

Secondary Aims Using Data Arising from a SMART

Using moderators to build a more deeply-tailored AI







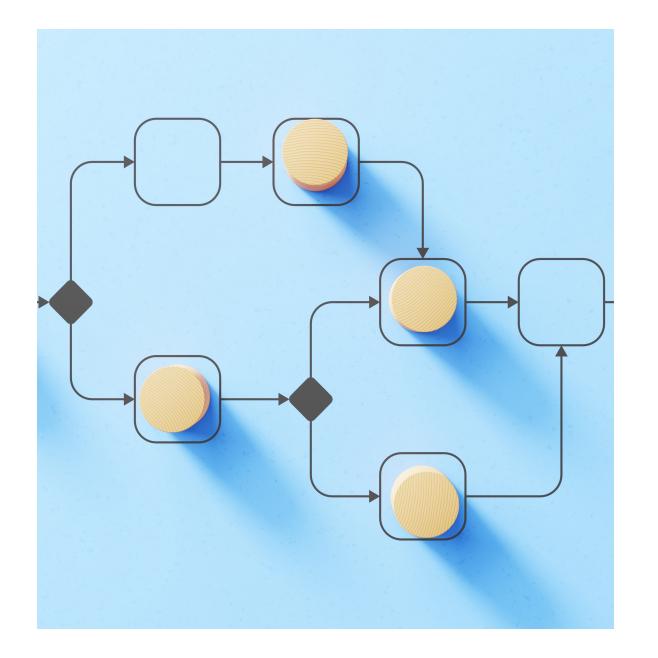
Learning Goals

Learn about one of the most innovative Secondary Aims of a SMART

 Constructing a proposal for a "more deeply-tailored" adaptive intervention.

Learn a new analysis method to use data from a SMART to propose a more deeply-tailored AI

The method is familiar and easy-to-use



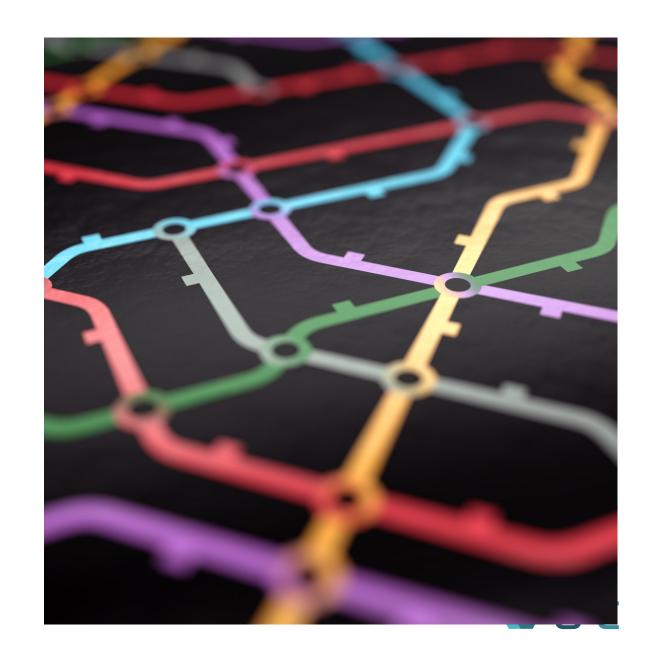
Outline

What is a more deeply-tailored adaptive intervention?

SMART secondary aims about more deeplytailored adaptive interventions

How can moderators analyses help construct a more deeply-tailored adaptive intervention?

Q-Learning: An extension of moderators analysis for data from a SMART



Outline

What is a more deeply-tailored adaptive intervention?

SMART secondary aims about more deeplytailored adaptive interventions

How can moderators analyses help construct a more deeply-tailored adaptive intervention?

Q-Learning: An extension of moderators analysis for data from a SMART



What is a more deeply tailored AI?

- A more deeply tailored AI is an adaptive intervention that
 - includes additional tailoring variables or decision rules
 - and leads to better outcomes
- For example, an AI that tailors second-stage treatment based on response status and other known variables



Recall the 4 embedded Als in the ADHD SMART Using if-then statements

AI #1:

Start with MED; if non-responder AUGMENT, else CONTINUE

AI #3:

Start with MED; if non-responder INTENSIFY, else CONTINUE

AI #2:

Start with BMOD; if non-responder AUGMENT, else CONTINUE

AI #4:

Start with BMOD; if non-responder INTENSIFY, else CONTINUE



Let's focus on embedded AI#2 Using if-then statements

AI #1:

Start with MED; if non-responder AUGMENT, else CONTINUE

AI #3:

Start with MED; if non-responder INTENSIFY, else CONTINUE

AI #2:

Start with BMOD; if non-responder AUGMENT, else CONTINUE

AI #4:

Start with BMOD; if non-responder INTENSIFY, else CONTINUE



Let's focus on AI#2
Using a schematic

Responders

Continue

Non-Responders

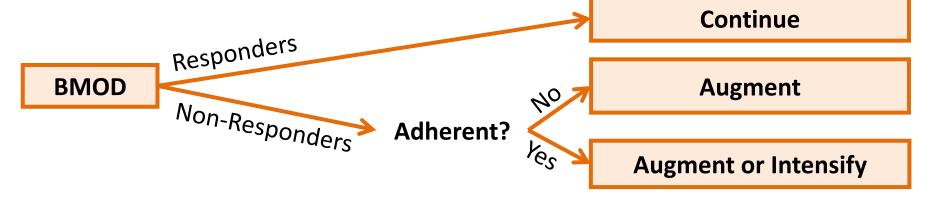
Augment

Let's focus on AI#2
Using a schematic

Responders Continue

Non-Responders Augment

An Al that is **more deeply tailored** than Al#2



Let's focus on AI#2 Using a schematic Responders **Continue BMOD** Non-Responders **Augment** An Al that is even more deeply tailored than AI#2 **Continue** Responders **BMOD Augment** 40 Non-Responders 40 Adherent? res **Augment or Intensify MED** in prior year? **Continue MED Augment** 40 Adherent? Yes **Augment or Intensify** 9

An embedded Al Using if-then statements

At the beginning of the school year:

stage
$$1 = \{BMOD\}$$

Then, every month, beginning at week 8:

IF response status to stage 1 = {NR}

THEN stage 2 = {AUGMENT}

ELSE continue with **stage 1**.



A more deeply-tailored Al

Using if-then statements

```
At the beginning of the school year:
      IF medication in the prior year = {YES}
            THEN stage 1 = \{MED\}.
      ELSE IF medication in the prior year = {NO}
             THEN stage 1 = \{BMOD\}.
Then, every month, beginning at week 8:
      IF response status to stage 1 = {NR},
            THEN IF adherence to stage 1 = {NO},
                  THEN stage 2 = {AUGMENT}
            ELSE stage 2 = {AUGMENT} or {INTENSIFY}.
      ELSE continue with stage 1.
```



Outline

What is a more deeply-tailored adaptive intervention?

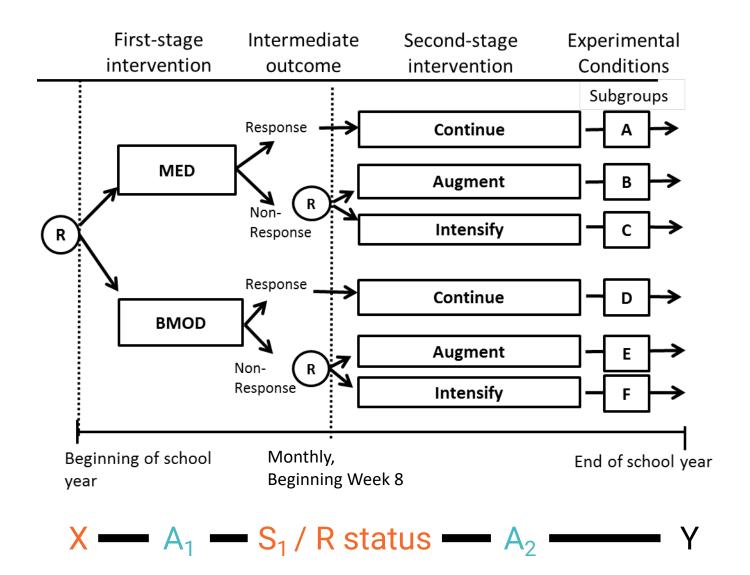
SMART secondary aims about more deeplytailored adaptive interventions

How can moderators analyses help construct a more deeply-tailored adaptive intervention?

