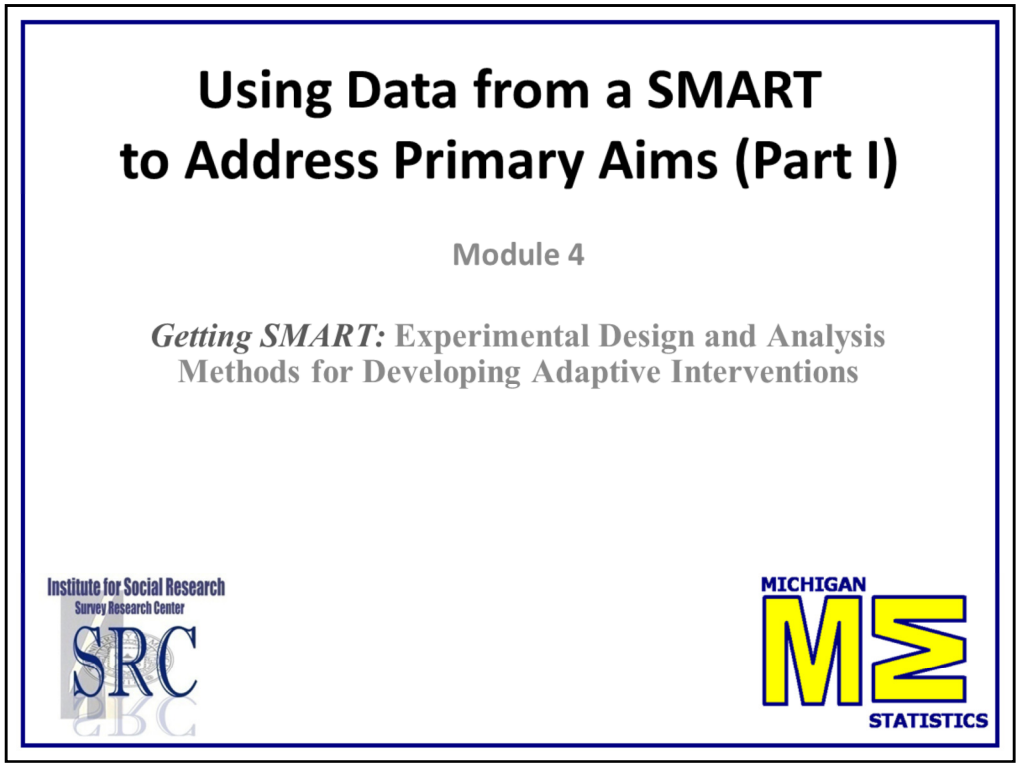


By now, you know what an AI is. You have discussed why they are important in terms of managing chronic disorders (indeed, an ATS formalizes the type of clinical practice taking place today). And, you have been introduced to the SMART clinical trial design, the rationale for SMARTs, and some important 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)

Outline

- Review of ADHD SMART study
- Learn how to analyze two typical primary research questions in a SMART design
  - (a): Main effect of first-stage options
  - (b): Main effect of second-stage options/tactics
- Prepare for a third primary aim analysis by
  - (c): Learning to estimate the mean outcome under each of the embedded AIs (separately) using an easy-to-use weighting approach.
- Summary & Discussion

Note about SAS code

- We will show you how to use SAS to preform analyses
- No time to go over the entire code and try it out.
- Step by step instructions in file:

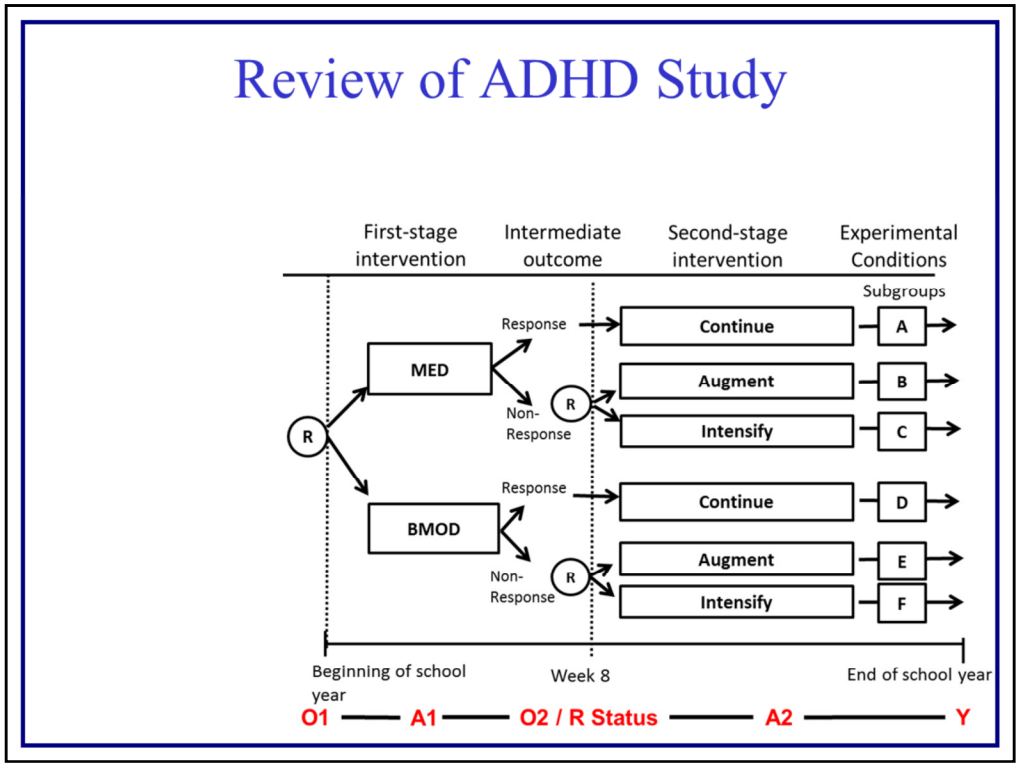
  "sas_code_modules_4_5_and_6_ADHD.doc"
- Try code on fake data:

  "ADHD_simulated_data.sas7bdat"
- Detailed description of data in:

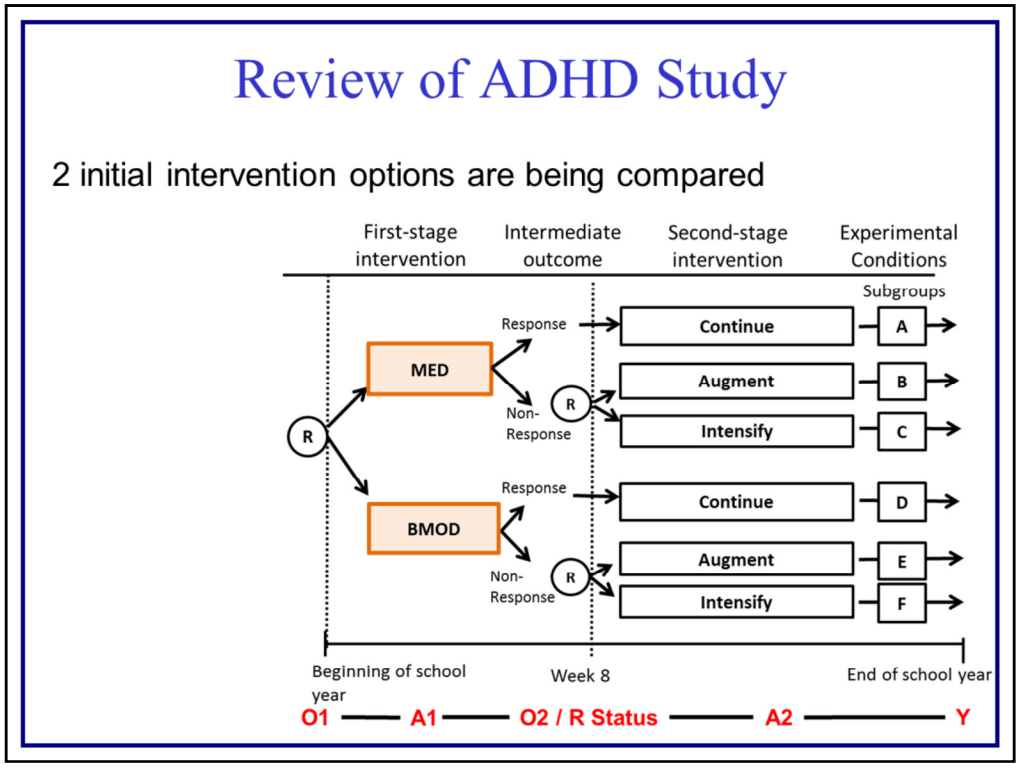
  "ADHD Data Description Handout.pdf"

Outline

- **Review of ADHD SMART study**
- Learn how to analyze two typical primary research questions in a SMART design
  - (a): Main effect of first-stage options
  - (b): Main effect of second-stage options/tactics
- Prepare for a third primary aim analysis by
  - (c): Learning to estimate the mean outcome under each of the embedded AIs (separately) using an easy-to-use weighting approach.
- Summary & Discussion



