

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

## Module 5

### *Getting SMART: Experimental Design and Analysis Methods for Developing Adaptive Interventions*



50 minutes

This is the 5<sup>th</sup> module of a 5-module workshop on experimental designs for building optimal adaptive interventions. By now, you know what an adaptive intervention 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.

By now you have also learned how to address 2 typical primary research questions (main effect of first-line treatment and effect of second-stage treatments (operationalized various ways, tactical and treatment))

By now you were also introduced to a weighting approach for estimating the mean outcome under 1 of the SMART design-embedded ATSs.

In this module, we are going to continue discussing the weighting approach. The goal is to continue to learn about this approach and learn how to use a weighting approach to estimate and compare the mean outcome for all of the design-embedded ATSs. This will be the final, 3<sup>rd</sup>, typical primary aim we discuss in this workshop.

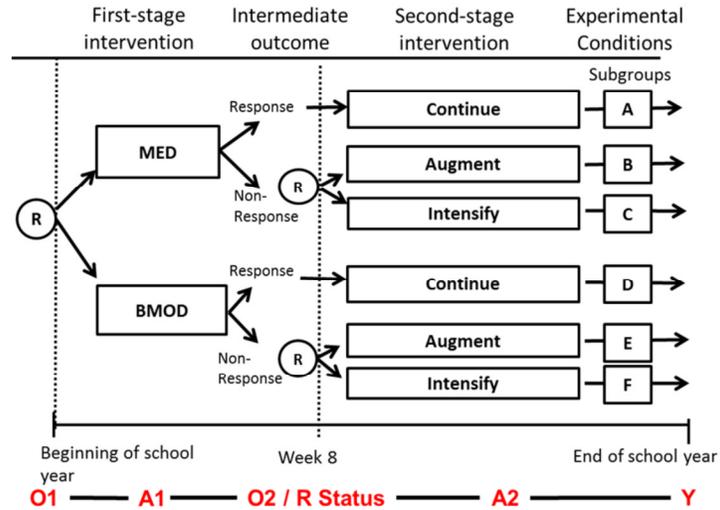
# Outline

- 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

In this module, we're going to build towards the last bullet point: using weighted-and-replicated regression to simultaneously compare all embedded adaptive interventions in a SMART. In your research, when analyzing data arising from a SMART, you'll want to use this simultaneous approach.

On our way there, we'll review how to estimate the mean outcome under one adaptive intervention and learn how to compare two AIs that start with different treatments. Then, we'll use this as a foundation to learn the weighted-and-replicated regression approach to simultaneously compare all embedded AIs.

## ADHD SMART



A1: Indicator for first-stage treatment. A1=-1 for MED, A1=1 for BMOD.

R: Indicator for response status. R=1 for responders, R=0 for non-responders.

A2: Indicator for second-stage treatment *among non-responders*. A2=-1 for AUGMENT, A2=1 for INTENSIFY. (Remember that A2 is only defined for non-responders!)

Y: Outcome variable of interest. In the following examples, we use school performance score. For this outcome, higher values are better.

## ADHD (Pelham)

### 4 Embedded Adaptive Interventions

- 1) Start with BMOD; if non-responder intensify BMOD, else continue BMOD
- 2) Start with BMOD; if non-responder BMOD +MED, else continue BMOD
- 3) Start with low-dose MED; if non-responder Intensify MED, else continue low-dose Med
- 4) Start with low-dose MED; if non-responder BMOD+MED, else continue low-dose Med.

You can also conceptualize the second-stage treatments in these embedded adaptive interventions as tactics rather than treatments.

Recall Typical Primary Aim 3:  
Best of 2 design-embedded AIs

**AI #1: Start with MED, then**

- If response, then **CONTINUE with medication**
- If non-response, then **AUGMENT with BMOD**