Q-Learning: An extension of moderators analysis for data from a SMART

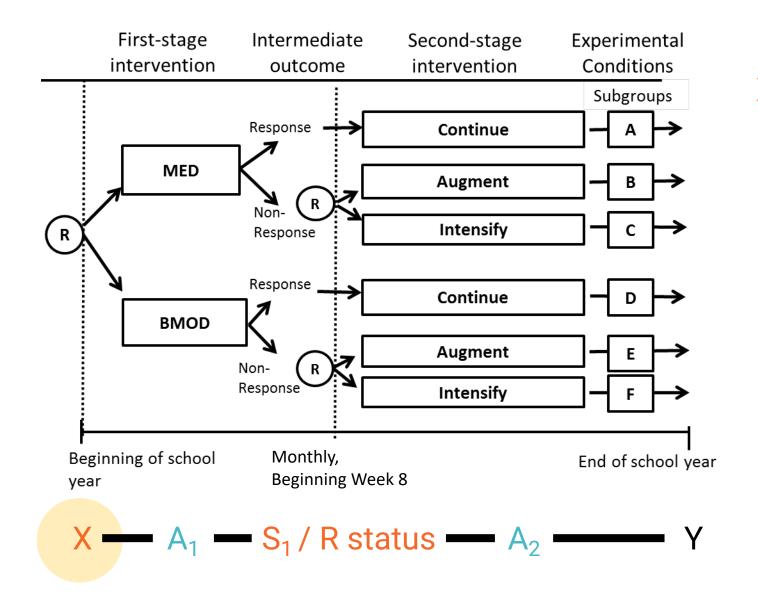


Measures collected in a SMART





Measures collected in a SMART

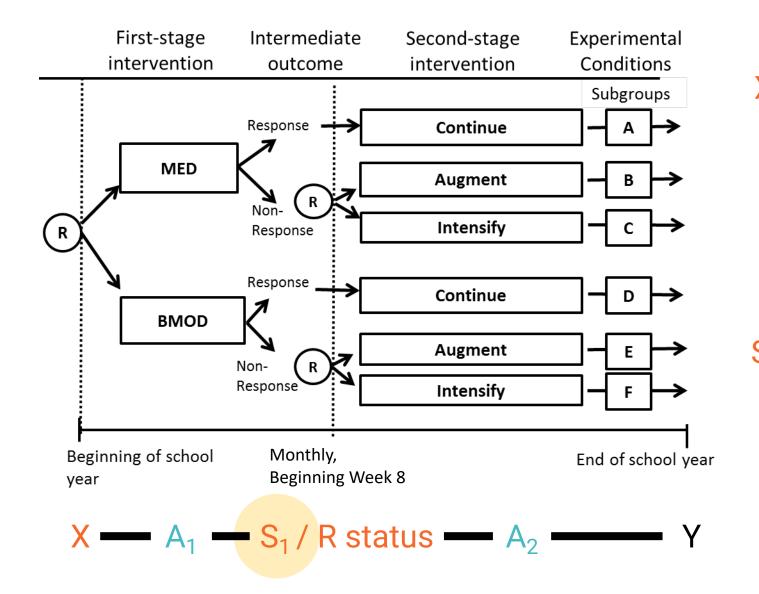


Baseline covariates

- Demographics
- MED before stage 1
- Baseline ADHD scores
- Baseline school performance
- ODD
- etc.



Measures collected in a SMART



Baseline covariates

- Demographics
- MED before stage 1
- Baseline ADHD scores
- Baseline school performance
- ODD
- etc.

Time-varying covariates

- Month of non-response
- Adherence to stage 1 treatment
- Parent function during stage 1

Example SMART secondary aims related to constructing a more deeply-tailored Al

For example, in the ADHD SMART, an investigator might be interested in

1) Whether **stage 1** of the intervention should be tailored according to whether the child has received **prior medication?**



Example SMART secondary aims related to constructing a more deeply-tailored Al

For example, in the ADHD SMART, an investigator might be interested in

- 1) Whether **stage 1** of the intervention should be tailored according to whether the child has received **prior medication?**
- 2) Whether, among non-responders, **stage 2** of the intervention should be tailored according to the child's level of **adherence?** to stage 1 treatment?



Example SMART secondary aims related to constructing a more deeply-tailored Al

For example, in the ADHD SMART, an investigator might be interested in

- 3) Whether **stage 2** intervention should be tailored according to change in ADHD symptoms from baseline to end of stage 1?
- 4) And, if so, what symptom cut-off do we use to make the **stage**2 intervention decision?

Outline

What is a more deeply-tailored adaptive intervention?

SMART secondary aims about more deeplytailored adaptive interventions

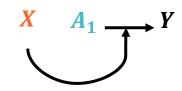
How can moderators analyses help construct a more deeply-tailored adaptive intervention?

Q-Learning: An extension of moderators analysis for data from a SMART



Review: What is a moderator variable?

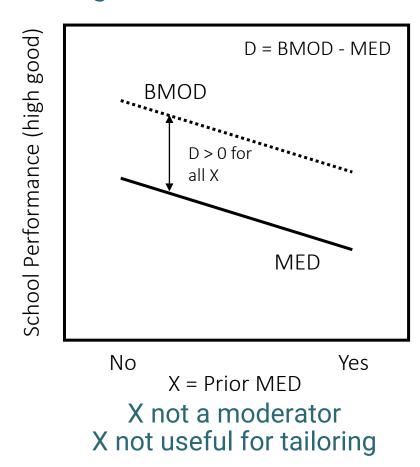
- A moderator is a variable that influences the individual causal effects of an intervention on an outcome.



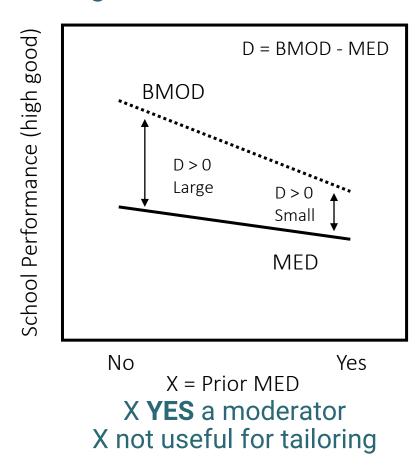
- Moderators can be useful for informing how to tailor an intervention
- Moderators are easy to examine using standard regression analyses



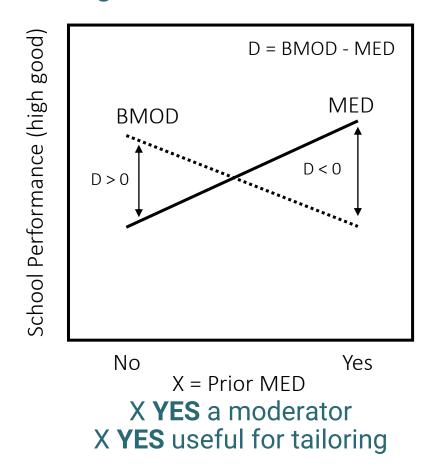
Hypothetical results examining whether X moderates the effect of A on Y



Hypothetical results examining whether X moderates the effect of A on Y



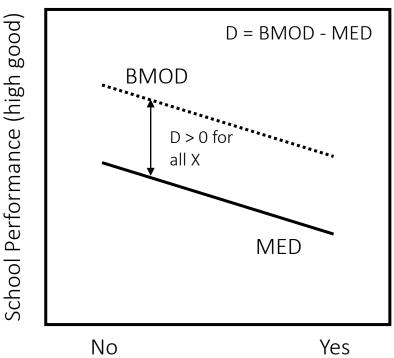
Hypothetical results examining whether X moderates the effect of A on Y



(high good)

School Performance

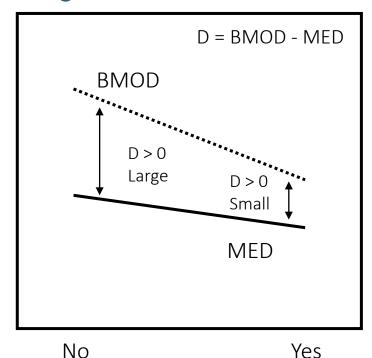
Hypothetical results examining whether X moderates the effect of A on Y



X = Prior MED

X not a moderator

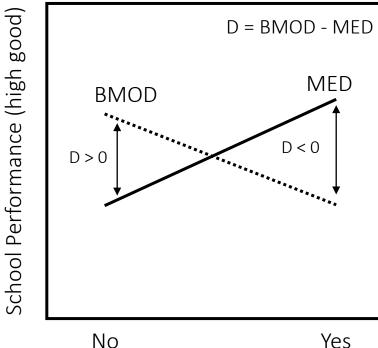
X not useful for tailoring



X = Prior MED

X YES a moderator

X not useful for tailoring



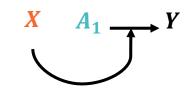
X = Prior MED

X YES a moderator

X YES useful for tailoring

Moderator analyses in a SMART requires extension because there are two stages

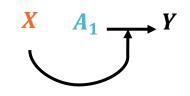
At the beginning of Stage 1: Does medication use in the prior year moderate the effect of starting with MED vs BMOD?