Review the characteristics of this SMART design

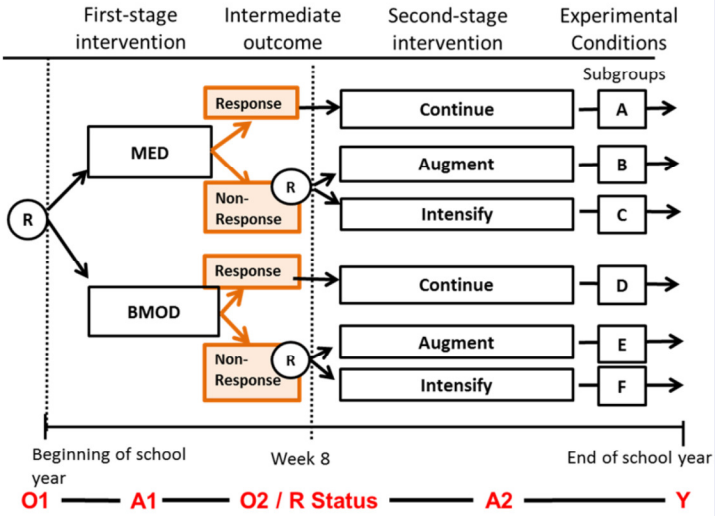


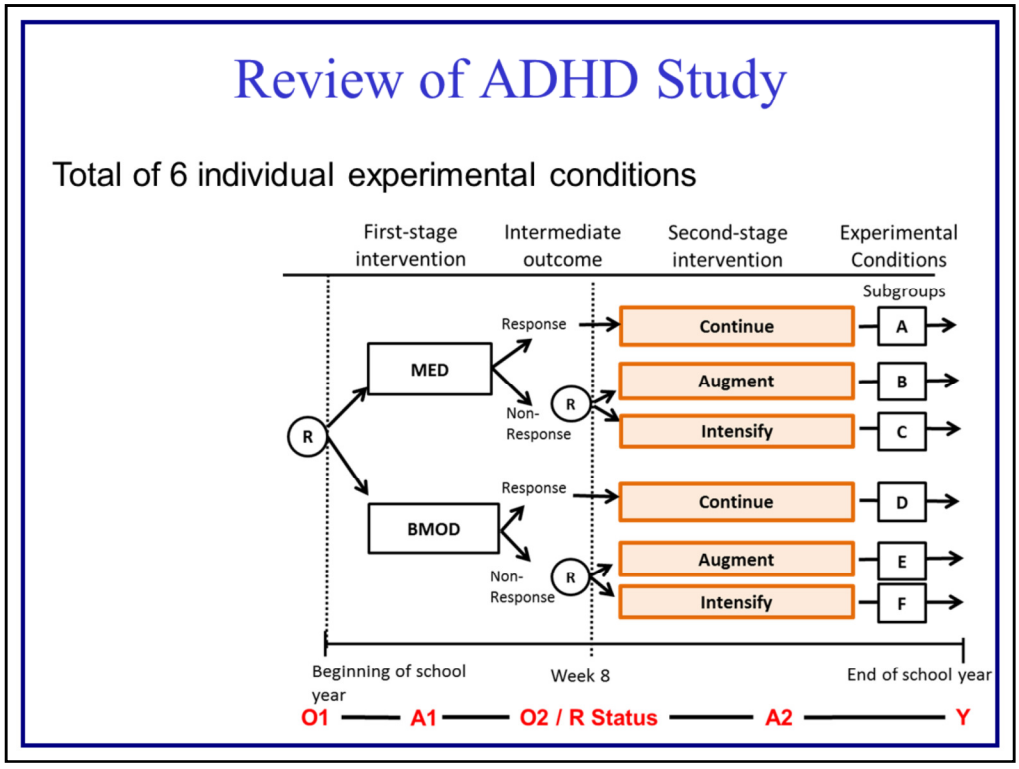
MED is Ritalin.

BMOD is behavioral modification, itself a multi-component behavioral intervention.

Review of ADHD Study

Embedded tailoring variable: Response status





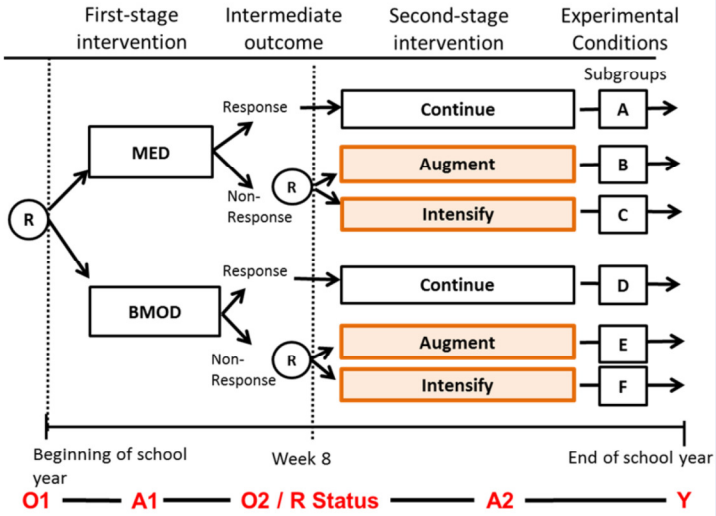
The interventions include differing doses of MED methylphenidate (a psychostimulant drug) and differing intensities of behavioral modification BMOD (consisting of a school-based component with the teacher, a Saturday treatment component involving social skills development, and a parent-training component targeted at helping parents to identify problematic behaviors with the relevant child-functioning domains).

Intensified MED: The higher-dose option for methylphenidate includes late-afternoon doses, if needed.

Intensified BMOD: The higher-intensity option for the behavioral modification includes more intensive training in social skills in the school-based component and, if needed, both additional individual parent training sessions that target specific behavior management issues and practice sessions with children.

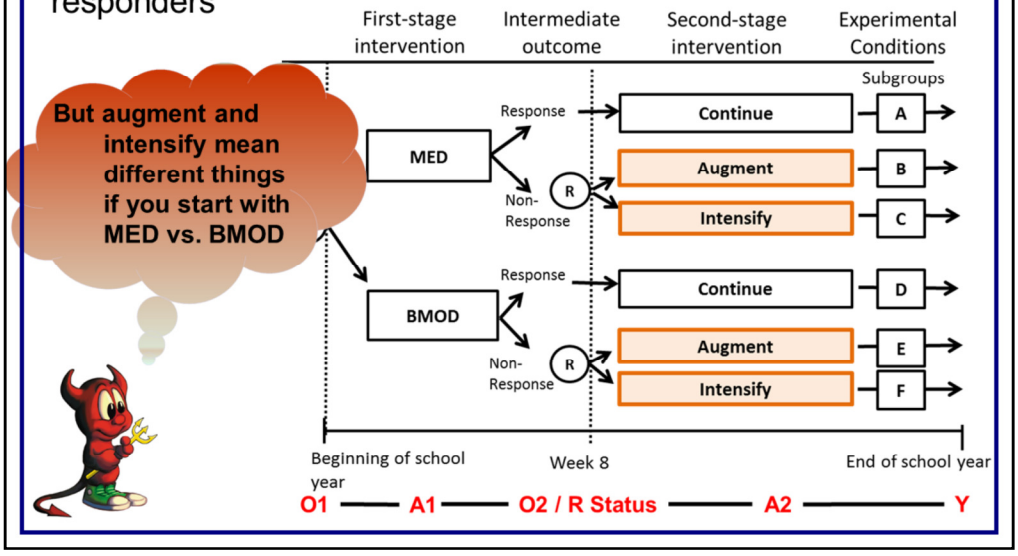
Review of ADHD Study

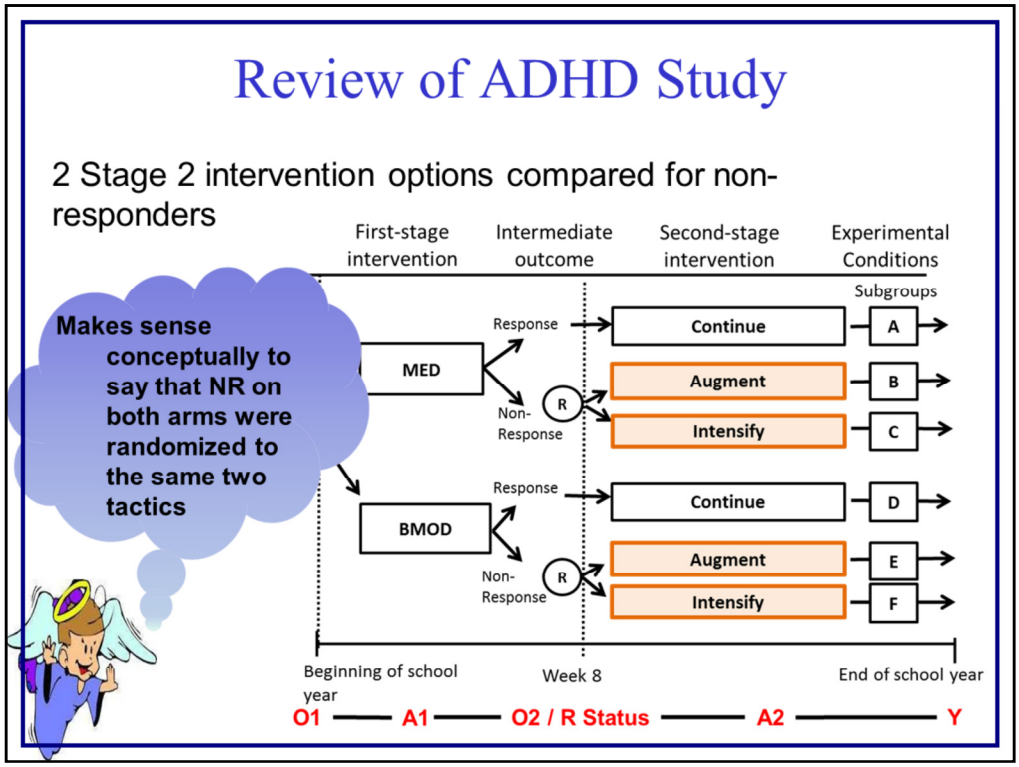
2 second stage intervention options compared for non-responders



Review of ADHD Study

2 Stage 2 intervention options compared for non-responders





Of course, augment and intensify mean different things for children who started with MED vs. BMOD.

But scientifically, augment and intensify are considered two tactical decisions: providing more MED or more BMOD is the same tactical decision (= Intensify) from a scientific/practical point of view.

So, conceptually, it makes sense to say that non-responders on both arms were randomized to the same two subsequent tactics.

