMENT, else CONTINUE

**AI #2:**
Start with BMOD;
if non-responder AUGMENT, else CONTINUE

**AI #3:**
Start with MED;
if non-responder INTENSIFY, else CONTINUE

**AI #4:**
Start with BMOD;
if non-responder INTENSIFY, else CONTINUE
Recall Typical Primary Aim 3

Compare 2 embedded adaptive interventions

**AI #1:**
Start with MED;
if non-responder AUGMENT,
else CONTINUE

**AI #2:**
Start with BMOD;
if non-responder AUGMENT,
else CONTINUE
This Aim is a Comparison of Mean Outcome Under AI #1 vs. mean outcome under AI #2
We Know How to Account for the Imbalance in Non-Responders Following AI #1

Assign $W = \text{weight} = 2$ to responders to MED: $2 \times \frac{1}{2} = 1$

Assign $W = \text{weight} = 4$ to non-responders to MED: $4 \times \frac{1}{4} = 1$

Then we take $W$-weighted mean of sample who ended up in boxes A & B.
A Similar Approach (and SAS Code) Can be Used to Obtain Mean Under AI #2

Assign $W = \text{weight} = 2$ to responders to MED: $2 \times \frac{1}{2} = 1$

Assign $W = \text{weight} = 4$ to non-responders to MED: $4 \times \frac{1}{4} = 1$

Then we take $W$-weighted mean of sample who ended up in boxes D & E.
Results for Estimated Mean Outcome had All Participants Followed AI#2 (BMOD, AUGMENT)

| Parameter     | Estimate | Standard Error | Pr > |Z| |
|---------------|----------|----------------|------|---|
| Intercept     | 3.0982   | 0.1070         | <.0001 | 
| Z1            | 0.4085   | 0.1070         | 0.0001 |

### Contrast Estimate Results

<table>
<thead>
<tr>
<th>Label</th>
<th>Mean Estimate</th>
<th>95% Confidence Limits</th>
<th>Standard Error</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Y under AI #2 (BMOD, AUGMENT)</td>
<td><strong>3.5067</strong></td>
<td>3.1643 - 3.8490</td>
<td>0.1747</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

**Interpretation:** The estimated mean school performance score for children consistent with AI #2 is ~3.51 (95% CI: (3.16, 3.85)).

This analysis is with simulated data.
Outline

Review

• ADHD SMART study
• Weighted regression approach for estimating the mean outcome under one AI

Learn

• Use weighted regression to compare the mean outcomes for two AIs that begin with different treatments
• Use weighted-and-replicated regression to simultaneously compare all embedded AIs in a SMART
An Intuitive (Yet Less Efficient) Approach to Comparing AI#1 vs AI#2
Reminder of Coding Scheme

First-stage intervention

- **A1= -1**
- **R= 1**

**MED**
- **A1= 1**
- **R= 0**

Beginning of school year

**O1**

**A1**

**O2 / R Status**

**A2**

End of school year

**Y**

Intermediate outcome

- **Response**
  - **R= 1**

**Non-Response**

- **R= 0**

Second-stage intervention

- **Continue**
  - **A2= .**
  - **A2= 1**

- **Augment**
  - **A2= -1**

- **Intensify**
  - **A2= 1**

Subgroups

**A**

**B**

**C**

**D**

**E**

**F**
data dat7; set dat1;
  Z1=-1;
  if A1=-1 and R=1 then Z1=1;
  if A1=-1 and R=0 and A2=-1 then Z1=1;
Z2=-1;
  if A1=1 and R=1 then Z2=1;
  if A1=1 and R=0 and A2=-1 then Z2=1;
W=2*R + 4*(1-R);
run;

data dat8;
  set dat7; if Z1=1 or Z2=1;
run;
An Intuitive (Yet Less Efficient) Approach to Comparing AI#1 vs AI#2