**VS.**

**AI#2: Start with BMOD, then**

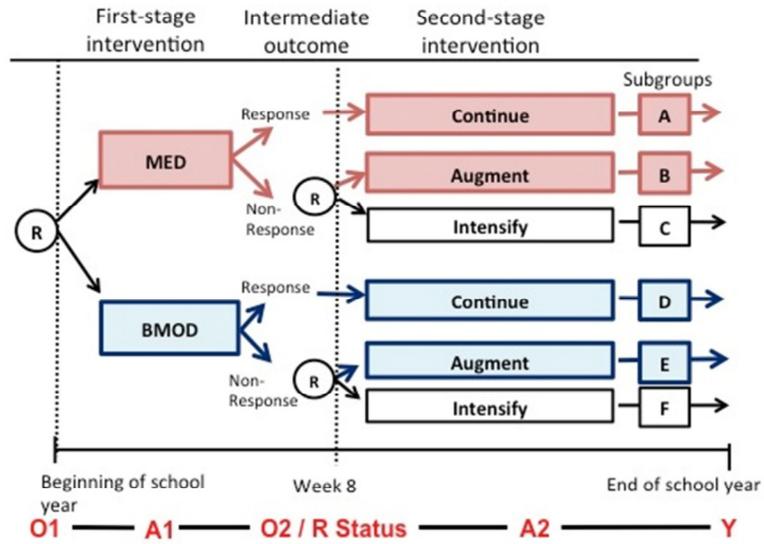
- If response, then **CONTINUE with BMOD**
- If non-response, then **AUGMENT with MED**

This primary aim is a comparison of 2 adaptive interventions that begin with *different* first-line treatments.

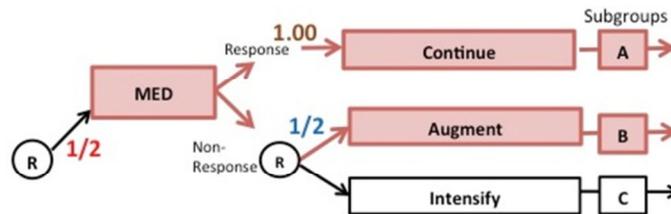
It is a comparison of two decision rules (notice the if/then). This comparison is like if you had a 2-arm RCT where 1 arm was randomized to AI#1 and the other arm was randomized to AI#2.

One could also do all remaining pair-wise comparisons between the 4 embedded AIs. Here we chose 1 pair for illustration.

## This Aim is a Comparison of the Mean Outcome under AI#1 vs the Mean Outcome AI#2



## You Already Learned How to Account for the Imbalance in Non-Responders Following AI#1

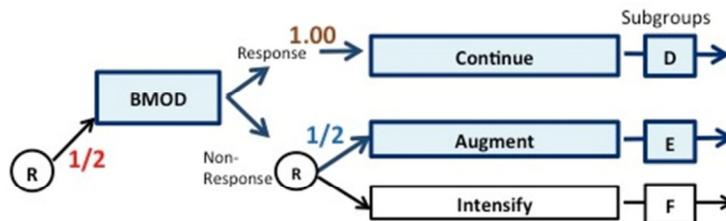


- Assign  $W = \text{weight} = 2$  to responders to MED  $\rightarrow 2 * \frac{1}{2} = 1$
- Assign  $W = \text{weight} = 4$  to non-responders to MED  $\rightarrow 4 * \frac{1}{4} = 1$
- Then we take  $W$ -weighted mean of sample who ended up in boxes A+B.

So we can just take a weighted mean (with weights define as above) of the outcomes for those participants falling into the A+B boxes above.

In the next slides we show how to do something equivalent to this using a regression approach.

## A Similar Approach (and SAS Code) Can be Used to Obtain Mean Under AI #2



- Assign  $W = \text{weight} = 2$  to responders to BMOD  $\rightarrow 2 * \frac{1}{2} = 1$
- Assign  $W = \text{weight} = 4$  to non-responders to BMOD  $\rightarrow 4 * \frac{1}{4} = 1$
- Then we take  $W$ -weighted mean of sample who ended up in boxes D+E.

So we can just take a weighted mean (with weights define as above) of the outcomes for those participants falling into D+E boxes above.

In the next slides we show how to do something equivalent to this using a regression approach.