Review of ADHD Study

4 embedded AIs: #1

At the beginning of school year

Stage 1 = {**MED**},

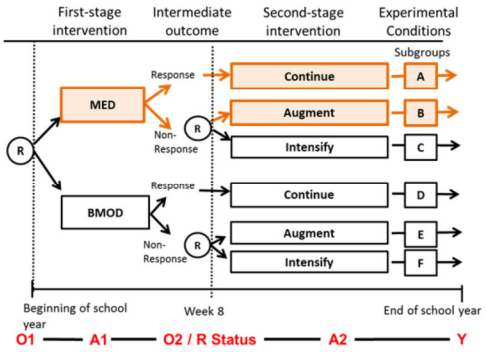
Then, every month,
starting at week 8

IF response status = {NR}

THEN Stage 2 = {**AUGMENT**}

ELSE IF response status = {R}

THEN **CONTINUE** Stage 1



Notice, AI is not randomized; it is a recommended decision rule

Notice that AIs are not randomized, it is a recommended policy– a recommended decision rule.

Review of ADHD Study

4 embedded AIs: #2

At the beginning of school year

Stage 1 = {**BMOD**},

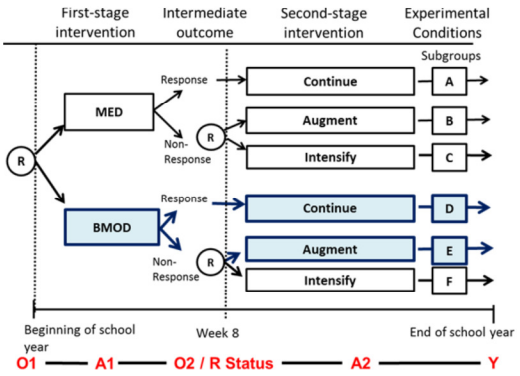
Then, every month,
starting at week 8

IF response status = {NR}

THEN Stage 2 = {**AUGMENT**}

ELSE IF response status = {R}

THEN **CONTINUE** Stage 1



Review of ADHD Study

4 embedded AIs: #3

At the beginning of school year

Stage 1 = {**MED**},

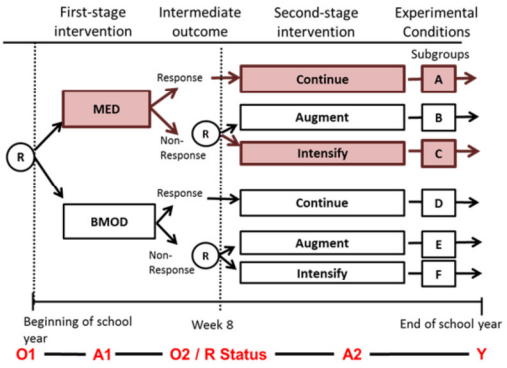
Then, every month,
starting at week 8

IF response status = {NR}

THEN Stage 2 = {**INTENSIFY**}

ELSE IF response status = {R}

THEN **CONTINUE** Stage 1



Review of ADHD Study

4 embedded AIs: #4

At the beginning of school year

Stage 1 = {**BMOD**},

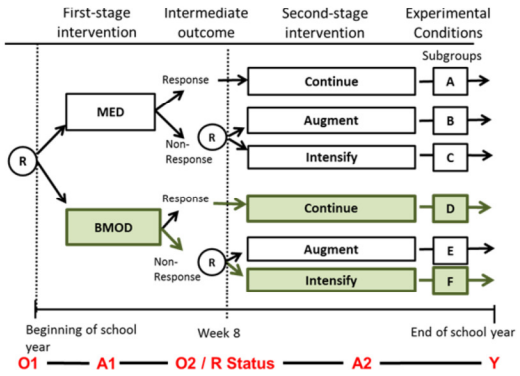
Then, every month,
starting at week 8

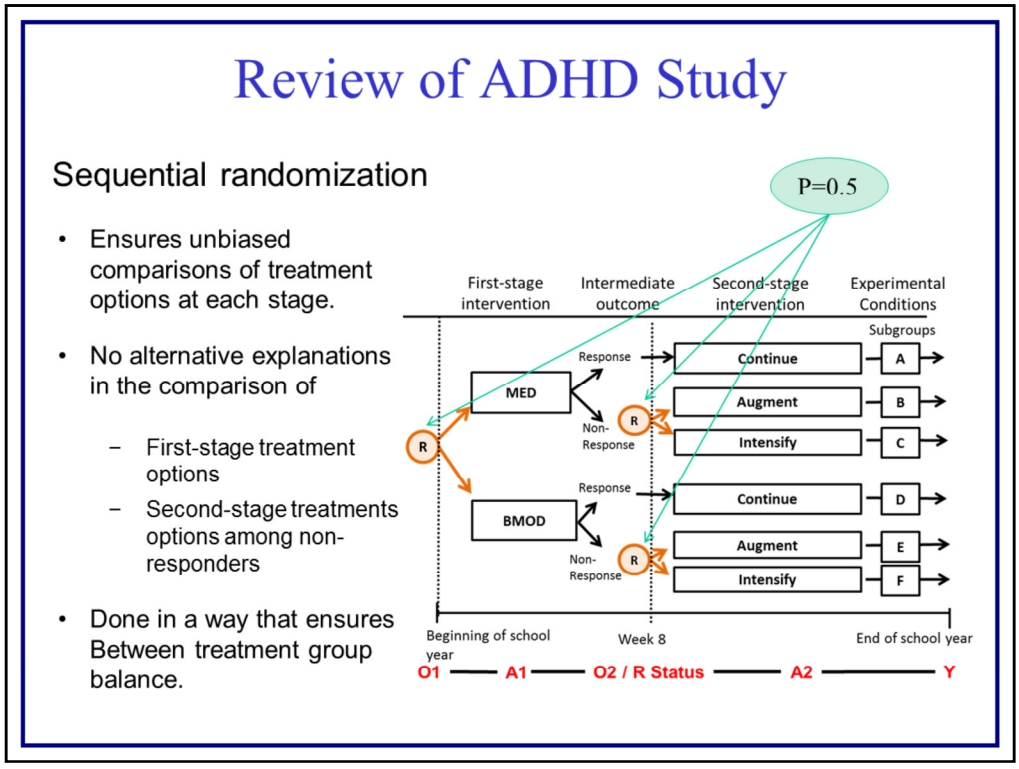
IF response status = {NR}

THEN Stage 2 = {**INTENSIFY**}

ELSE IF response status = {R}

THEN **CONTINUE** Stage 1





The sequential randomizations ensure unbiased comparisons (no alternative explanations) between assigned treatments both initially (at the first line) and in the future (at the second stage) among non-responders.

Example of alternative explanation for stage 1: in the real-world, kids with high severity symptoms might be more likely to receive mediation, but these kids are also likely to fail at the end of treatment. Hence the results will indicate that MED is less effective than BMOD, but we observe this not because MED is in fact less effective, but because it was given to kids with high symptom severity which are less likely to improve.

Example of alternative explanation for stage 2: in the real-world, non-responders who do not adhere to stage 1 are likely to get augment rather than intensify. But its really difficult to re-motivate non-adherent and hence they are likely to fail at the end of the school year. What we see at the end is that augment is a less beneficial second-stage tactic for non-responders. However, we observe this not because augment is in fact less effective, but because augment was given to non-adherent who are highly likely to fail.

Sequential randomizations are done in a way that ensures between treatment group balance, like in a factorial design. Balanced in a sense that within each initial option I have ½ of the non-responders on augment and ½ of the non-responders on intensify.

This makes this design efficient because when comparing stage 2 options for non-responders I can compare half of non-responders vs. the other half. Also, half of the sample received MED and half BMOD, so the comparison of first stage options can be done by comparing ½ of the sample vs. the other ½.

Review of ADHD Study

What the data looks like; ***Part 1***

	ODD at baseline?	Baseline ADHD Score	Prior Med?	Stage 1 Option	Response/ Non-Response	Stage 2 Tactic	School Perfm
ID	O11	O12	O13	A1	R	A2	Y
1	1 (YES)	1.18	0 (NO)	-1 (MED)	1 (R)	.	3
2	0 (NO)	-0.567	0	-1	0 (NR)	1 (INTSFY)	4
3	0	0.553	1 (YES)	1 (BMOD)	0	-1 (ADD)	4
4	0	-0.013	0	1	0	-1	4
5	0	-0.571	1	1	0	1	2
6	0	-0.684	1	1	0	-1	4
7	0	1.169	0	-1	1	.	3

***This data is simulated

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 (Oppositional defiant behavior– yes or no– whether the child was diagnosed with ODD at baseline) ODD (disobedient/hostile pattern of behavior towards authority)

O12 = baseline ADHD scores (based on teacher evaluation from previous school year– standardized)

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

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.