Create Z1:
Indicator for whether or not the person is consistent with AI#1

\[
\text{Z1} = \begin{cases} 
-1 & \text{if } \text{A1} = -1 \text{ and } \text{R} = 1 \\
1 & \text{if } \text{A1} = -1 \text{ and } \text{R} = 0 \text{ and } \text{A2} = -1 \\
-1 & \text{if } \text{A1} = 1 \text{ and } \text{R} = 1 \\
1 & \text{if } \text{A1} = 1 \text{ and } \text{R} = 0 \text{ and } \text{A2} = -1 
\end{cases}
\]

\[W = 2 \times \text{R} + 4 \times (1 - \text{R})\]

run;

data dat8;
set dat7; if Z1 = 1 or Z2 = 1;
run;
An Intuitive (Yet Less Efficient) Approach to Comparing AI#1 vs AI#2

data dat7; set dat1;
Z1=-1;
  if A1=-1 and R=1 then Z1=1;
  if A1=-1 and R=0 and A2=-1 then Z1=1;
Z2=-1;
  if A1=1 and R=1 then Z2=1;
  if A1=1 and R=0 and A2=-1 then Z2=1;
W=2*R + 4*(1-R);
run;

data dat8;
  set dat7; if Z1=1 or Z2=1;
run;

Create Z2:
Indicator for whether or not the person is consistent with AI#2
An Intuitive (Yet Less Efficient) Approach to Comparing AI#1 vs AI#2

```
data dat7; set dat1;
Z1=-1;
    if A1=-1 and R=1 then Z1=1;
    if A1=-1 and R=0 and A2=-1 then Z1=1;
Z2=-1;
    if A1=1 and R=1 then Z2=1;
    if A1=1 and R=0 and A2=-1 then Z2=1;
W=2*R + 4*(1-R);
run;
```

Assign weights:
2 for responders
4 for non-responders

```
data dat8;
    set dat7; if Z1=1 or Z2=1;
run;
```
An Intuitive (Yet Less Efficient) Approach to Comparing AI#1 vs AI#2

data dat7; set dat1;
Z1=-1;
    if A1=-1 and R=1 then Z1=1;
    if A1=-1 and R=0 and A2=-1 then Z1=1;
Z2=-1;
    if A1=1 and R=1 then Z2=1;
    if A1=1 and R=0 and A2=-1 then Z2=1;
W=2*R + 4*(1-R);
run;

data dat8;
    set dat7; if Z1=1 or Z2=1;
run;

Subset Data:
Keep only participants consistent with either AI#1 or AI#2
An Intuitive (Yet Less Efficient) Approach to Comparing AI#1 vs AI#2

The Regression and Contrast Coding Logic:

Recall:

$Z_1$ is now an indicator for whether the person is consistent with AI#1 or with AI#2:

$\rightarrow Z_1 = 1 = \text{AI#1}$
$\rightarrow Z_1 = -1 = \text{AI#2}$

To compare the 2 AIs, we can fit the Model:

$$E(Y|Z_1) = \beta_0 + \beta_1 Z_1$$

Overall Mean $Y$ under AI#1 $= \beta_0 + \beta_1 \times 1$

Overall Mean $Y$ under AI#2 $= \beta_0 + \beta_1 \times -1$

Diff Between AIs $= \beta_0 + \beta_1 - (\beta_0 - \beta_1) = 2\beta_1$
An Intuitive (Yet Less Efficient) Approach to Comparing AI#1 vs AI#2

\begin{verbatim}
proc genmod data = dat8;
  class id;
  model Y = Z1;
  weight W;
  repeated subject = id / type = ind;
  estimate 'Mean Y AI#1(MED, Add BMOD)' intercept 1 Z1 1;
  estimate 'Mean Y AI#2(BMOD, Add MED)' intercept 1 Z1 -1;
  estimate 'Diff: AI#1 - AI#2' Z1 2;
run;
\end{verbatim}

\begin{align*}
\text{Mean Y under AI#1} & = \beta_0 + \beta_1 \times 1 \\
\text{Mean Y under AI#2} & = \beta_0 + \beta_1 \times -1 \\
\text{Diff Between AIs} & = 2\beta_1
\end{align*}
An Intuitive (Yet Less Efficient) Approach to Comparing AI#1 vs AI#2

| Parameter | Estimate | Standard Error | Pr > |Z| |
|-----------|----------|----------------|------|---|
| Intercept | 3.1858   | 0.1221         | <.0001 |
| Z1        | -0.3209  | 0.1221         | 0.0086 |