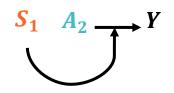


Moderator analyses in a SMART requires extension because there are two stages

At the beginning of Stage 1: Does medication use in the prior year moderate the effect of starting with MED vs BMOD?



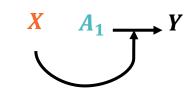
At the beginning of Stage 2: Does level of adherence to first-stage intervention moderate the effect of AUGMENT vs INTENSIFY among non-responders?



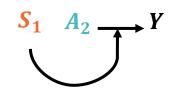


Moderator analyses in a SMART requires extension because there are two stages

At the beginning of Stage 1: Does medication use in the prior year moderate the effect of starting with MED vs BMOD?



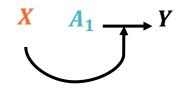
At the beginning of Stage 2: Does level of adherence to first-stage intervention moderate the effect of AUGMENT vs INTENSIFY among non-responders?



Results from such moderator analyses in a SMART can be used to suggest a more deeply-tailored AI.

Examining baseline **moderators** of stage 1 intervention in a SMART

 For the first question, we could use the following regression:



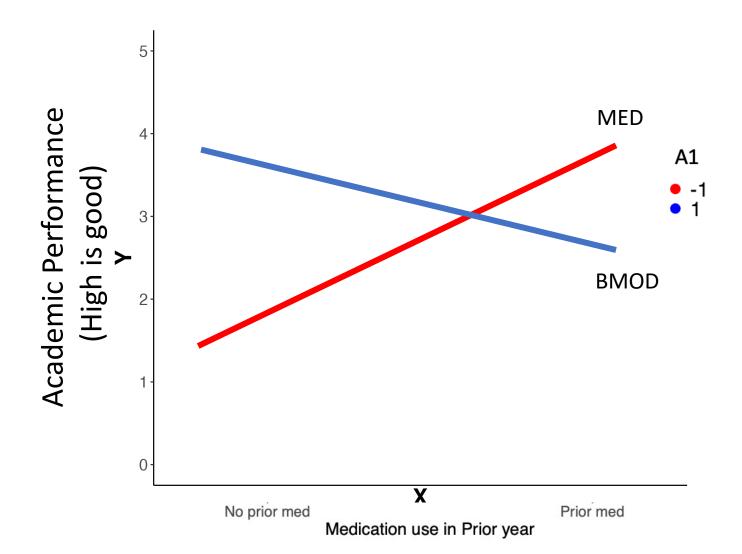
$$E[Y | X, A_1] = \beta_0 + \beta_1 X + \beta_2 A_1 + \beta_3 X A_1$$

Covariate-by-treatment Interaction term

 This regression examines whether X is a moderator of the effect of first-stage treatment



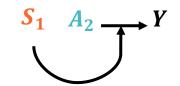
Examining baseline **moderators** of stage 1 intervention in a SMART

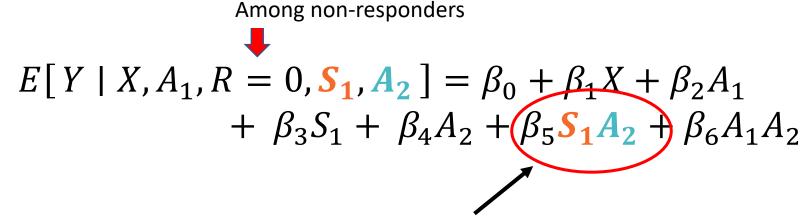




Examine baseline & time-varying moderators of stage 2 intervention in a SMART

 For the second question, we could use the following regression:



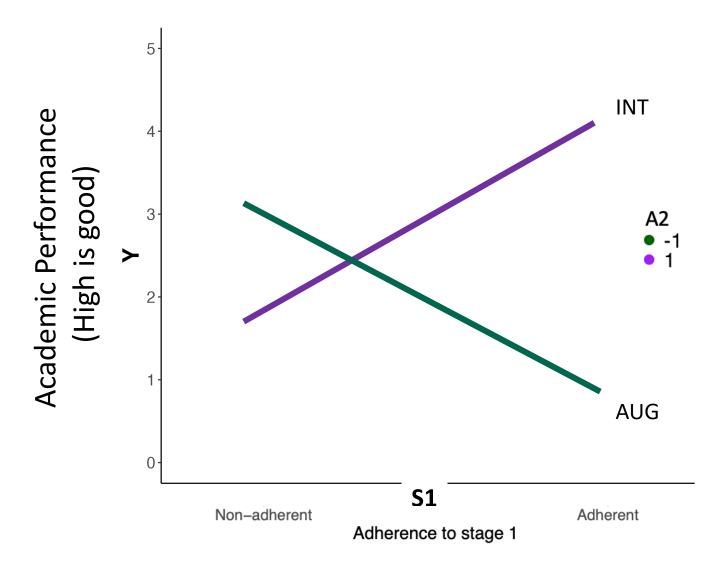


Covariate-by-treatment Interaction term

 This regression examines whether S₁ is a moderator of the effect of second-stage treatment, among non-responders



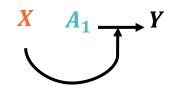
Examine baseline & time-varying moderators of stage 2 intervention in a SMART



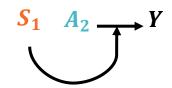


Now, you may be wondering...

$$E[Y | X, A_1] = \beta_0 + \beta_1 X + \beta_2 A_1 + \beta_3 X A_1$$



$$E[Y \mid X, A_1, R = 0, S_1, A_2] = \beta_0 + \beta_1 X + \beta_2 A_1 + \beta_3 S_1 + \beta_4 A_2 + \beta_5 S_1 A_2 + \beta_6 A_1 A_2$$



...what if, instead of these two separate regressions, we used a single regression model to answer both questions simultaneously?



What if we did a single regression?

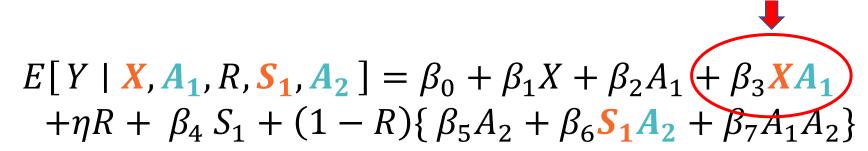
For example:

$$E[Y \mid X, A_1, R, S_1, A_2] = \beta_0 + \beta_1 X + \beta_2 A_1 + \beta_3 X A_1 + \eta R + \beta_4 S_1 + (1 - R) \{ \beta_5 A_2 + \beta_6 S_1 A_2 + \beta_7 A_1 A_2 \}$$



What if we did a single regression?

• For example:



- But there are two causal problems with this approach!
 - Both result from the possibility that S_1 can be impacted by A_1
 - Both cause bias in β_2 and β_3



Causes bias

What if we did a single regression?