Review of ADHD Study

What the data looks like; ***Part 2***

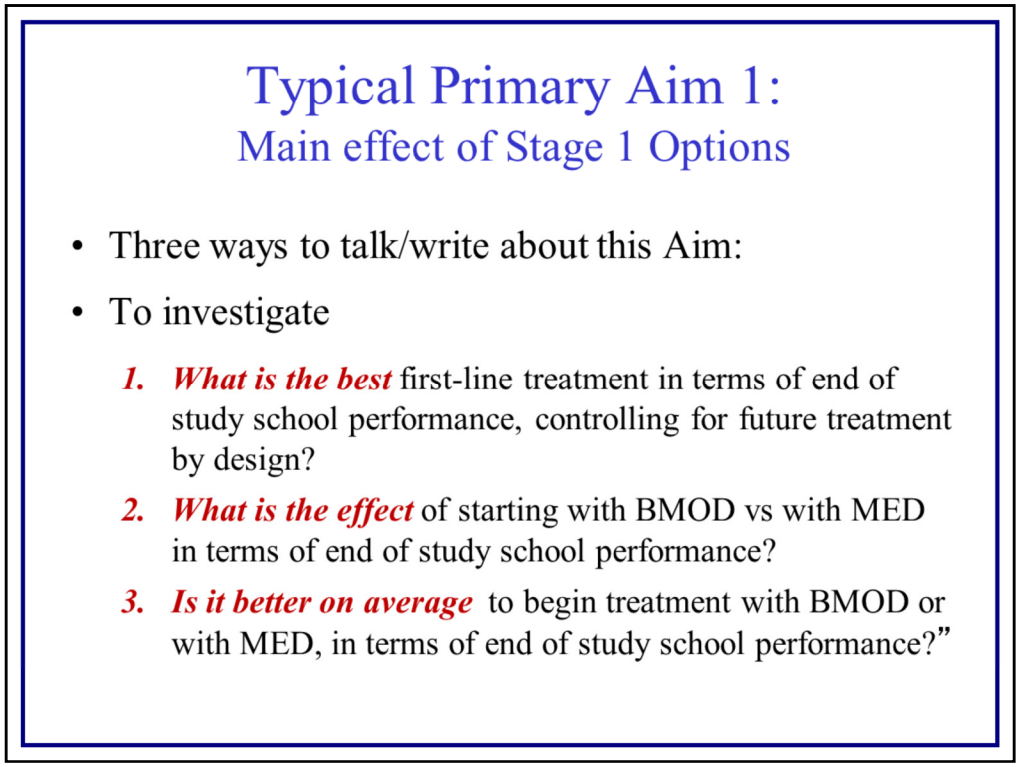
	Race	Stage 1 Option	Response/ Non-Response	Time until NR (months)	Adherence	Stage 2 Tactic	School Perfm
ID	O14	A1	R	O21	O22	A2	Y
1	1 (White)	-1 (MED)	1 (R)	.	0 (NO)	.	3
2	0 (Other)	-1	0 (NR)	6	0	1 (INTSFY)	4
3	0	1 (BMOD)	0	1	1 (YES)	-1 (ADD)	4
4	0	1	0	7	0	-1	4
5	0	1	0	5	1	1	2
6	0	1	0	3	1	-1	4
7	0	-1	1	.	0	.	3

***This data is simulated

In addition to R, there can be other covariates measured after A1 but before A2, such as
O21 = Time in months until non-response (only measured for those with R=0)
O22 = Adherence to first-line treatment = YES(1) or NO(0).

Outline

- Review of ADHD SMART study
- **Learn how to analyze two typical primary research questions in a SMART design**
  - **(a): Main effect of first-stage options**
  - (b): Main effect of second-stage options/tactics
- Prepare for a third primary aim analysis by
  - (c): Learning to estimate the mean outcome under each of the embedded AIs (separately) using an easy-to-use weighting approach.
- Summary & Discussion



These are different ways to talk/write about Typical Primary Question #1:

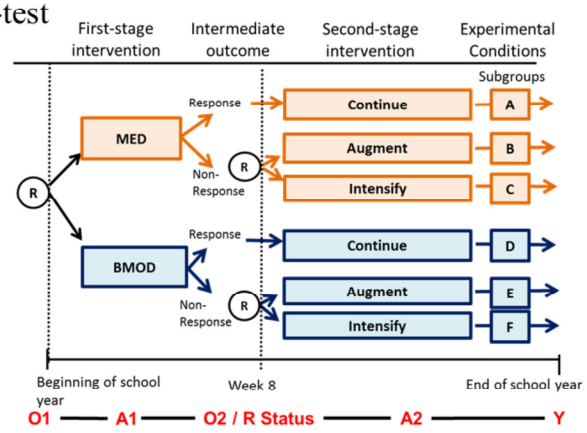
On average, how do longitudinal outcomes differ between children assigned first to medication versus children assigned first to behavioral intervention?

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?

For more information about comparison of growth curves, see book (Longitudinal Data Analysis) by Donald Hedeker, Robert D. Gibbons

Typical Primary Aim 1:
Main effect of Stage 1 Options

- Simply a comparison of two groups:
  - A two-sample t-test

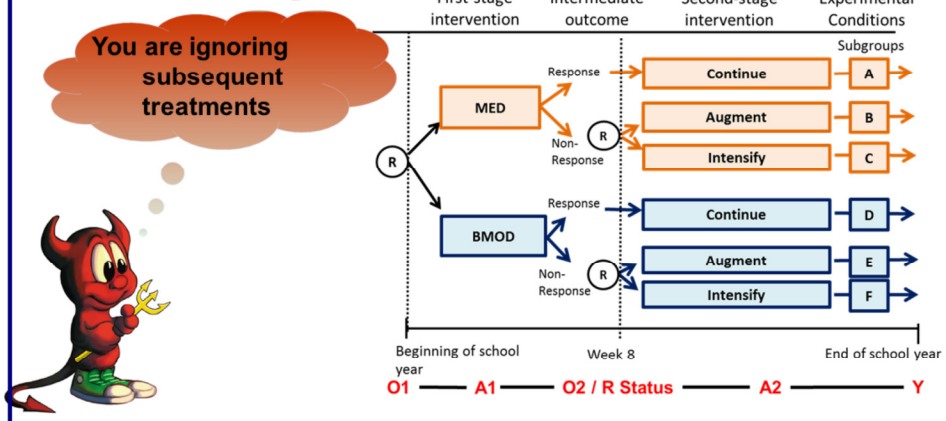


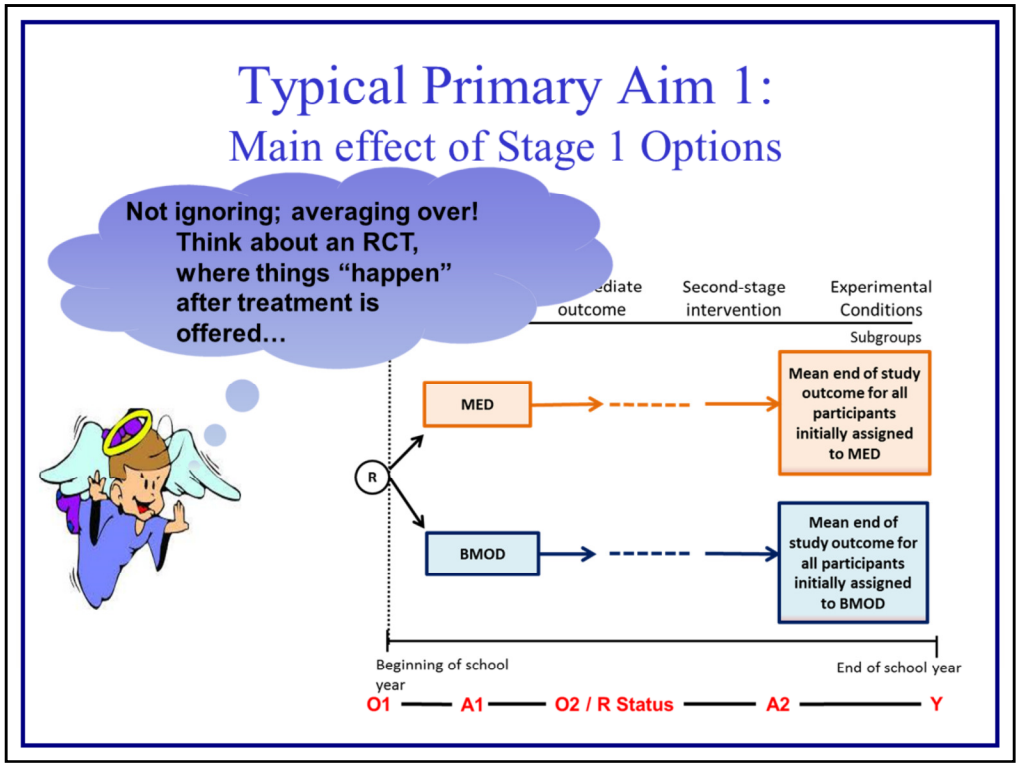
Given a continuous, end of study (e.g., 8 months) outcome, then a two-sample t-test is all that is needed.

This is just a comparison of two groups of study participants (the blue participants versus the orange participants).

Typical Primary Aim 1:
Main effect of Stage 1 Options

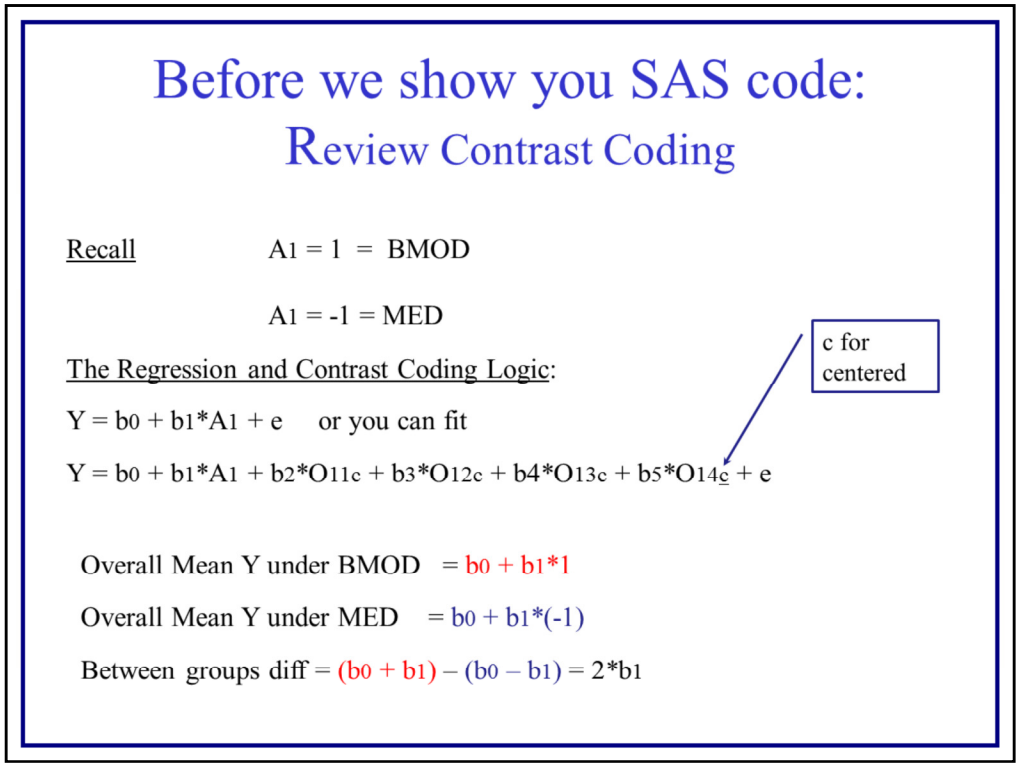
- Simply a comparison of two groups:
  - A two-sample t-test