**Contrast Estimate Results**

<table>
<thead>
<tr>
<th>Label</th>
<th>Mean Estimate</th>
<th>95% Confidence Limits</th>
<th>Standard Error</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Y under AI #1 (MED, AUGMENT)</td>
<td>2.8649</td>
<td>2.5305 - 3.1992</td>
<td>0.1706</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Mean Y under AI #2 (BMOD, AUGMENT)</td>
<td>3.5067</td>
<td>3.1643 - 3.8490</td>
<td>0.1747</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Diff: AI#1 – AI#2</td>
<td>-0.6418</td>
<td>-1.1203 - 0.1633</td>
<td>0.2442</td>
<td>0.0086</td>
</tr>
</tbody>
</table>

This analysis is with simulated data.
An Intuitive (Yet Less Efficient) Approach to Comparing AI#1 vs AI#2

```
proc genmod data = dat8;
  class id;
  model Y = Z1 012c 014c;
  weight w;
  repeated subject = id / type = ind;
  estimate 'Mean Y AI#1(MED, AUGMENT)' intercept 1 Z1 1;
  estimate 'Mean Y AI#2(BMOD,AUGMENT)' intercept 1 Z1 -1;
  estimate 'Diff: AI#1 - AI#2' intercept 1 Z1 2;
run;
```

Add baseline control covariates
# An Intuitive (Yet Less Efficient) Approach to Comparing AI#1 vs AI#2

## Analysis Of GEE Parameter Estimates

| Parameter | Estimate | Standard Error | Pr > |Z| |
|-----------|----------|----------------|-------|---|
| Intercept | 3.1858   | 0.1221         | <.0001|
| Z1        | -0.2442  | 0.1122         | 0.0295|
| O12c      | -0.5153  | 0.0971         | <.0001|
| O14c      | 0.4905   | 0.2774         | 0.0770|

## Contrast Estimate Results

<table>
<thead>
<tr>
<th>Label</th>
<th>Mean Estimate</th>
<th>95% Confidence Limits</th>
<th>Standard Error</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Y under AI #1</td>
<td>2.8842</td>
<td>2.5919 - 3.1765</td>
<td>0.1491</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Mean Y under AI #2</td>
<td>3.3727</td>
<td>3.0542 - 3.6912</td>
<td>0.1625</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Diff: AI#1 – AI#2</td>
<td>-0.4884</td>
<td>-0.9282 - -0.0487</td>
<td>0.2244</td>
<td>0.0295</td>
</tr>
</tbody>
</table>

Notice SE: Slightly smaller compared to the analysis without control covariates.

This analysis is with simulated data.
Outline

Review
• ADHD SMART study
• Weighted regression approach for estimating the mean outcome under one AI

Learn
• Use weighted regression to compare the mean outcomes for two AIs that begin with different treatments
• Use weighted-and-replicated regression to simultaneously compare all embedded AIs in a SMART
What about a Regression to Compare AI#1 (MED, AUGMENT) vs...
AI#2 (BMOD, AUGMENT) vs...
... AI#3 (MED, INTENSIFY) vs...

First-stage intervention | Intermediate outcome | Second-stage intervention
-------------------------|----------------------|-------------------------
MED                      | Continue A           | Subgroups
Non-Response            | Augment B            | A
Response                 | Intensify C          | B
Non-Response             | Continue D           | C
Response                 | Augment E            | D
Non-Response             | Intensify F          | E

Beginning of school year | Week 8 | End of school year
O1 A1 O2 / R Status A2 Y
... AI#4 (BMOD, INTENSIFY), All In One Swoop?
Notice that AI#1 and AI#3 (start MED) Share Responders (Box A)
Notice that AI#1 and AI#3 (start MED) Share Responders (Box A)
Similarly: AI#2 and AI#4 (start BMOD) Share Responders (Box D)
Similarly: AI#2 and AI#4 (start BMOD) Share Responders (Box D)
So, What’s Going On?