## Results for Estimated Mean Outcome under AI #2

*The SAS System*

*The GENMOD Procedure*

Analysis Of GEE Parameter Estimates			
Parameter	Estimate	Standard Error	P-value
Intercept	3.0982	0.1070	<.0001
Z2	0.4085	0.1070	0.0001

Contrast Estimate Results				
	Mean Estimate	Mean 95% Confidence Limits		Standard Error
Mean Y under AI#2 (BMOD, Add MED)	3.5067	3.1643	3.8490	0.1747

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

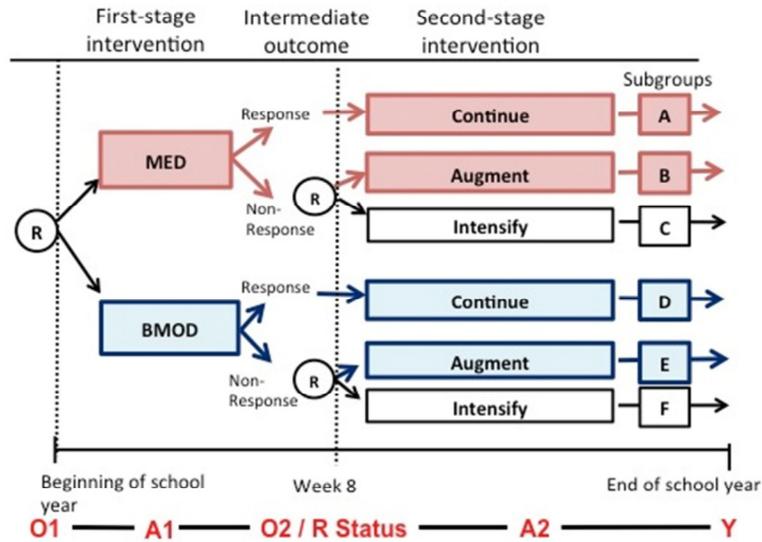
This analysis is with simulated data.

## 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

Outline

## Compare the Mean Outcome under AI#1 vs AI#2 with Just One Regression Analysis



Your initial approach to this comparison might be to just take the mean across participants in the AI#1 sub-groups and compare to the mean outcome of participants in the AI#2 sub-groups. But this approach is not appropriate. This is because, as before, we need to take *weighted* means, and also use robust standard errors.

## SAS Code to Compare the Mean Outcome under AI#1 vs AI#2 with Just One Regression Analysis

```
data dat7; set dat1;  
  if A2=1 then delete;  
  W=2*R + 4*(1-R);  
run;
```

Exclude participants who ended up in boxes C or F. These are the participants who did not respond and received INTENSIFY.

Regression should use information only from those following AI#1 or AI#2!

Assign weights

This analysis is with simulated data.

## SAS Code to Compare the Mean Outcome under AI#1 vs AI#2 with Just One Regression Analysis

```
proc genmod data = dat7;
  class id;
  model y = A1;
  scwgt w;
  repeated subject = id / type = ind;
  estimate 'Mean Y AI#1 (MED, Add BMOD)' intercept 1 A1 -1;
  estimate 'Mean Y AI#2 (BMOD, Add MED)' intercept 1 A1 1;
  estimate 'Diff: AI#1 - AI#2' A1 -2;
run;
```

This analysis is with simulated data.

## Results Comparing Mean Outcome under AI #1 vs AI #2 with Just One Regression Analysis

*The SAS System*

*The GENMOD Procedure*

Analysis Of GEE Parameter Estimates			
Parameter	Estimate	Standard Error	Pr >  Z
Intercept	3.1858	0.1221	<.0001
A1	0.3209	0.1221	0.0086

Contrast Estimate Results				
	Mean Estimate	Mean 95%		Standard Error
		Confidence Limits		
Mean Y under AI#1 (MED, ADD BMOD)	2.8649	2.5305	3.1992	0.1706
Mean Y under AI#2 (BMOD, ADD MED)	3.5067	3.1643	3.8490	0.1747
Diff: AI#1 - AI#2	-0.6418	-1.1203	-0.1633	0.2442

This analysis is with simulated data.

Following AI#2 (BMOD, Add MED) leads to better school performance than following AI#1 (MED, Add BMOD). The difference is statistically significant ( $p$ -value = 0.0086) at 5% Type-I error.