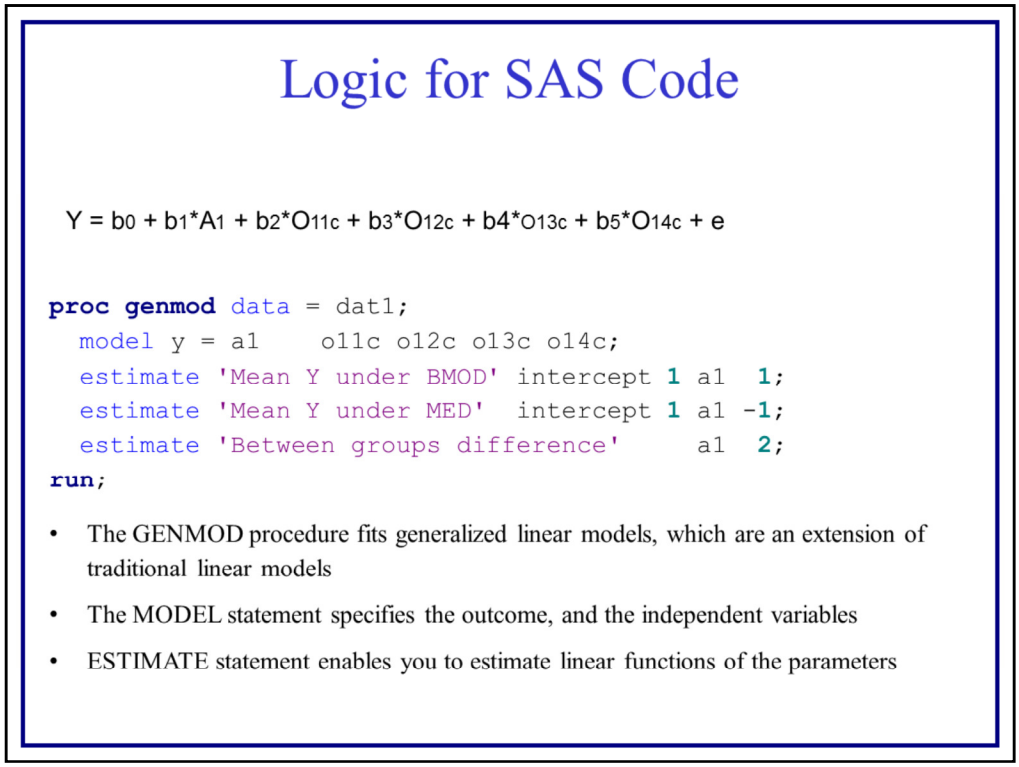


The way to think about this is to think for the moment of the 2 arm RCT and imagine that even in those studies "we do things" or "things happen" even after we offer treatment.

This is no different from a typical intent to treat approach (where the results of an experiment is based on the initial treatment assignment and not on the treatment eventually received). But here, it is more like "we do things" because we actually control the future treatments by design.



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.



This analysis is on Page 2 of your SAS code Word document.

The GENMOD procedure fits generalized linear models, which are an extension of traditional linear models that allows the mean of a population to depend on a linear predictor through a nonlinear link function.

Logic for SAS Code

```
proc genmod data = dat1;
  model y = a1    o11c o12c o13c o14c;
  estimate 'Mean Y under BMOD' intercept 1 a1  1 o11c 0;
  estimate 'Mean Y under MED'  intercept 1 a1 -1;
  estimate 'Between groups difference'     a1  2;
run;
```

The Regression Logic:

$Y = b_0 + b_1*A_1 + b_2*O_{11c} + b3*O_{12c} + b_4*O_{13c} + b_5*O_{14c} + e$

Mean Y under BMOD $= E(Y \mid A_1 = 1) = b_0 + b_1*1$

Mean Y under MED $= E(Y \mid A_1 = -1) = b_0 + b_1*(-1)$

Between groups diff $= E(Y \mid A_1 = 1) - E(Y \mid A_1 = 1)$

$= (b_0 + b_1) - (b_0 - b_1) = 2*b_1$

In ESTIMATE statements, If I leave a coefficient blank, it means I set it to zero.

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!

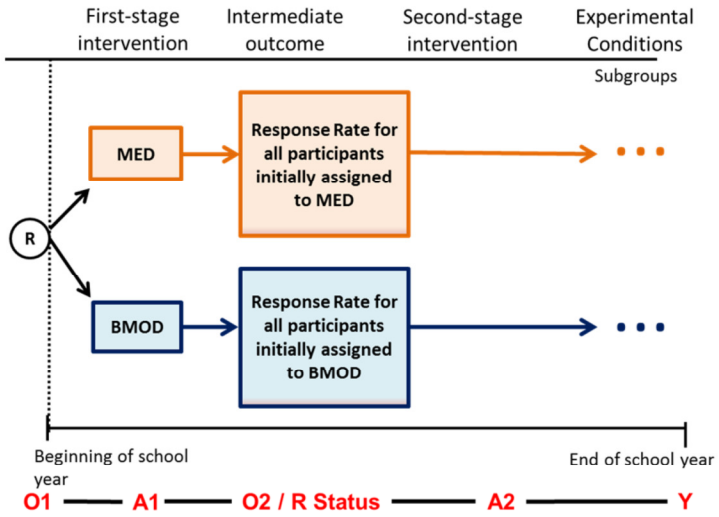
Aim 1 Results

Contrast Estimate Results

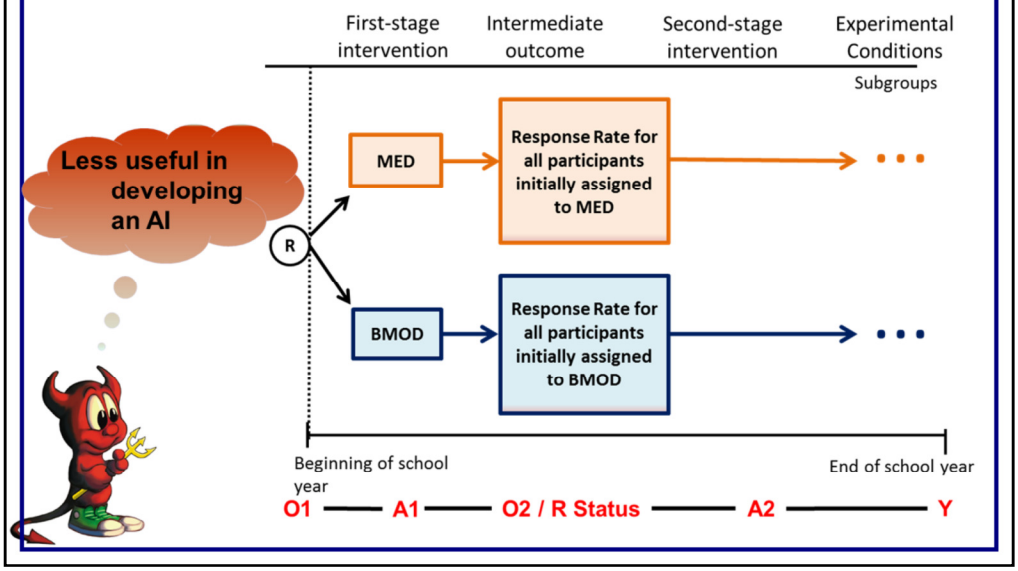
Label	Estimate	95% Conf Limits Lower	Upper	P-value
Mean Y under BMOD	3.0459	2.7859	3.3059	<.0001
Mean Y under MED	2.8608	2.6008	3.1208	<.0001
Between groups diff	**0.1851**	-0.1849	0.5551	**0.3269**
(SE = standard err)	**(0.1889)**			

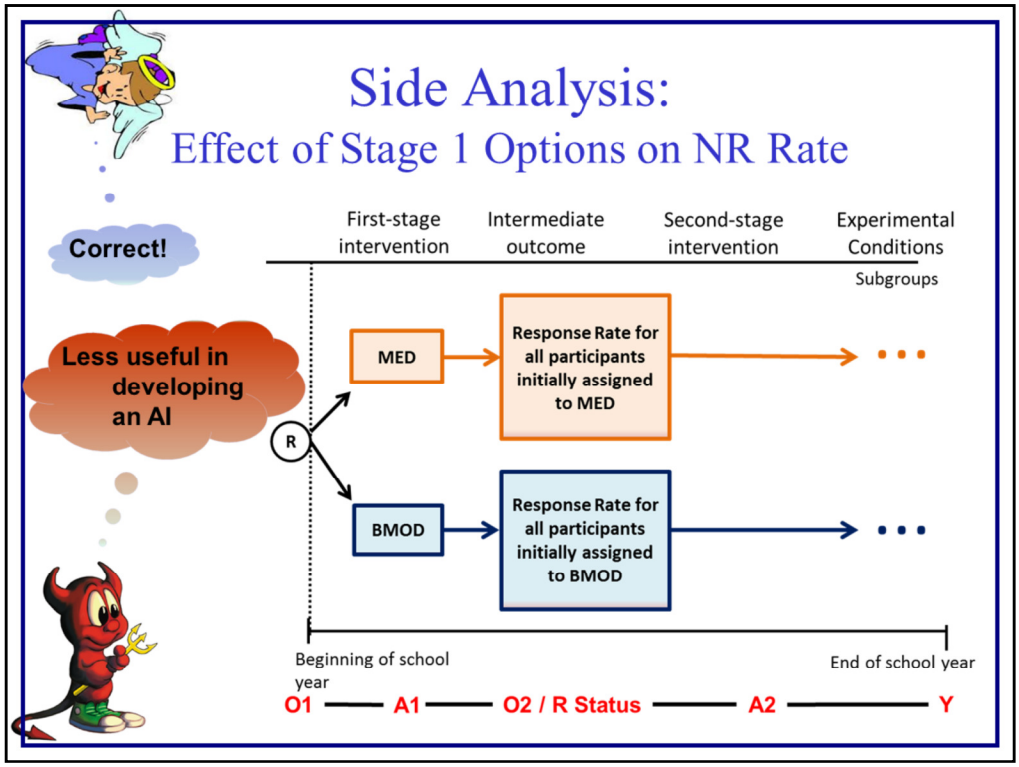
- Results are from simulated dataset
- Slightly better to begin with BMOD (vs MED) in terms of school performance at end of study, but not statistically significant (p-value = 0.33).