For example:

$$E[Y \mid X, A_1, R, S_1, A_2] = \beta_0 + \beta_1 X + \beta_2 A_1 + \beta_3 X A_1 + \eta R + \beta_4 S_1 + (1 - R) \{ \beta_5 A_2 + \beta_6 S_1 A_2 + \beta_7 A_1 A_2 \}$$

- Problem 1: Wrong moderating effect of A1 on Y
 - We may have unintentionally cut off the effect of A_1 on Y via S_1 .



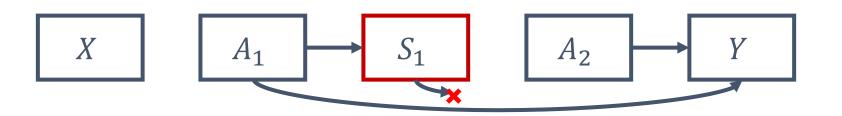


What if we did a single regression?

For example:

$$E[Y \mid X, A_1, R, S_1, A_2] = \beta_0 + \beta_1 X + \beta_2 A_1 + \beta_3 X A_1 + \eta R + \beta_4 S_1 + (1 - R) \{ \beta_5 A_2 + \beta_6 S_1 A_2 + \beta_7 A_1 A_2 \}$$

- Problem 1: Wrong causal effect of A1 on Y
 - We unintentionally cut off any effect of A_1 on Y that is mediated by S_1 .
 - · This is not the moderator effect we want.



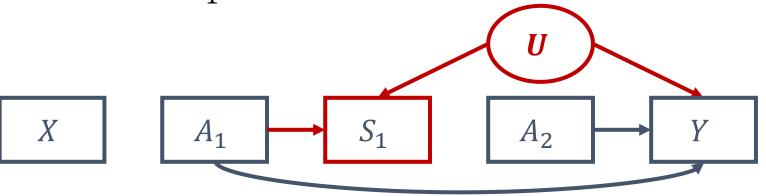


What if we did a single regression?

For example:

$$E[Y \mid X, A_1, R, S_1, A_2] = \beta_0 + \beta_1 X + \beta_2 A_1 + \beta_3 X A_1 + \eta R + \beta_4 S_1 + (1 - R) \{ \beta_5 A_2 + \beta_6 S_1 A_2 + \beta_7 A_1 A_2 \}$$

- Problem 2: Collider Bias (a.k.a Causal Bias)
 - We may get unintended spurious effects due to known or unknown common causes of S_1 and Y





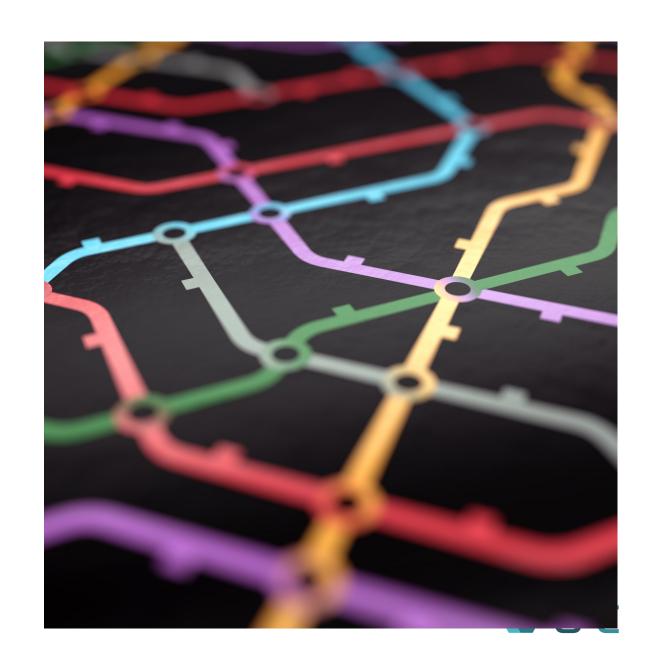
Outline

What is a more deeply-tailored adaptive intervention?

SMART secondary aims about more deeplytailored adaptive interventions

How can moderators analyses help construct a more deeply-tailored adaptive intervention?

Q-Learning: An extension of moderators analysis for data from a SMART



What is Q-Learning? Why is it useful?

Q = quality of the adaptive intervention

- Uses the two regression models presented earlier which are easy to use and interpret
- Bypasses the two problems associated with the single regression approach
- Leads to a better proposal for a more deeply-tailored Al
 - Appropriately accounts for the optimal second-stage intervention when determining the optimal first-stage intervention

Q-learning: Three simple steps

Step 1

Stage 2 regression → Obtain optimal stage 2 decision

 $E[Y_i \mid X, \mathbf{A_1}, \mathbf{S_1}, \mathbf{A_2}, R = 0] = \beta_0 + \beta_1 X + \beta_2 A_1 + \beta_3 S_1 + \beta_4 \mathbf{A_2} + \beta_5 \mathbf{S_1} \mathbf{A_2} + \beta_6 \mathbf{A_1} \mathbf{A_2}$

Step 2

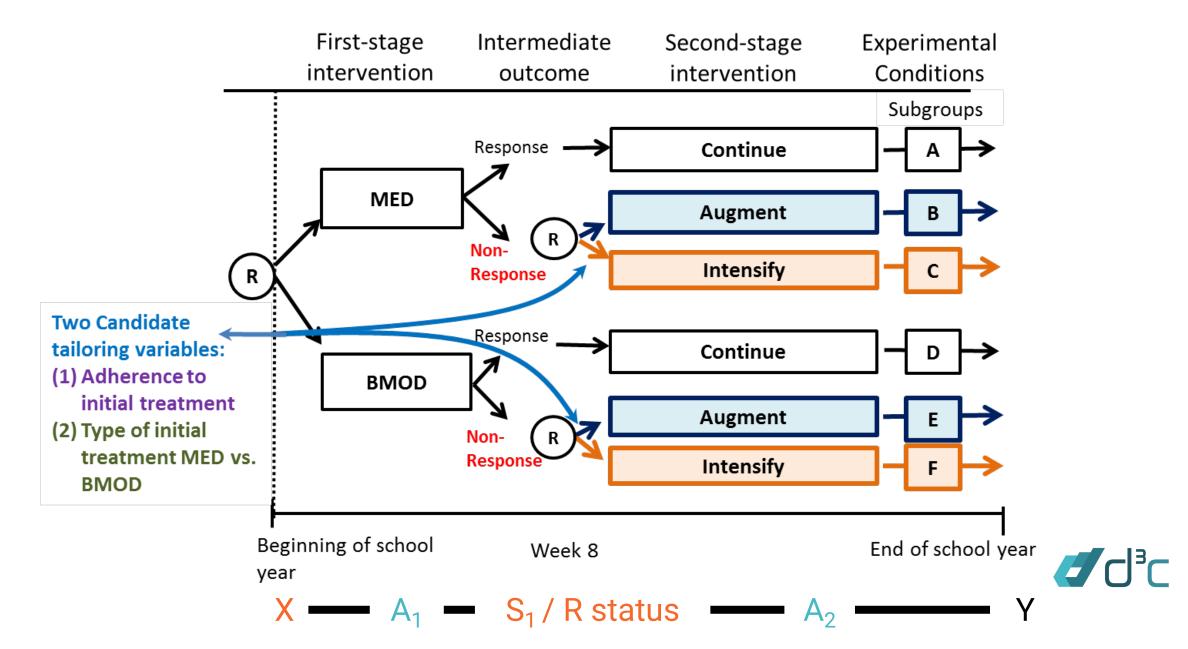
- Calculate Y_i^{opt}
 - For non-responders: $\widehat{Y_i^{opt}}$ is the estimated predicted outcome (based on step 1) had non-responder i been offered the best stage 2 intervention given X, A₁, and S₁
 - For responders: we use $\widehat{Y_i^{opt}} = Y_i$

Step 3

Stage 1 regression → Obtain optimal stage 1 decision

$$E\left[\widehat{Y_i^{opt}} \mid X, A_1\right] = \beta_0 + \beta_1 X + \beta_2 A_1 + \beta_3 X A_1$$

Recall the ADHD SMART



Let's write this moderators analysis in terms of the ADHD SMART:

$$E[Y \mid X, \mathbf{A_1}, \mathbf{S_1}, \mathbf{A_2}, R = 0] = \beta_0 + \dots + \beta_2 A_1 + \beta_3 \mathbf{adherence} + \beta_4 \mathbf{A_2} + \beta_5 (\mathbf{A_2} \times \mathbf{A_1}) + \beta_6 (\mathbf{A_2} \times \mathbf{adherence})$$