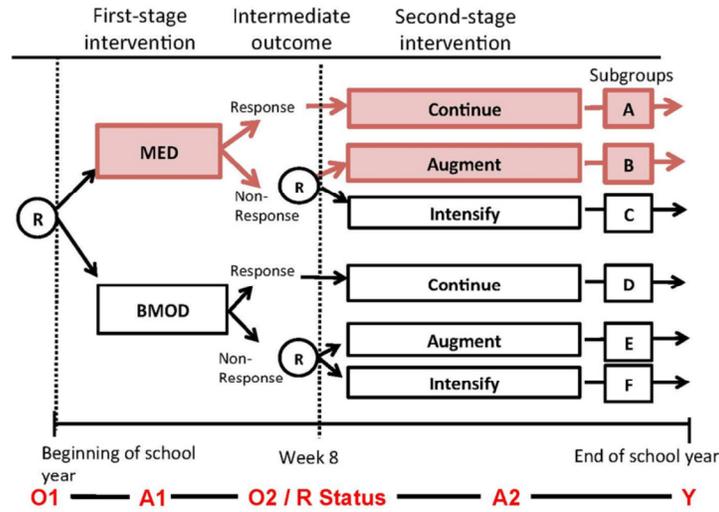
It is also possible to adjust for baseline (pre-A1) covariates in this regression. This usually leads to more efficient (more statistically powerful) comparisons, if the covariate is predictive of the outcome. We discuss this more later...

## 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**

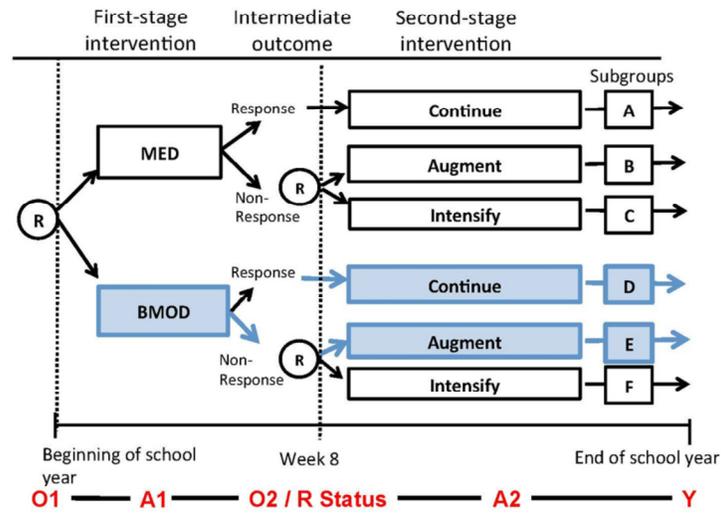
We're now able to pull together our existing knowledge to learn how to simultaneously compare all embedded AIs in a SMART using a "weighted-and-replicated" regression. This is what you'll want to use in your research.

## What About a Regression To Compare AI #1 (MED, Add BMOD) vs...

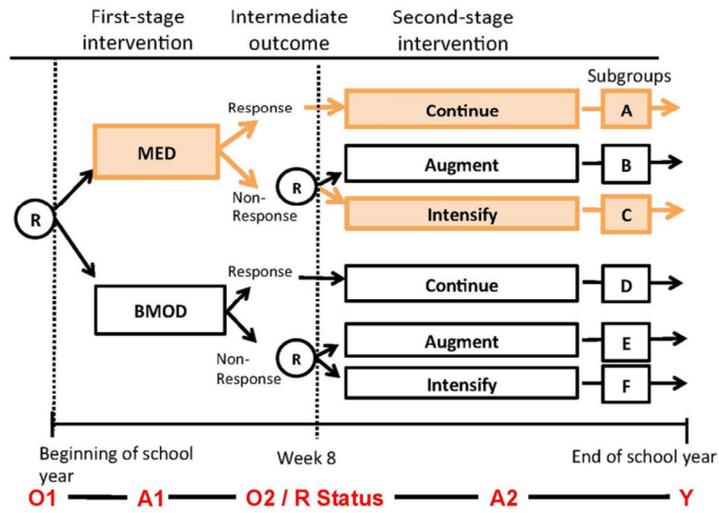


Remember that the AI does not involve randomization.