Side Analysis:
Effect of Stage 1 Options on NR Rate



Side Analysis:
Effect of Stage 1 Options on NR Rate





This is NOT a primary aim. But useful nonetheless.

Note that this analysis is less useful in terms of building an AI 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.

Results of Side Analysis

Effect of Stage 1 Options on NR Rate

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

Frequency Raw pct	0 (Non-Response)	1 (Response)	Total
-1 (MED)	47 62.67%	28 37.33%	75
1 (BMOD	52 69.33%	23 30.67%	75
Total	99	51	150

In terms of early NR rate, initial MED is slightly better (vs. BMOD) by 7%, but *NS* (p-value = 0.39) .

This analysis is on Page 3 of your SAS code Word document.

NOPERCENT: Suppresses display of percentages

NOCOL: suppresses the display of column percentages in cross-tabulation table cells.

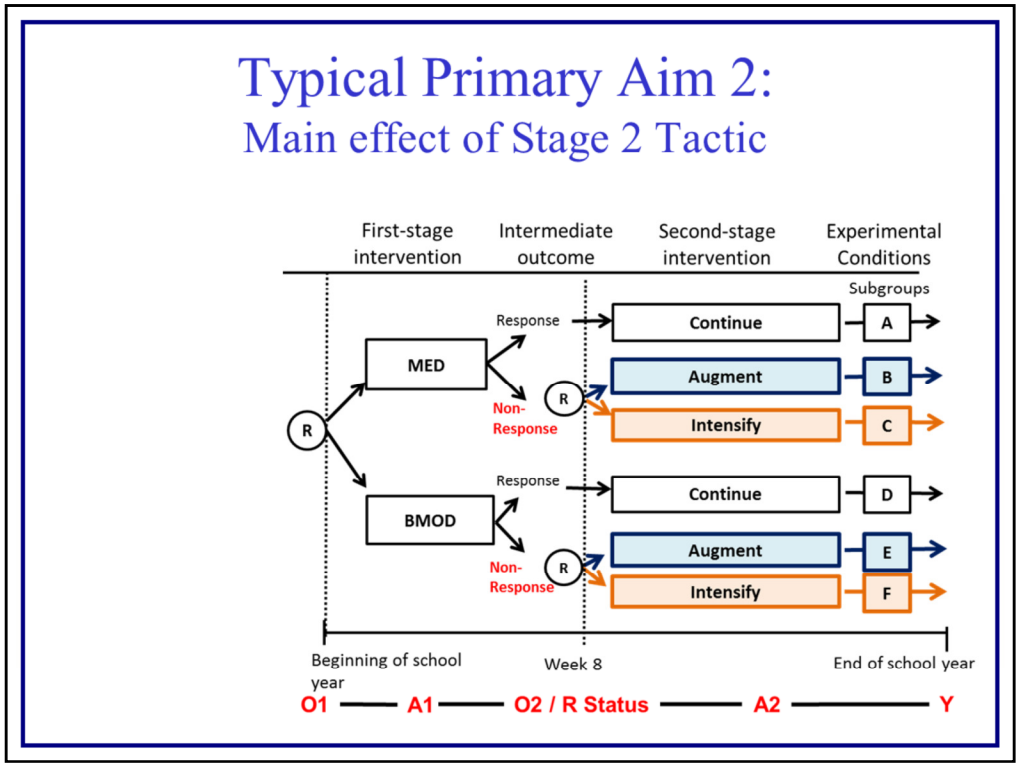
Outline

- Review of ADHD SMART study
- **Learn how to analyze two typical primary research questions in a SMART design**
  - (a): Main effect of first-stage options
  - (b): **Main effect of second-stage options/tactics**
- Prepare for a third primary aim analysis by
  - (c): Learning to estimate the mean outcome under each of the embedded AIs (separately) using an easy-to-use weighting approach.
- Summary & Discussion

Typical Primary Aim 2:
Main effect of Stage 2 Tactic

How to talk about this Aim?

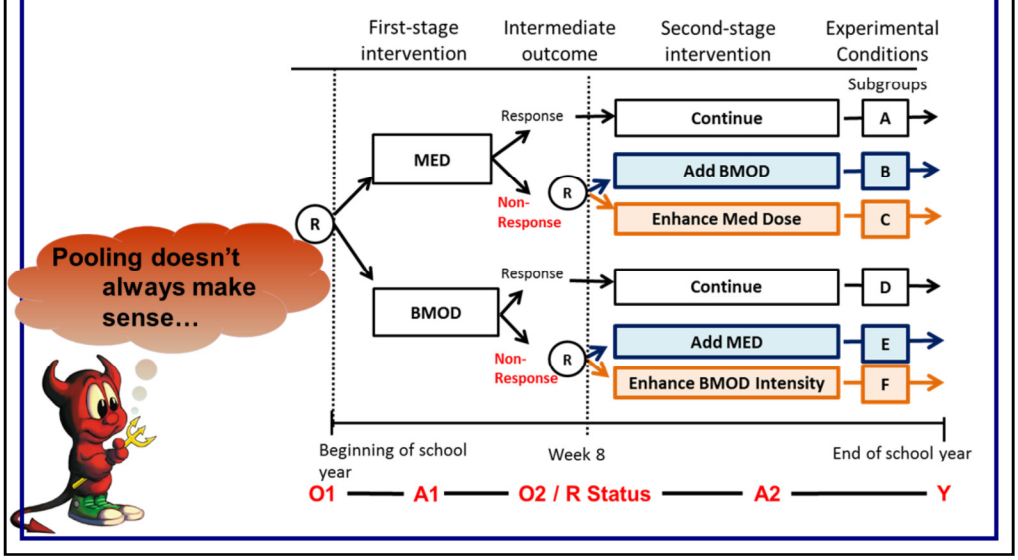
- To investigate whether, among children who do not respond to (either) first-line treatments, it is better to **intensify** or **augment** the initial treatment
  - Regardless of history of treatment
  - Controlling for first-stage intervention options

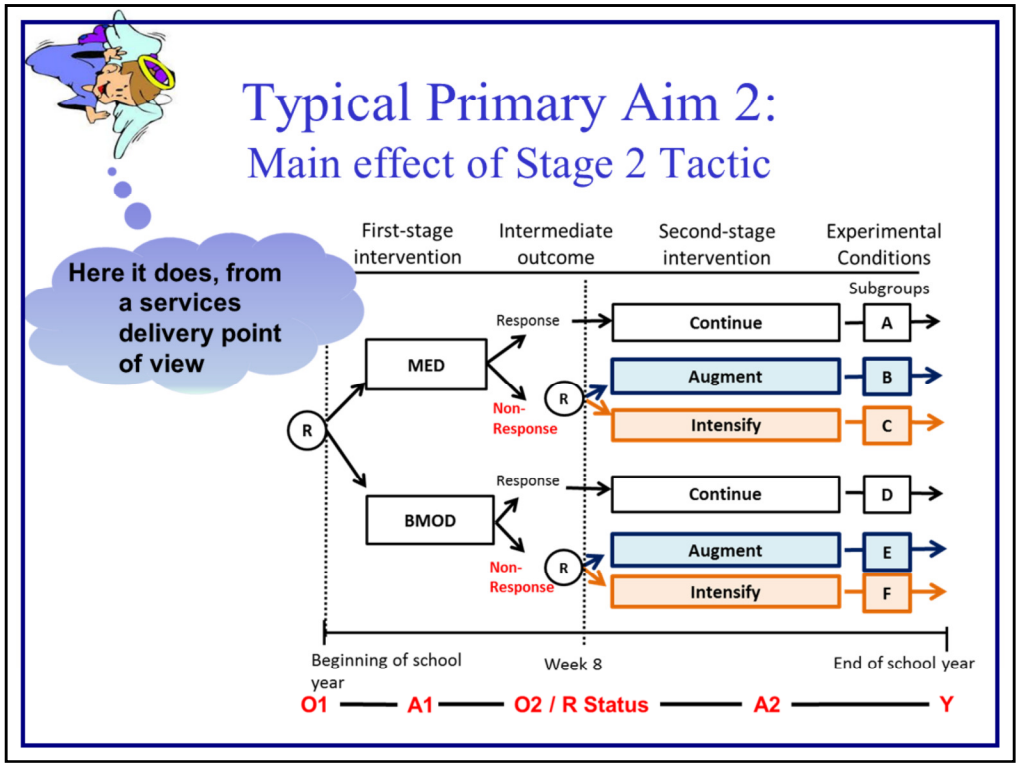


This is not a comparison of adaptive intervention, per se. Rather it informs the tactical decision often made in clinical practice of whether to add to the treatment with something new versus increase the dosage/intensity of treatment.

Note that this is a comparison of the blue cells versus the orange cells, pooled over (averaged over) first-line. The pooling leads to more power (i.e., larger sample size for the comparison of tactics).