 A_1 = Stage 1 options: -1=MED; 1=BMOD

 S_{11} = Adherence to stage 1: 1=yes; 0=no

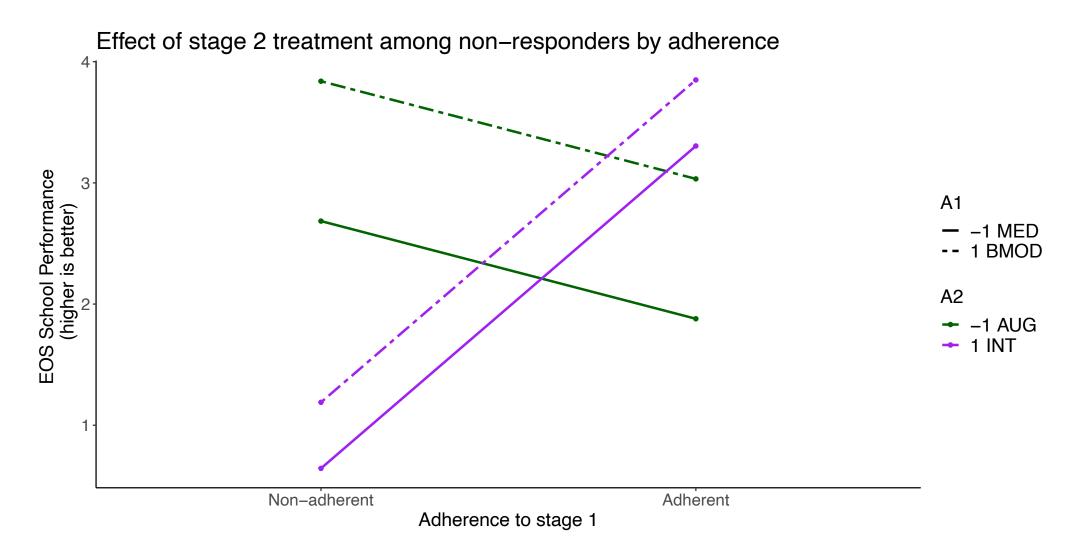
A₂ = Stage 2 options: -1=AUGMENT; 1=INTENSIFY

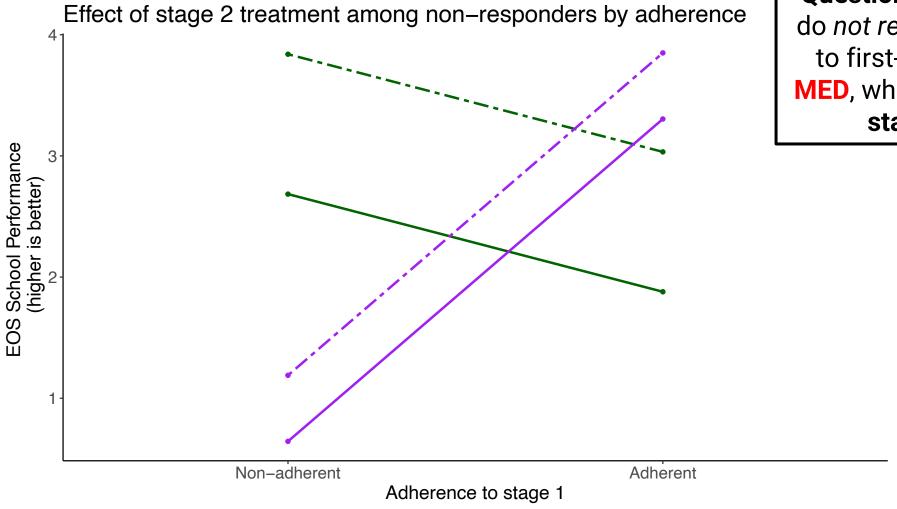
Y = End of year school performance

This model will help us to...

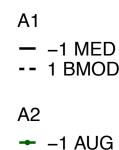
- a) Determine if the best second-stage tactic depends on adherence; and
- b) Identify the best second-stage tactic for each level of adherence and first-stage treatment

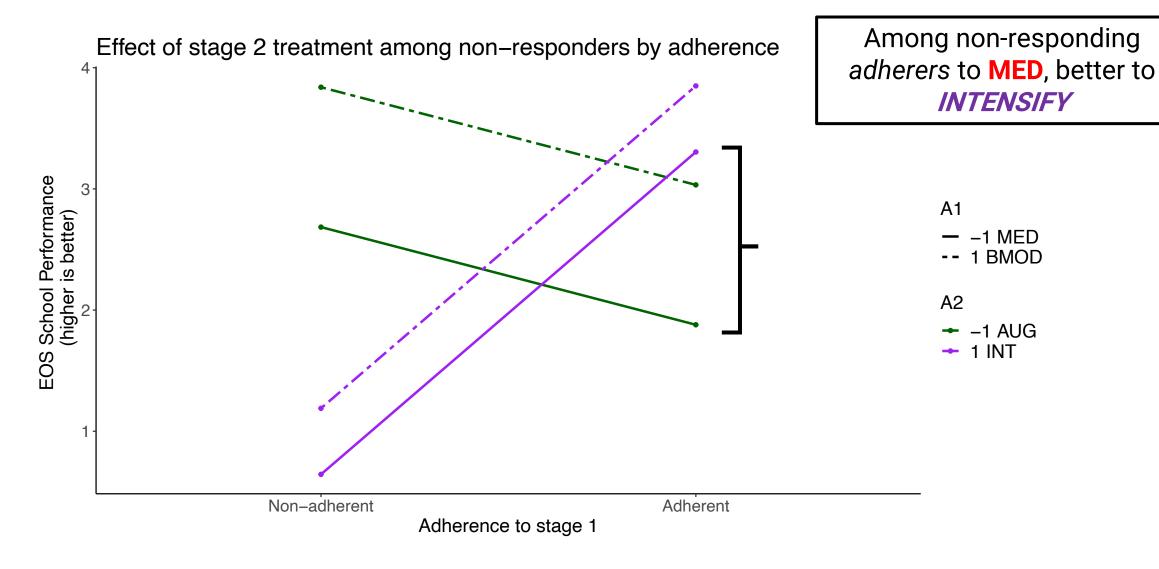


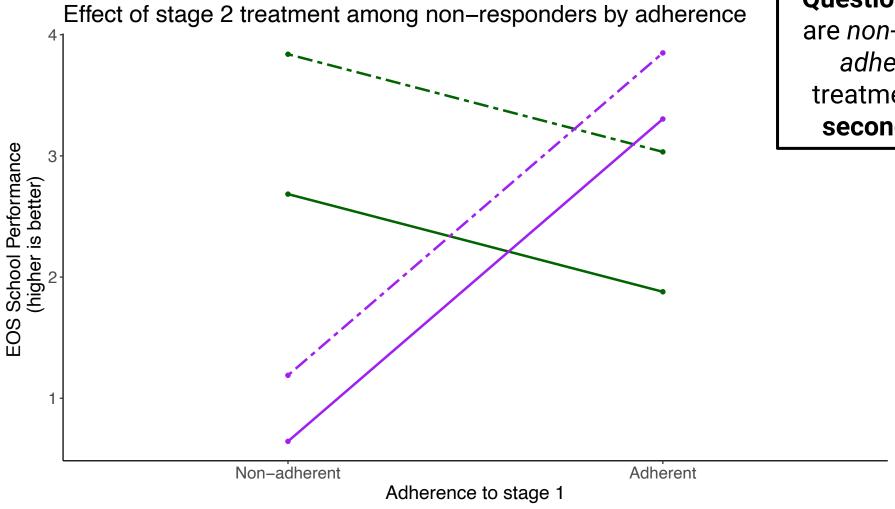




Question: Among those who do not respond but do adhere to first-stage treatment of MED, what is the best second stage treatment?