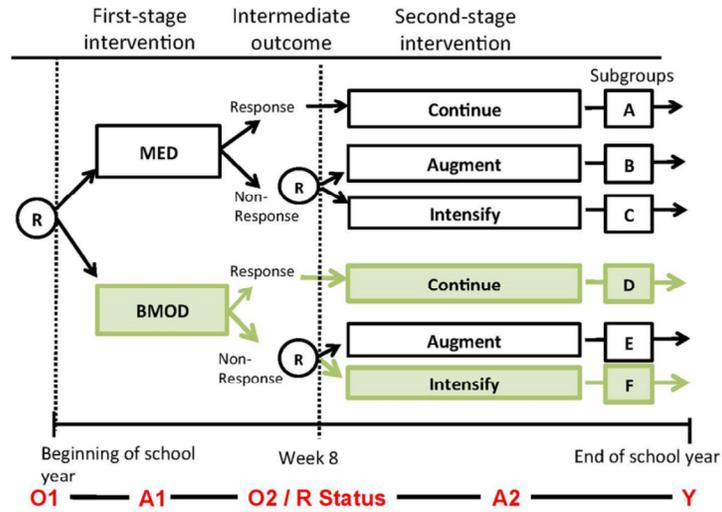
## ...AI #2 (BMOD, Add MED) vs...



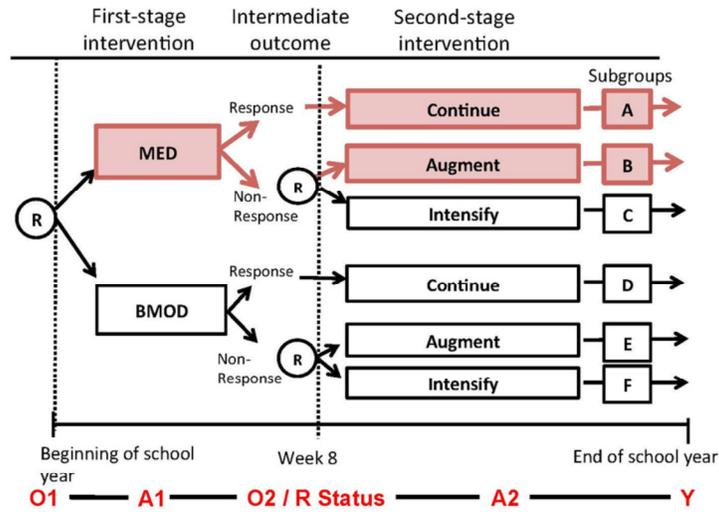
...AI #3 (MED, INTENSIFY) vs...



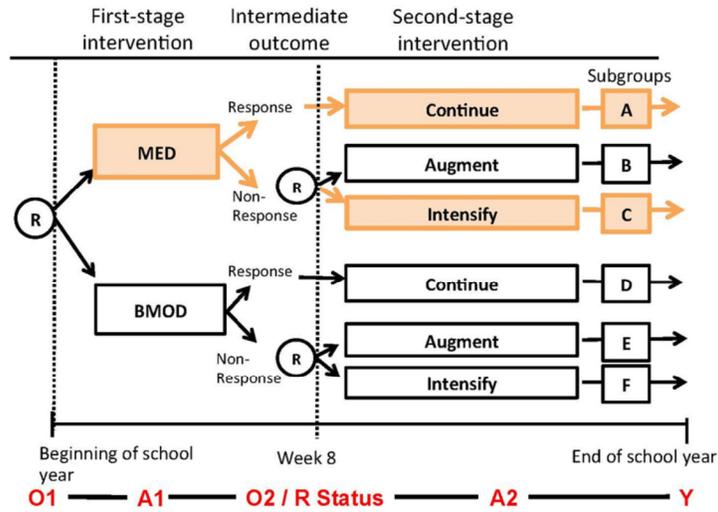
...AI#4 (BMOD, INTENSIFY), all in one swoop!



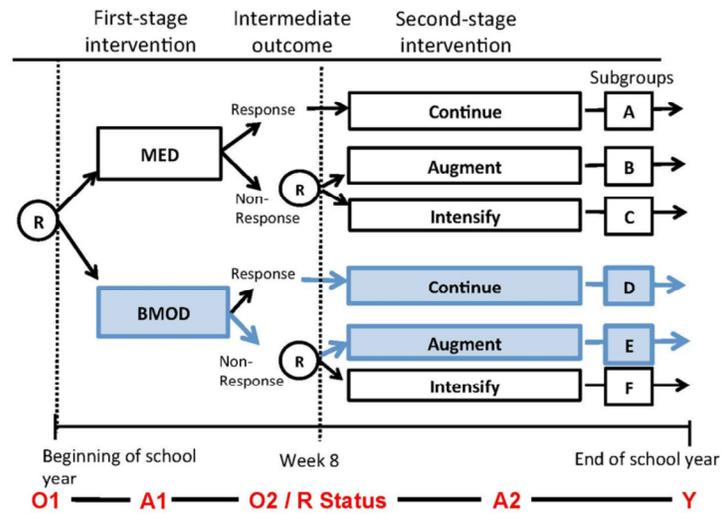
Notice that AI#1 and AI#3 (start MED) share responders (box A)