In this SMART, all responders are consistent with two AIs

- Responders to MED are part of AI#1 and AI#3
- Responders to BMOD are part of AI#2 and AI#4

If our goal is to estimate the mean outcome under all AIs simultaneously,

We must share responders somehow.

- But how?
What Do We Do?

We “trick” SAS into using the responders twice

We do this by replicating responders:
  • Create 2 observations for each responder
  • We assign half of them $A2=1$, the other half $A2=-1$

$w=2$ to responders and $w=4$ to non-responders

Robust standard errors account for weighting and the fact that responders are “re-used”. No cheating here!
Weighting and Replicating Serve Different Purposes

**Weighting**
- Accounts for over/underrepresentation of responders or non-responders
- Because of the randomization scheme

**Replicating**
- Allows us to use standard software to do simultaneous estimation and comparison
- Because participants are consistent with more than one AI
data dat9; set dat1;
  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;
## Replicated Data

<table>
<thead>
<tr>
<th>Obs</th>
<th>ID</th>
<th>A1</th>
<th>R</th>
<th>A2</th>
<th>Y</th>
<th>o11c</th>
<th>o12c</th>
<th>o13c</th>
<th>o14c</th>
<th>ob</th>
<th>weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>45</td>
<td>32</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>5</td>
<td>-0.35333</td>
<td>-2.73889</td>
<td>-0.31333</td>
<td>0.19333</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>46</td>
<td>32</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>-0.35333</td>
<td>-2.73889</td>
<td>-0.31333</td>
<td>0.19333</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>47</td>
<td>33</td>
<td>1</td>
<td>0</td>
<td>-1</td>
<td>3</td>
<td>0.64667</td>
<td>-1.07820</td>
<td>0.68667</td>
<td>0.19333</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>48</td>
<td>34</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>-1</td>
<td>0.21662</td>
<td>-0.31333</td>
<td>0.19333</td>
<td>1</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>35</td>
<td>1</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>-0.35333</td>
<td>0.58276</td>
<td>-0.31333</td>
<td>0.19333</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>50</td>
<td>36</td>
<td>-1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0.64667</td>
<td>-0.03527</td>
<td>-0.31333</td>
<td>0.19333</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>51</td>
<td>37</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-0.35333</td>
<td>0.99556</td>
<td>-0.31333</td>
<td>0.19333</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>52</td>
<td>37</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-0.35333</td>
<td>0.99556</td>
<td>-0.31333</td>
<td>0.19333</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>53</td>
<td>38</td>
<td>-1</td>
<td>0</td>
<td>-1</td>
<td>3</td>
<td>-0.35333</td>
<td>0.14034</td>
<td>0.68667</td>
<td>-0.80667</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>54</td>
<td>39</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>0.64667</td>
<td>1.64983</td>
<td>0.68667</td>
<td>0.19333</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>55</td>
<td>39</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>0.64667</td>
<td>1.64983</td>
<td>0.68667</td>
<td>0.19333</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>
After Weighting & Replicating: SAS Code for the Weighted Regression

The Regression and Contrast Coding Logic:

Recall:

Our goal is to compare all 4 embedded AIs
We have 2 indicators: $A_1, A_2$:

<table>
<thead>
<tr>
<th>$A_1$</th>
<th>$A_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>BMOD</td>
</tr>
<tr>
<td>-1</td>
<td>MED</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1</th>
<th>INTENSIFY</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1</td>
<td>AUGMENT</td>
</tr>
</tbody>
</table>

To compare all 4 AIs, we can fit the following model:

$$E(Y|A_1, A_2) = \beta_0 + \beta_1 A_1 + \beta_2 A_2 + \beta_3 A_1 A_2$$
After Weighting & Replicating: SAS Code for the Weighted Regression