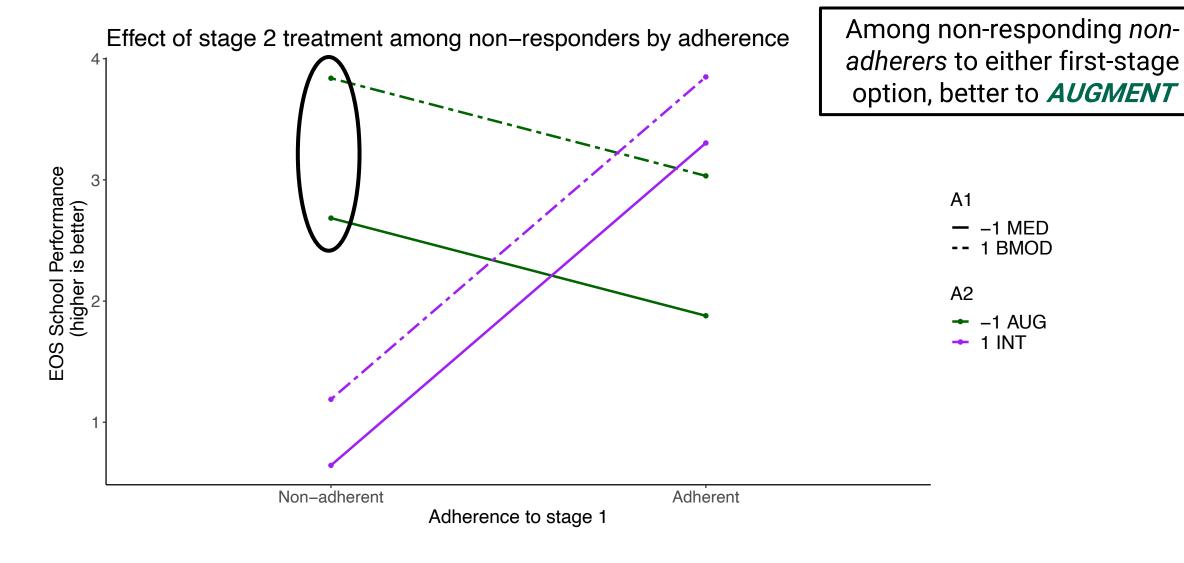






Question: Among those who are non-responders and non-adherers to first-stage treatment, what is the **best** second stage treatment?

A1 — −1 MED -- 1 BMOD A2 — −1 AUG



Q-learning: Step 2

Step 1

Stage 2 regression → Obtain optimal stage 2 decision

$$E[Y_i \mid X, \mathbf{A_1}, \mathbf{S_1}, \mathbf{A_2}, R = 0] = \beta_0 + \beta_1 X + \beta_2 A_1 + \beta_3 S_1 + \beta_4 \mathbf{A_2} + \beta_5 \mathbf{S_1} \mathbf{A_2} + \beta_6 \mathbf{A_1} \mathbf{A_2}$$

Step 2

- Calculate $\widehat{Y_i^{opt}}$
 - For non-responders: $\widehat{Y_i^{opt}}$ is the estimated predicted outcome (based on step 1) had non-responder i been offered the best stage 2 intervention given X, A₁, and S₁
 - For responders: we use $\widehat{Y_i^{opt}} = Y_i$

Step 3

Stage 1 regression → Obtain optimal stage 1 decision

$$E\left[\widehat{Y_i^{opt}} \mid X, A_1\right] = \beta_0 + \beta_1 X + \beta_2 A_1 + \beta_3 X A_1$$

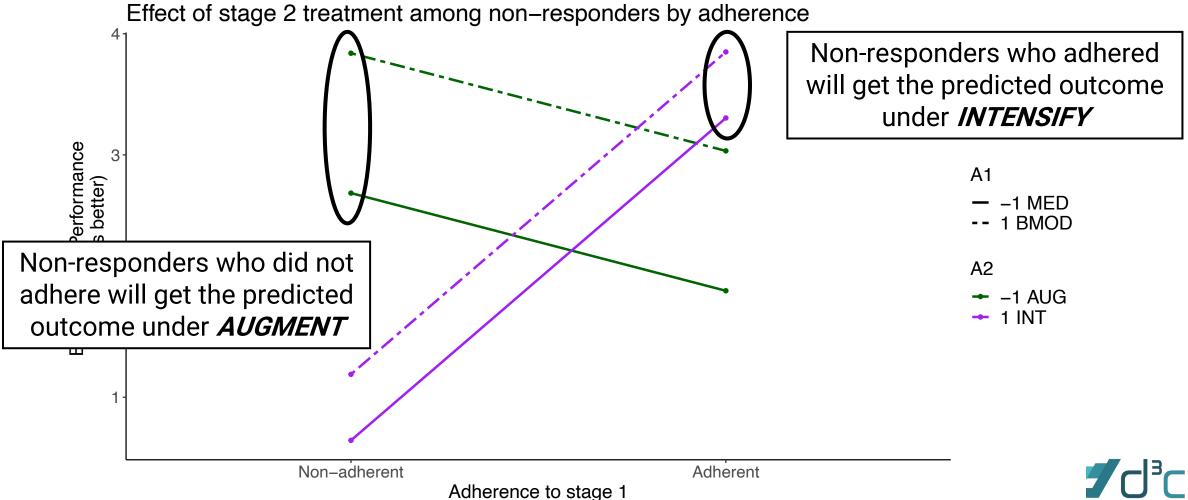
Step 2: predict the outcome under the best second-stage option

 Next, we use the regression from step 1 to estimate the outcome for each non-responder if they received the best second-stage tactic, given their observed values on the tailoring variables

- We assign each **non-responder** the value Y_i^{opt}
 - $\widehat{Y_i^{opt}}$ = the expected outcome if each non-responder received the **best** second-stage tactic given their <u>initial treatment</u> and <u>adherence</u>
- We assign each **responder** the value $\widehat{Y_i^{opt}} = Y_i$



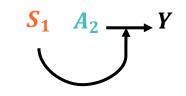
Step 2: predicted outcome under the best second-stage option



Step 2: predicted outcome under the best second-stage option

Recall the model for our stage 2 moderators analysis:

$$E[Y \mid X, \mathbf{A_1}, \mathbf{S_1}, \mathbf{A_2}, R = 0] = \beta_0 + \dots + \beta_2 A_1 + \beta_3 adherence + \beta_4 \mathbf{A_2} + \beta_5 (\mathbf{A_2} \times \mathbf{A_1}) + \beta_6 (\mathbf{A_2} \times adherence)$$



- Suppose that: $\beta_0 = 2.2$, $\beta_2 = 0.35$, $\beta_3 = 0.47$, $\beta_4 = -1.0$, $\beta_5 = -0.13$, $\beta_6 = 1.7$
- Suppose John was a non-responding, non-adhering (adherence = 0) participant who had mean values for all baseline variables and received MED (A₁ = 1) at stage 1 and INT (A₂ = -1) at stage 2.
- Predict John's scores:

$$\hat{Y} = 2.2 + 0.35(A_1) + 0.47(adherence) - 1.0(A_2) - 0.13(A_2 \times A_1) + 1.7(A_2 \times adherence)$$

$$\hat{Y}_{A2=INT} = 2.2 + 0.35(1) + 0.47(0) - 1.0(1) - 0.13(1 \times 1) + 1.7(1 \times 0) = 3.1$$
(which he received)
$$\hat{Y}_{A2=AUG} = 2.2 + 0.35(1) + 0.47(0) - 1.0(-1) - 0.13(-1 \times 1) + 1.7(-1 \times 0) = 2.5$$
John's score under AUG (which he did *not* receive)

Q-learning: Step 3

Step 1

Stage 2 regression → Obtain optimal stage 2 decision

 $E[Y_i \mid X, \mathbf{A_1}, \mathbf{S_1}, \mathbf{A_2}, R = 0] = \beta_0 + \beta_1 X + \beta_2 A_1 + \beta_3 S_1 + \beta_4 \mathbf{A_2} + \beta_5 \mathbf{S_1} \mathbf{A_2} + \beta_6 \mathbf{A_1} \mathbf{A_2}$