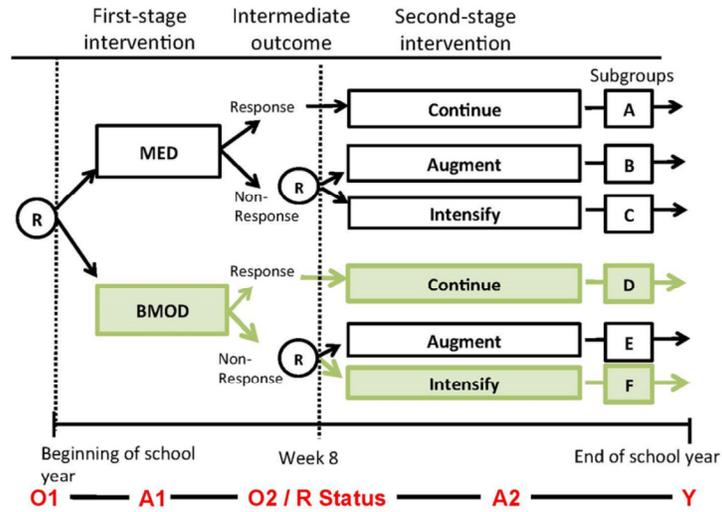
Notice that AI#1 and AI#3 (start MED) share responders (box A)



Similarly: Notice that AI#2 and AI#4 (start BMOD) share responders (box D)



Similarly: Notice that AI#2 and AI#4 (start BMOD) share responders (box D)



## So, what's going on?

- In the ADHD 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  $\frac{1}{2}$  of them  $A_2 = 1$ , the other  $\frac{1}{2}$   $A_2 = -1$
- $W=2$  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 patients may be consistent with more than one AI

## SAS code for replication-and-weighting to compare means under all four ATs

```
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;
```

Basically we require an extra step to replicate observations (i.e., rows in the data set) of responders, such that instead of one observation per responder, there are 2 observations per responder (one with  $A2=1$  and the other with  $A2=-1$ ).

The working intuition is that since a responder's treatment is consistent with this person having been assigned either of two ATs, then we need use each responder's data twice. The first time to estimate the mean for the first AT and the second time to estimate the mean for the second AT.

## Replicated Data

Obs	ID	A1	R	A2	Y	o11c	o12c	o13c	o14c	ob	weight
45	32	1	1	-1	5	-0.35333	-2.73889	-0.31333	0.19333	1	2
46	32	1	1	1	5	-0.35333	-2.73889	-0.31333	0.19333	2	2
47	33	1	0	1	3	0.64667	-1.07820	0.68667	0.19333	1	4
48	34	1	0	1	2	0.35333	0.21667	-0.31333	0.19333	1	4
49	35	1	0	-1	5	-0.35333	-2.73889	-0.31333	0.19333	1	4
50	36	-1	0	1	1	0.64667	-1.07820	0.68667	0.19333	1	4
51	37	-1	1	-1	1	-0.35333	0.99556	-0.31333	0.19333	1	2
52	37	-1	1	1	1	-0.35333	0.99556	-0.31333	0.19333	2	2
53	38	-1	0	-1	3	-0.35333	0.14034	0.68667	-0.80667	1	4
54	39	-1	1	-1	3	0.64667	1.64983	0.68667	0.19333	1	2
55	39	-1	1	1	3	0.64667	1.64983	0.68667	0.19333	2	2

Responders are replicated!

Non-Responders aren't!

Remember that, until now, A2 was not defined for responders. Now, we've replicated our responders and we assign half of the replicates to A2=1 (which makes them consistent with AI#1) and the other half to A2=-1 (which makes them consistent with AI#3). Non-responders are not replicated.