The Regression and Contrast Coding Logic:

\[ E(Y|A_1, A_2) = \beta_0 + \beta_1 A_1 + \beta_2 A_2 + \beta_3 A_1 A_2 \]

<table>
<thead>
<tr>
<th>AI</th>
<th>Mean Y Under AI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1  (MED, AUGMENT)</td>
<td>(\beta_0 + \beta_1(-1) + \beta_2(-1) + \beta_3(-1)(-1))</td>
</tr>
<tr>
<td>2  (BMOD, AUGMENT)</td>
<td>(\beta_0 + \beta_1(1) + \beta_2(-1) + \beta_3(1)(-1))</td>
</tr>
<tr>
<td>3  (MED, INTENSIFY)</td>
<td>(\beta_0 + \beta_1(-1) + \beta_2(1) + \beta_3(-1)(1))</td>
</tr>
<tr>
<td>4  (BMOD, INTENSIFY)</td>
<td>(\beta_0 + \beta_1(1) + \beta_2(1) + \beta_3(1)(1))</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>A_1</th>
<th>A_2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>BMOD</td>
</tr>
<tr>
<td>-1</td>
<td>MED</td>
</tr>
<tr>
<td>1</td>
<td>INTENSIFY</td>
</tr>
<tr>
<td>-1</td>
<td>AUGMENT</td>
</tr>
</tbody>
</table>
After Weighting & Replicating: SAS Code for the Weighted Regression

The Regression and Contrast Coding Logic:

\[ E(Y|A_1, A_2) = \beta_0 + \beta_1 A_1 + \beta_2 A_2 + \beta_3 A_1 A_2 \]

<table>
<thead>
<tr>
<th>AI</th>
<th>Mean Y Under AI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( \beta_0 + \beta_1(-1) + \beta_2(-1) + \beta_3(-1)(-1) )</td>
</tr>
<tr>
<td>2</td>
<td>( \beta_0 + \beta_1(1) + \beta_2(-1) + \beta_3(1)(-1) )</td>
</tr>
<tr>
<td>3</td>
<td>( \beta_0 + \beta_1(-1) + \beta_2(1) + \beta_3(-1)(1) )</td>
</tr>
<tr>
<td>4</td>
<td>( \beta_0 + \beta_1(1) + \beta_2(1) + \beta_3(1)(1) )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( A_1 )</th>
<th>( A_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>BMOD</td>
</tr>
<tr>
<td>-1</td>
<td>MED</td>
</tr>
<tr>
<td>1</td>
<td>INTENSIFY</td>
</tr>
<tr>
<td>-1</td>
<td>AUGMENT</td>
</tr>
</tbody>
</table>
After Weighting & Replicating: SAS Code for the Weighted Regression

The Regression and Contrast Coding Logic:

\[ E(Y|A_1, A_2) = \beta_0 + \beta_1 A_1 + \beta_2 A_2 + \beta_3 A_1 A_2 \]

<table>
<thead>
<tr>
<th>AI</th>
<th>Mean Y Under AI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (-1, -1)</td>
<td>(\beta_0 - \beta_1 - \beta_2 + \beta_3)</td>
</tr>
<tr>
<td>2 (1, -1)</td>
<td>(\beta_0 + \beta_1 - \beta_2 - \beta_3)</td>
</tr>
<tr>
<td>3 (-1, 1)</td>
<td>(\beta_0 - \beta_1 + \beta_2 - \beta_3)</td>
</tr>
<tr>
<td>4 (1, 1)</td>
<td>(\beta_0 + \beta_1 + \beta_2 + \beta_3)</td>
</tr>
</tbody>
</table>