Typical Primary Aim 2:
Main effect of Stage 2 Tactic




Typical Primary Aim 2:
Main effect of Stage 2 Tactic
Here it does, from a services delivery point of view
First-stage intervention
Intermediate outcome
Second-stage intervention
Experimental Conditions
Subgroups
R
MED
Response
Non-Response
R
Continue
A
Augment
B
Intensify
C
BMOD
Response
Non-Response
R
Continue
D
Augment
E
Intensify
F
Beginning of school year
Week 8
End of school year
O1
A1
O2 / R Status
A2
Y

SAS Code for Aim 2 Analysis

```
* use only non-responders;
data dat3;
  set dat2; if R=0;
run;

* run the regression;
proc genmod data = dat3;
        model y = a2     o11c o12c o13c o14c o21c o22c;
        estimate 'Mean Y w/INTENSIFY tactic' intercept 1 a2  1;
        estimate 'Mean Y w/ADD tactic'       intercept 1 a2 -1;
       estimate 'Between groups difference'             a2  2;
run;
```

This analysis is on Page 4 of your SAS code Word document.

Recall: A2 is coded 1 for intensity and -1 for augment The Regression Logic:

Logic for SAS Code

```
proc genmod data = dat3;
   model y = a2      o11c o12c o13c o14c o21c o22c;
   estimate 'Mean Y w/INTENSIFY tactic' intercept 1 a2  1;
   estimate 'Mean Y w/ADD TXT tactic'   intercept 1 a2 -1;
    estimate 'Between groups difference'             a2  2;
run;
```

The Regression Logic:

$$Y = b_0 + b_1*A_2 + b_2*O_{11c} + b_3*O_{12c} + b_4*O_{13c} + b_5*O_{14c} + b_6*O_{21c} + b_6*O_{22c} + e$$

Mean Y under INTENSIFY $= E(Y \mid A_2 = 1) = b_0 + b_1*1$

Mean Y under AUGMENT/ADD $= E(Y \mid A_1 = -1) = b_0 + b_1*(-1)$

Between groups diff $= E(Y \mid A_2 = 1) - E(Y \mid A_2 = -1)$

$= (b_0 + b_1) - (b_0 - b_1) = 2*b_1$

Aim 2 Results

Contrast Estimate Results

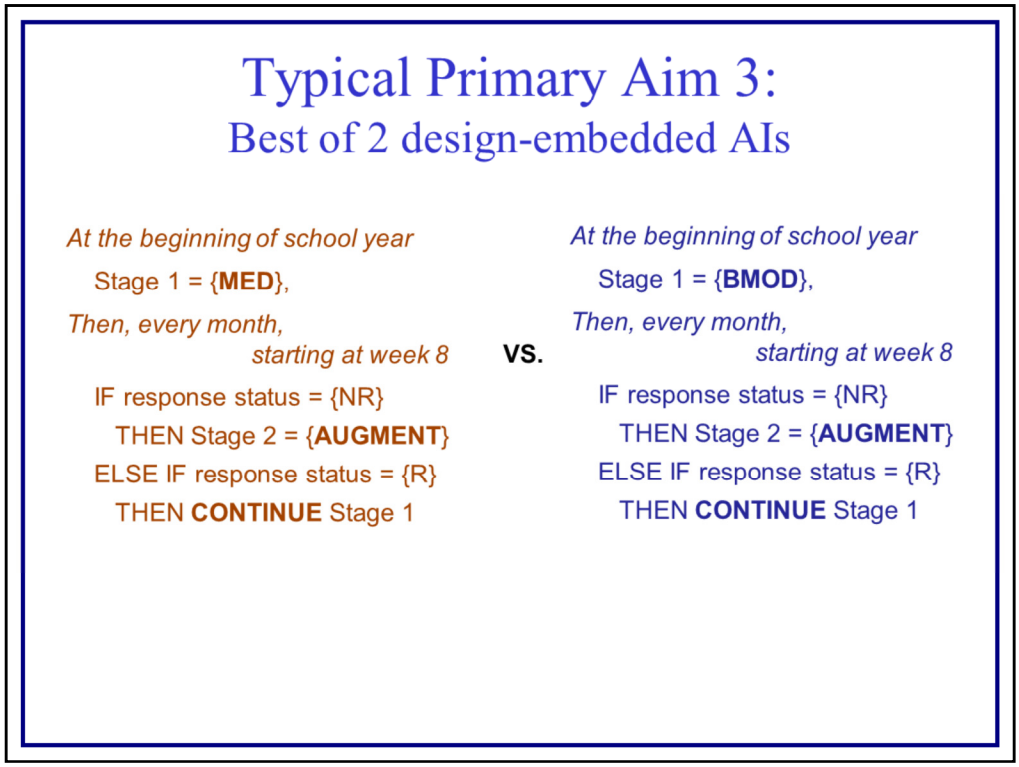
Label	Estimate	95% Conf Limits Lower	Upper	P-value
Mean Y w/INTENSIFY tactic	2.6064	2.3055	2.9072	<.0001
Mean Y w/ADD tactic	3.1942	2.8904	3.4981	<.0001
Between groups difference	**-0.5879**	-1.0206	-0.1552	**0.0077**
(SE = standard error)	**(0.2208)**			

- Results are from simulated dataset
- On average, Augment is a better tactic (vs. Intensify) for non-responders to either MED or BMOD in terms of school performance at end of study.
- Difference is statistically significant

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.

Outline

- Review of ADHD SMART study
- Learn how to analyze two typical primary research questions in a SMART design
  - (a): Main effect of first-stage options
  - (b): Main effect of second-stage options/tactics
- **Prepare for a third primary aim analysis by**
  - **(c): Learning to estimate the mean outcome under each of the embedded AIs (separately) using an easy-to-use weighting approach.**
- Summary & Discussion



This primary aim is a comparison of 2 adaptive intervention 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 AIs. Here we chose 1 pair for illustration.

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

- How to talk about this Aim?

To investigate whether, an AI that recommends to

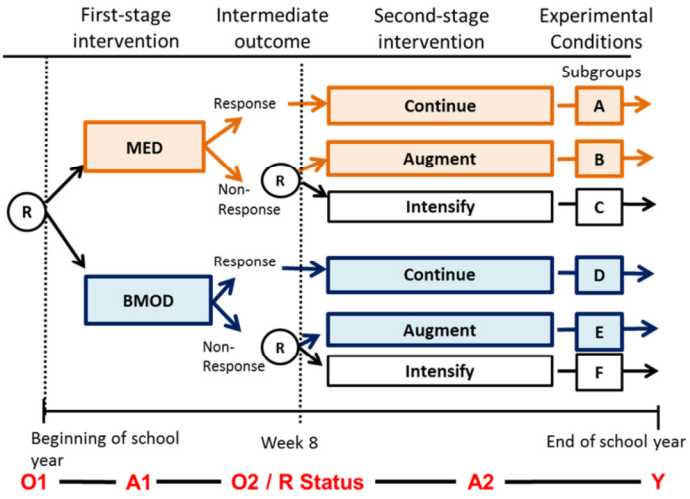
start with BMOD; if non-responder augment (BMOD+MED), else continue (BMOD),