## After Replication-and-Weighting, the SAS code for the weighted regression

```
proc genmod data = dat9;
  class id;
  model y = a1 a2 a1*a2;
  scwgt weight;
  repeated subject = id / type = ind;
  estimate 'MeanY:AI#1 (MED, Add BMOD)' int 1 a1 -1 a2 -1 a1*a2 1;
  estimate 'MeanY:AI#2 (BMOD, Add MED)' 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;
```

Why only four parameters? Because there are only 4 means in total that we wish to estimate (i.e., a saturated regression model).

## After Replication-and-Weighting, the SAS code for the weighted regression

```
proc genmod data = dat9;
  class id;
  model y = a1 a2 a1*a2;
  swgt weight;
  repeated subject = id / type = ind;
  estimate 'MeanY:AI#1 (MED, Add BMOD)' int 1 a1 -1 a2 -1 a1*a2 1;
  estimate 'MeanY:AI#2 (BMOD, Add MED)' 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;
```

## Results for Weighted-and-Replicated Regression: Comparing Mean Outcome for all AIs Simultaneously

Contrast Estimate Results				
Label	Mean Estimate	Mean		Standard Error
		Confidence Limits		
Mean Y: AI#1(MED, Add BMOD)	2.8649	2.5305	3.1992	0.1706
Mean Y: AI#2(BMOD, Add MED)	3.5067	3.1643	3.8490	0.1747
Mean Y: AI#3(MED, INTNSFY)	2.7895	2.4644	3.1145	0.1658
Mean Y: AI#4(BMOD, INTNSFY)	2.6533	2.2515	3.0552	0.2050
Diff: AI#1 - AI#2	-0.6418	-1.1203	-0.1633	0.2442
Diff: AI#1 - AI#3	0.0754	-0.3106	0.4614	0.1969
Diff: AI#1 - AI#4	0.2115	-0.3112	0.7343	0.2667

**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-  
and-replicated regression than just  
convenience!

## Replicated-and-Weighted Regression is More Powerful Statistically

```
proc genmod data = dat9;
  class id;
  model y = a1 a2 a1*a2 O12c O14c;
  scwgt weight;
  repeated subject = id / type = ind;
  estimate `MeanY:AI#1 (MED, Add BMOD)` int 1 a1 -1 a2 -1 a1*a2 1;
  estimate `MeanY:AI#2 (BMOD, Add MED)` 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;
```

Improve power: Adjusting for baseline covariates that are associated with outcome leads to more efficient estimates (lower standard error = more power = smaller p-value).

## Results for Weighted-and-Replicated Regression: Comparing Mean Outcome for all AIs Simultaneously

Improved efficiency: Adjusting for baseline covariates resulted in lower standard error and tighter confidence intervals. Point estimates remained about the same, as expected.

Contrast Estimate Results					
	Mean Estimate	Mean 95%		Standard Error	P-value
		Confidence Limits			
Mean Y: AI#1(MED, Add BMOD)	2.8801	2.5869	3.1733	0.1496	<.0001
Mean Y: AI#2(BMOD, Add MED)	3.3854	3.0689	3.7018	0.1614	<.0001
Mean Y: AI#3(MED, INTNSFY)	2.8149	2.5163	3.1135	0.1524	<.0001
Mean Y: AI#4(BMOD, INTNSFY)	2.7338	2.3596	3.1081	0.1909	<.0001
Diff: AI#1 - AI#2	-0.5053	-0.9401	-0.0704	0.2219	0.0228
Diff: AI#1 - AI#3	0.0652	-0.2811	0.4115	0.1767	0.7120
Diff: AI#1 - AI#4	0.1463	-0.3271	0.6107	0.2415	0.5447

Standard error in unadjusted model was 0.2442

This analysis is with simulated data.

## Citations

- Murphy, S. A. (2005). An experimental design for the development of adaptive intervention. *Statistics in Medicine*, 24, 455-1481.
- Nahum-Shani, I., Qian, M., Almirall, D., Pelham, W. E., Gnagy, B., Fabiano, G. A., ... & Murphy, S. A. (2012). Experimental design and primary data analysis methods for comparing adaptive interventions. *Psychological methods*, 17(4), 457.