The difference between AI#1 and AI#2:

\[
(\beta_0 - \beta_1 - \beta_2 + \beta_3) - (\beta_0 + \beta_1 - \beta_2 - \beta_3) = -2\beta_1 + 2\beta_3
\]
After Weighting & Replicating: SAS Code for the Weighted Regression

```sas
proc genmod data = dat9;
  class id;
  model Y = A1 A2 A1*A2;
  weight weight;
  repeated subject = id / type = ind;
  estimate 'MeanY:AI#1(MED,AUGMENT)' int 1 A1 -1 A2 -1 A1*A2 1;
  estimate 'MeanY:AI#2(BMOD,AUGMENT)' int 1 A1 1 A2 -1 A1*A2 -1;
  estimate 'MeanY:AI#3(MED,INTNSFY)' int 1 A1 -1 A2 1 A1*A2 -1;
  estimate 'MeanY:AI#4(BMOD,INTNSFY)' int 1 A1 1 A2 1 A1*A2 1;
  estimate 'Diff: AI#1 - AI#2' int 0 A1 -2 A2 0 A1*A2 2;
  estimate 'Diff: AI#1 - AI#3' int 0 A1 0 A2 -2 A1*A2 2;
  estimate 'Diff: AI#1 - AI#4' int 0 A1 -2 A2 -2 A1*A2 0;
*etc...;
run;
```
After Weighting & Replicating: SAS Code for the Weighted Regression

```
proc genmod data = dat9;
  class id;
  model Y = A1 A2 A1*A2;
  weight weight;
  repeated subject = id / type = ind;
  estimate 'MeanY:AI#1(MED,AUGMENT)' int 1 A1 -1 A2 -1 A1*A2 1;
  estimate 'MeanY:AI#2(BMOD,AUGMENT)' int 1 A1 1 A2 -1 A1*A2 -1;
  estimate 'MeanY:AI#3(MED,INTENSFY)' int 1 A1 -1 A2 1 A1*A2 -1;
  estimate 'MeanY:AI#4(BMOD,INTNSFY)' int 1 A1 1 A2 1 A1*A2 1;
  estimate 'Diff: AI#1 - AI#2' int 0 A1 -2 A2 0 A1*A2 2;
  estimate 'Diff: AI#1 - AI#3' int 0 A1 0 A2 -2 A1*A2 2;
  estimate 'Diff: AI#1 - AI#4' int 0 A1 -2 A2 -2 A1*A2 0;
*etc...;
run;
```

**Estimate Difference:**

\[ \text{Diff AI #1} - \text{AI #2} = -2\beta_1 + 2\beta_3 \]
## Results for Weighted & Replicated Regression: Comparing Mean Outcome for all AIs Simultaneously

<table>
<thead>
<tr>
<th>Label</th>
<th>Mean Estimate</th>
<th>95% Confidence Limits</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Y under AI #1 (MED, AUGMENT)</td>
<td>2.8649</td>
<td>2.5305 - 3.1992</td>
<td>0.1706</td>
</tr>
<tr>
<td>Mean Y under AI #2 (BMOD, AUGMENT)</td>
<td>3.5067</td>
<td>3.1643 - 3.8490</td>
<td>0.1747</td>
</tr>
<tr>
<td>Mean Y under AI #3 (MED, INTENSIFY)</td>
<td>2.7895</td>
<td>2.4644 - 3.1145</td>
<td>0.1658</td>
</tr>
<tr>
<td>Mean Y under AI #4 (BMOD, INTENSIFY)</td>
<td>2.6533</td>
<td>2.2515 - 3.0552</td>
<td>0.2050</td>
</tr>
<tr>
<td>Diff: AI#1 – AI#2</td>
<td>-0.6418</td>
<td>-1.1203 - -0.1633</td>
<td>0.2442</td>
</tr>
<tr>
<td>Diff: AI#1 – AI#3</td>
<td>0.0754</td>
<td>-0.3106 - 0.4614</td>
<td>0.1969</td>
</tr>
<tr>
<td>Diff: AI#1 – AI#4</td>
<td>0.2115</td>
<td>-0.3112 - 0.7343</td>
<td>0.2667</td>
</tr>
</tbody>
</table>

**NOTE:** We get the exact same results as before when we compared AI#1 vs AI#2, but now we can simultaneously make inference for all the comparisons.