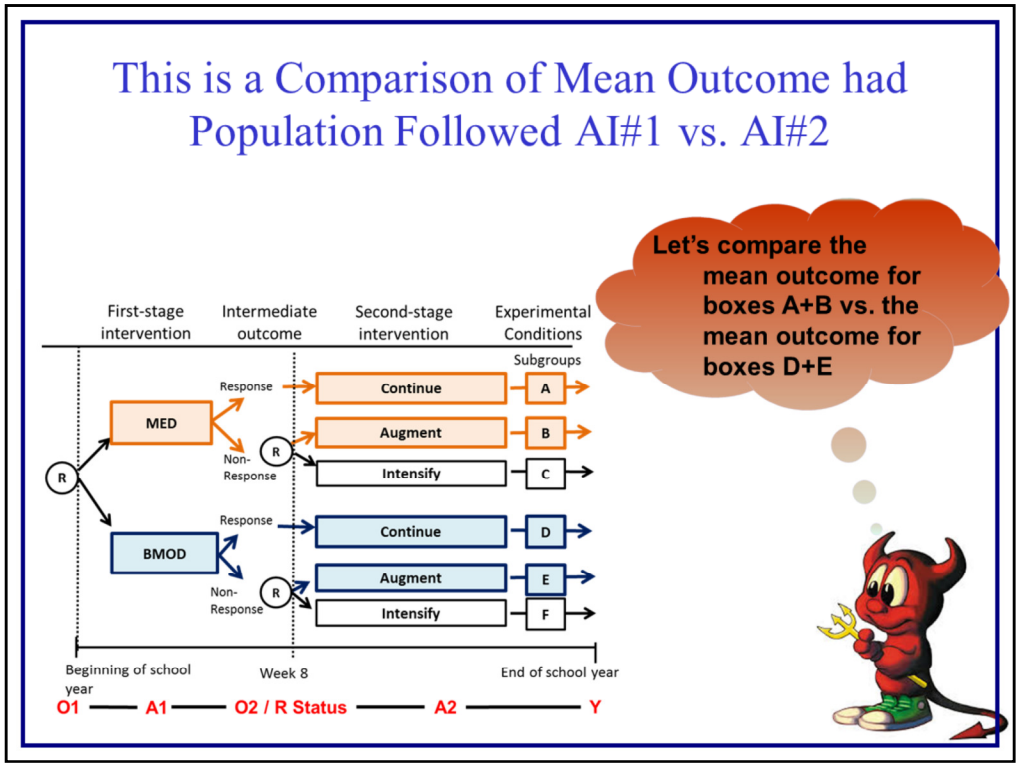
is better than an AI that recommends to

start with MED; if non-responder augment (BMOD+MED), else continue (MED),

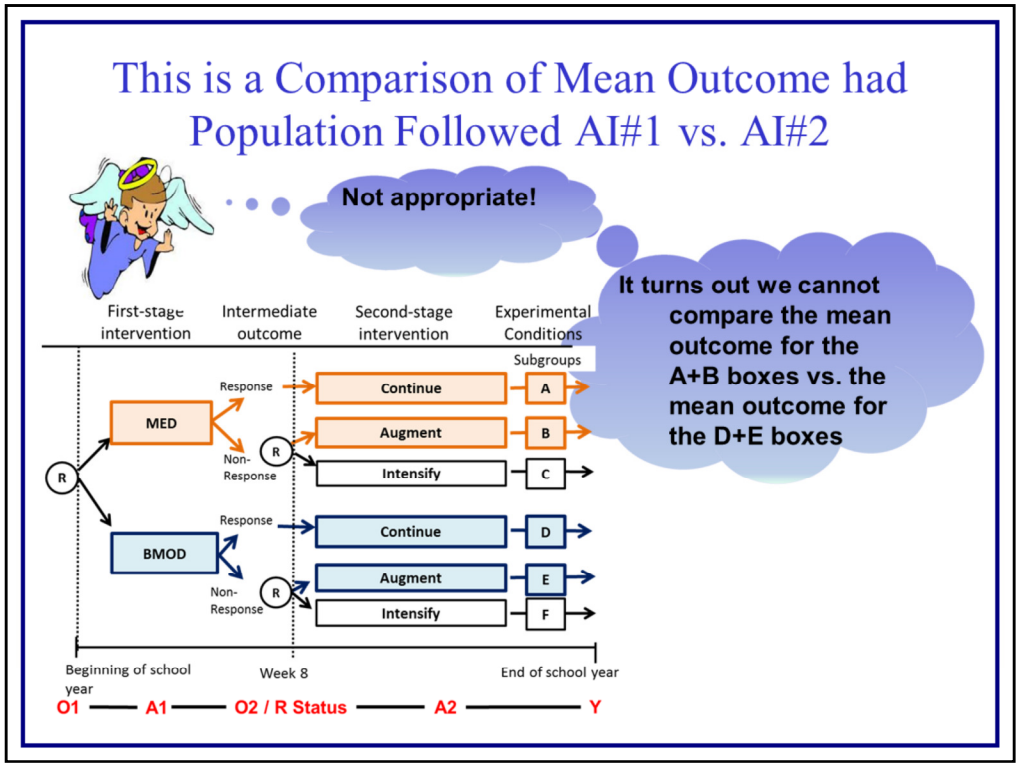
in terms of end of study school performance.

This is a Comparison of Mean Outcome had Population Followed AI#1 vs. AI#2

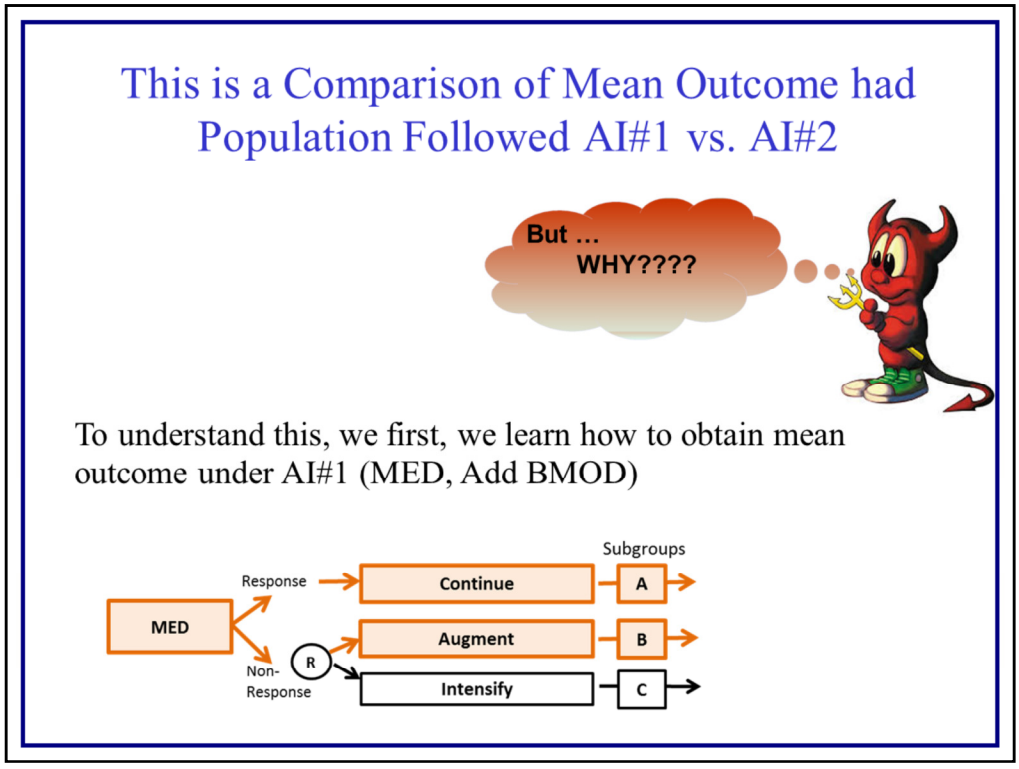




Your initial approach to this comparison might be to just take the mean across participants in boxes A+B and compare to the mean outcome of participants in boxes D+E.



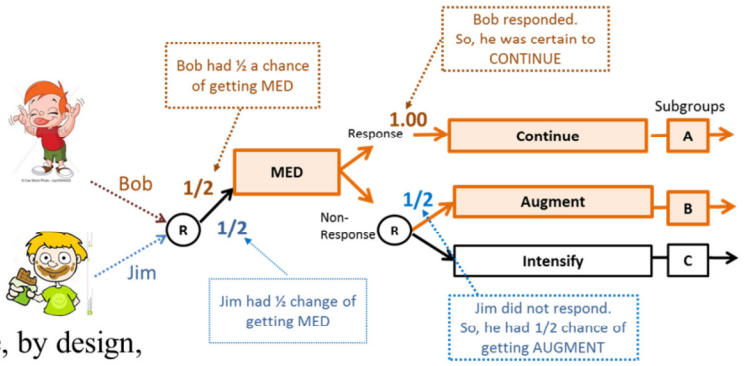
But this approach is not appropriate.



We cant do this because there is imbalance in the responders and non-responders who followed the (MED, Add BMOD) AI#1.

For example, let's first consider estimating the mean outcome had all participants followed AI#1. The issue is…[next slide]

There is Imbalance in the Non/Responding Participants Following this AI



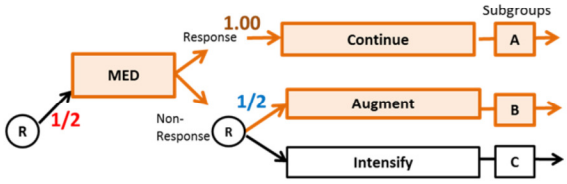
…because, by design,

- Responders had ½ chance of following AI #1, whereas
- Non-responders had a ½ x ½ = ¼ chance of following AI #1
- So, we want to estimate mean outcome had all participants followed AI#1
- But, responders are over-represented in this data, by design.
- We want all participants to be equally represented in this data

… that responders are over-represented in the data BY DESIGN.

This picture is just heuristic. Actually, responders have R*1/2 probability of being part of this AI and non-responders have (1-R)*1/4.
So if response rate is 0.5 responders have ¼ probability and non-responders have 1/8 probability of being part of this AI.

There is Imbalance in the Non/Responding Participants Following this AI



…by design,

- Responders had ½ = **0.5** chance of following this AI, whereas
- Non-responders had a 0.5 x 0.5 = ¼ = **0.25** chance of following this AI

What can we do? We can fix this imbalance by

- Assigning W = weight = 2 to responders to MED → **2* ½ = 1**
- Assign W = weight = 4 to non-responders to MED → **4* ¼ =1**
- 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 A+B boxes above.

The weights are different for participants in Box A vs participants in Box B.

In the next slides we show how to do something equivalent to this using a regression approach.
The weights are similar conceptually to the use of weights in survey methodology, where weights are being used to correct for certain subgroups being over or under represented in a survey because of survey non-response or self-selection. So, survey methodologists give higher weights to survey responders who are less represented in the survey.

These weights however differ from those used in propensity score, in terms of their purpose. The weights used in the context of propensity scores allows one to design and analyze an observational (nonrandomized) study so that it mimics some of the particular characteristics of a randomized controlled trial. In observational studies, treatment selection is often influenced by subject characteristics. As a result, baseline characteristics of treated subjects often differ systematically from those of untreated subjects. Therefore, the weights are used to account for systematic differences in baseline characteristics between treated and untreated subjects when estimating the effect of treatment on outcomes.

SAS Code to Estimate Mean Outcome had all Participants Followed AI#1 (MED, Add BMOD)

```
* First, create indicator and assign weights;

data dat5; set dat2;
  Z1=-1;
  if A1=-1 and R=1              then Z1=1;
  if A1=-1 and R=0 and A2=-1    then Z1=1;
  W=2*R + 4*(1-R);
run;
```

- The indicator Z1 differentiates between participants who followed AI#1 (Z1 = 1) and those who did not (Z1 = -1)
- W will equal 2 if R=1 (responder) and 4 if R=0 (non-responder)

This analysis is on Page 5 of your SAS code Word document.

SAS Code to Estimate Mean Outcome had all Participants Followed AI#1 (MED, Add BMOD)

```
* Second, run W-weighted regression
                                  Y = b0 + b1*z1 + e;
* b0 + b1 will represent the mean outcome under AI#1;
proc genmod data = dat5;
  class id;
  model y = z1;
  scwgt w;
  repeated subject = id / type = ind;
  estimate 'Mean Y under AI#1' intercept 1 z1  1;
run;
```

This is how we ask SAS to provide robust standard errors:
Why do we need that?
Weights depend on response status, which is unknown ahead of time.
Robust SE account for this uncertainty (i.e., for sampling error in the "estimation" of the weights).

Instead of a regression, you can also calculate the W-weighted mean outcomes for all participants following AI #1
Weighted Mean = $\Sigma w_i Y_i / \Sigma w_i$;

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 for Estimated Mean Outcome had All Participants Followed AI#1 (MED, Add BMOD)

Analysis Of GEE Parameter Estimates

Parameter	Estimate	SError	P-value
Intercept	2.9153	0.1084	<.0001
Z1	-0.0504	0.1084	0.6417

Contrast Estimate Results

	Estimate	95% Conf Limits Lower	Upper	SError
Mean Y under AI#1 (MED, Add BMOD)	**2.8649**	2.5305	3.1992	**0.1706**

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.