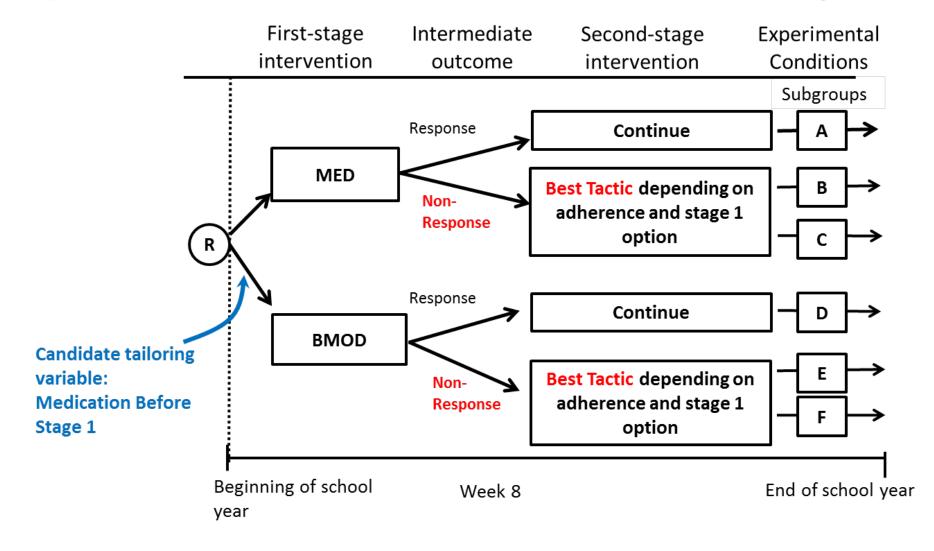
Step 2

- Calculate $\widehat{Y_i^{opt}}$
 - For non-responders: $\widehat{Y_i^{opt}}$ is the estimated predicted outcome (based on step 1) had non-responder i been offered the best stage 2 intervention given X, A₁, and S₁
 - For responders: we use $\widehat{Y_i^{opt}} = Y_i$

Step 3

Stage 1 regression → Obtain optimal stage 1 decision

$$E\left[\widehat{Y_i^{opt}} \mid X, A_1\right] = \beta_0 + \beta_1 X + \beta_2 A_1 + \beta_3 X A_1$$





• Fit the following regression model:

$$E\left[\widehat{Y_i^{opt}} \mid X, A_1\right] = \beta_0 + \beta_1 \text{priorMed} + \beta_2 A_1 + \beta_3 (\text{priorMed} \times A_1)$$

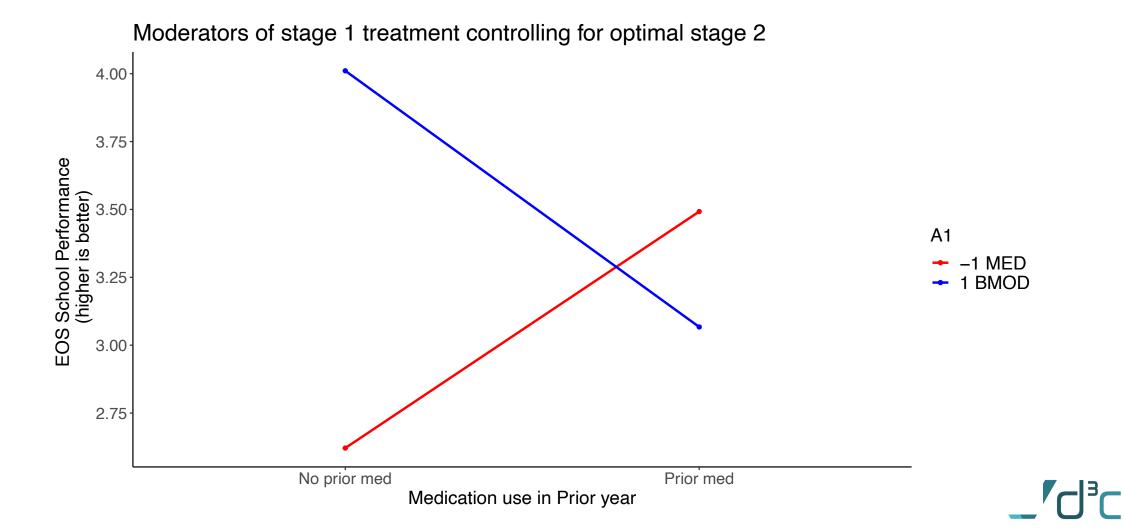
$$\uparrow \qquad \qquad \text{priorMed: } 1 = \text{yes; } 0 = \text{no}$$

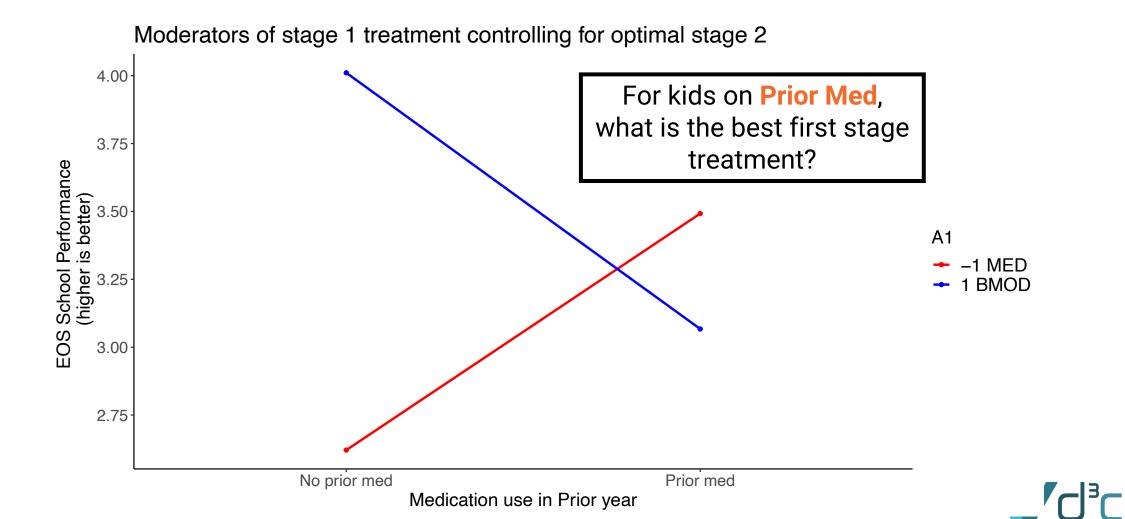
Estimated outcome, which controls for optimal stage 2 intervention, was calculated for all individuals in step 2

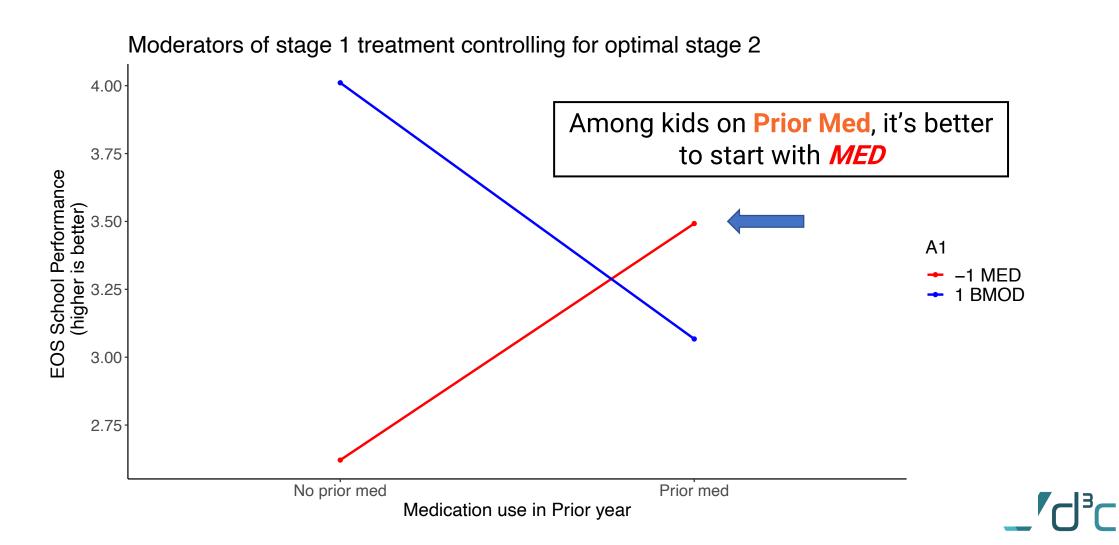
This model will help us to...