This analysis is with simulated data.
But wait!...
There’s More to Weighted & Replicated Regression Than Just Convenience!
Weighted & Replicated Regression is More Efficient Statistically

```
proc genmod data = dat9;
  class id;
  model Y = A1 A2 A1*A2 012c 014c;
  weight weight;
  repeated subject = id / type = ind;
  estimate 'MeanY:AI#1(MED,AUGMENT)' int 1 A1 -1 A2 -1 A1*A2 1;
  estimate 'MeanY:AI#2(BMOD,AUGMENT)' int 1 A1 1 A2 -1 A1*A2 -1;
  estimate 'MeanY:AI#3(MED,INTENSFY)' int 1 A1 -1 A2 1 A1*A2 -1;
  estimate 'MeanY:AI#4(BMOD,INTNSFY)' int 1 A1 1 A2 1 A1*A2 1;
  estimate 'Diff: AI#1 - AI#2' int 0 A1 -2 A2 0 A1*A2 2;
  estimate 'Diff: AI#1 - AI#3' int 0 A1 0 A2 -2 A1*A2 2;
  estimate 'Diff: AI#1 - AI#4' int 0 A1 -2 A2 -2 A1*A2 0;
*etc...;
run;
```

**Improve power:**
Adjusting for baseline covariates that are associated with outcome leads to more efficient estimates (lower standard error = more power = smaller p-value).
### Improved efficiency: Adjusting for baseline covariates resulted in lower standard error and tighter confidence intervals. Point estimates remained about the same, as expected.

### Weighted & Replicated Regression is More Efficient Statistically

<table>
<thead>
<tr>
<th>Label</th>
<th>Mean Estimate</th>
<th>95% Confidence Limits</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Y under AI #1 (MED, AUGMENT)</td>
<td>2.8801</td>
<td>Lower 2.5869, Upper 3.1733</td>
<td>0.1496</td>
</tr>
<tr>
<td>Mean Y under AI #2 (BMOD, AUGMENT)</td>
<td>3.3854</td>
<td>Lower 3.0689, Upper 3.7018</td>
<td>0.1614</td>
</tr>
<tr>
<td>Mean Y under AI #3 (MED, INTENSIFY)</td>
<td>2.8149</td>
<td>Lower 2.5163, Upper 3.1135</td>
<td>0.1524</td>
</tr>
<tr>
<td>Mean Y under AI #4 (BMOD, INTENSIFY)</td>
<td>2.7338</td>
<td>Lower 2.3596, Upper 3.1081</td>
<td>0.1909</td>
</tr>
<tr>
<td>Diff: AI#1 – AI#2</td>
<td>-0.5053</td>
<td>Lower -0.9401, Upper -0.0704</td>
<td>0.2219</td>
</tr>
</tbody>
</table>

SE in unadjusted model was 0.2442
### Weighted & Replicated Regression is More Efficient Statistically

#### Contrast Estimate Results

<table>
<thead>
<tr>
<th>Label</th>
<th>Mean Estimate</th>
<th>95% Confidence Limits</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Y under AI #1 (MED, AUGMENT)</td>
<td>2.8801</td>
<td>2.5869 - 3.1733</td>
<td>0.1496</td>
</tr>
<tr>
<td>Mean Y under AI #2 (BMOD, AUGMENT)</td>
<td>3.3854</td>
<td>3.0689 - 3.7018</td>
<td>0.1614</td>
</tr>
<tr>
<td>Mean Y under AI #3 (MED, INTENSIFY)</td>
<td>2.8149</td>
<td>2.5163 - 3.1135</td>
<td>0.1524</td>
</tr>
<tr>
<td>Mean Y under AI #4 (BMOD, INTENSIFY)</td>
<td>2.7338</td>
<td>2.3596 - 3.1081</td>
<td>0.1909</td>
</tr>
</tbody>
</table>

**Diff: AI#1 – AI#2**

-0.5053, -0.9401, -0.0704, 0.2219

*This analysis is with simulated data.*

SE in unadjusted model was **0.2442**

SE in adjusted model including only data from participants in AI #1 and AI #2 was **0.2244**
Citations