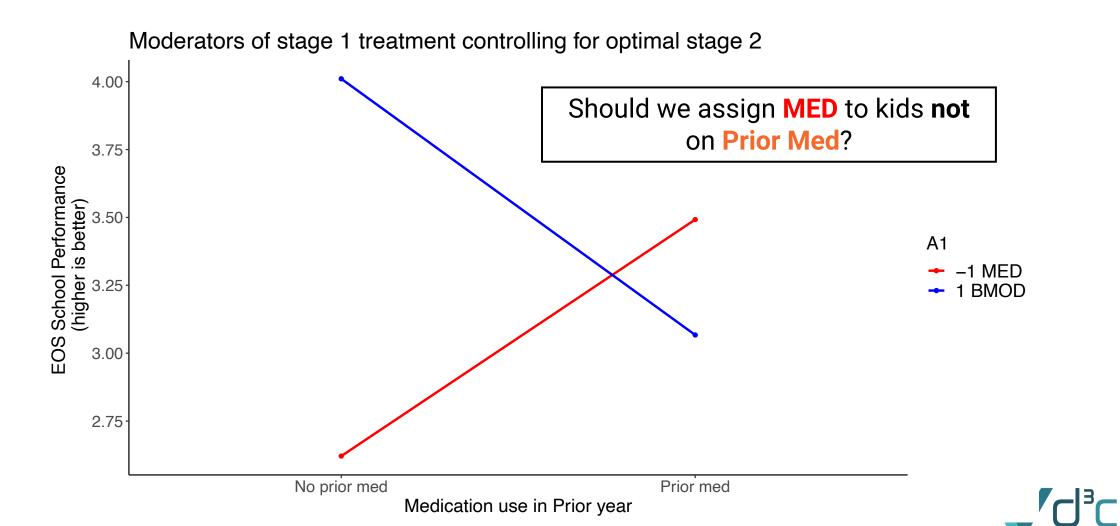
- a) Determine whether the best first-stage option depends on medication in prior year; and
- b) Identify the best first-stage option for children who received med in prior year vs. those that did not.

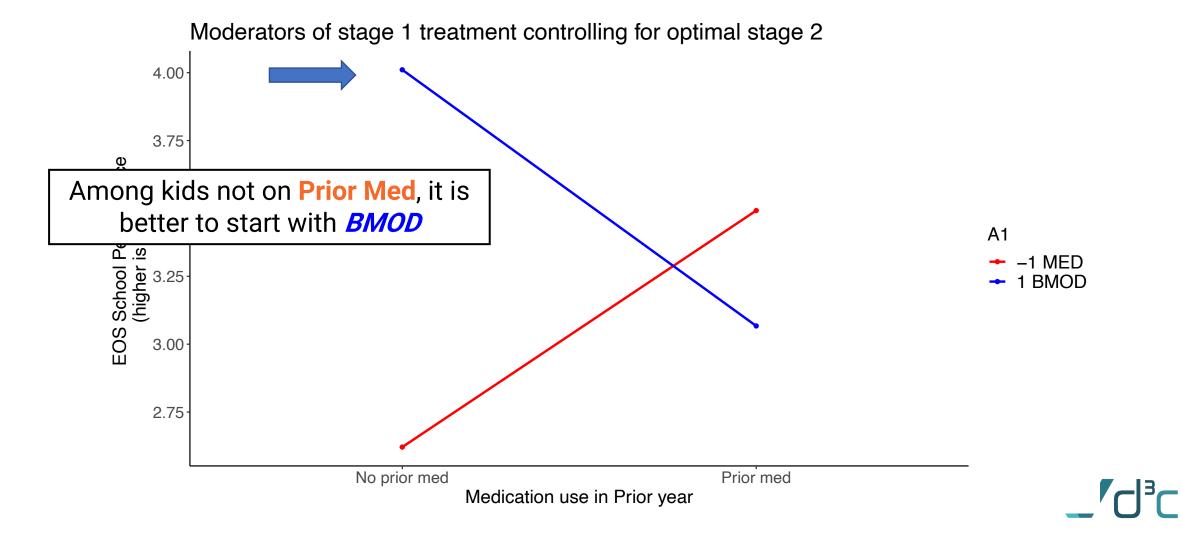




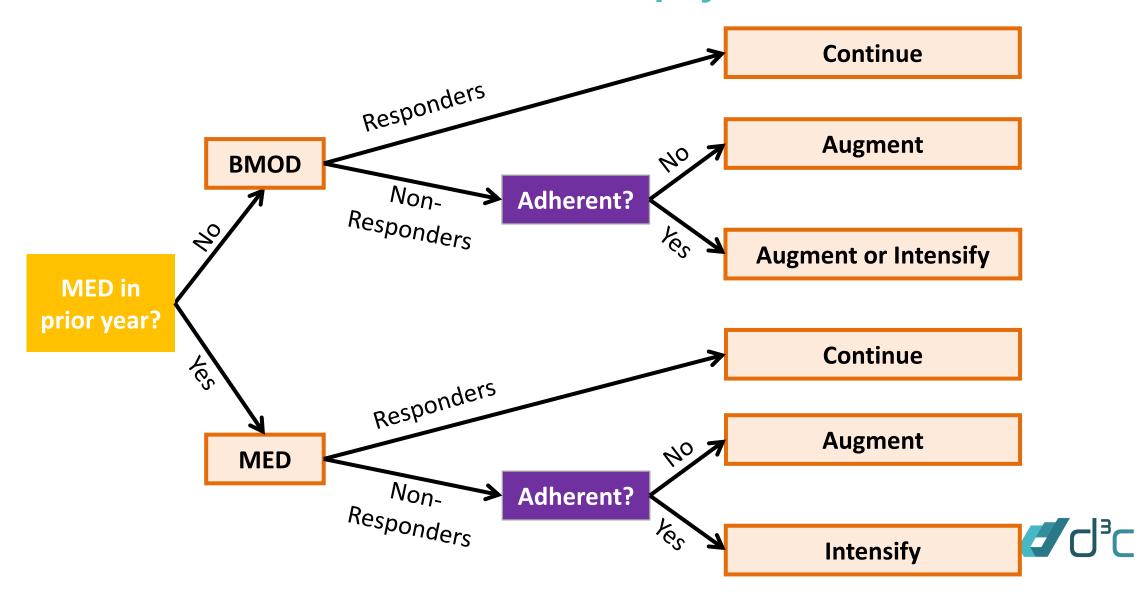








The estimated more-deeply tailored AI is



The estimated more-deeply tailored AI is

```
At the beginning of the school year:

IF medication in the prior year = {YES}

THEN stage 1 = {MED}.

ELSE IF medication in the prior year = {NO}

THEN stage 1 = {BMOD}.

Then, every month, beginning at week 8...
```



The estimated more-deeply tailored AI is

Then, every month, beginning at week 8...

```
IF response status to stage 1 = {NR}
      THEN
            IF adherence to MED or BMOD = {NO},
                   THEN stage 2 = {AUGMENT}.
            ELSE IF adherence to MED = {YES},
                   THEN stage 2 = {INTENSIFY}.
            ELSE IF adherence to BMOD = {YES},
                   THEN stage 2 = {AUGMENT} or {INTENSIFY}.
      ELSE IF response status to stage 1 = {R}
            THEN continue stage 1
```



Estimated mean of more deeply tailored Al

Estimated optimal AI tailoring on prior med and adherence

	est	low upp
Mean Y under bmod, prior med	3.37	2.76 3.89
Mean Y under med, prior med	3.80	3.23 4.39
Mean diff (bmod-med) for prior med	-0.43	-1.17 0.33
Mean Y under bmod, no prior med	4.32	3.89 4.70
Mean Y under med, no prior med	2.93	2.53 3.36
Mean diff (bmod-med) for no prior med	1.39	0.78 1.89
Mean Y Deeply tailor AI	4.16	3.78 4.45
•		-

Estimated means of the four embedded AIs

	Estimate	95% LCL	95% UCL SE
Mean Y: AI#1 (MED, AUGMENT)	2.58	1.96	3.20 0.32
Mean Y: AI#2 (BMOD, AUGMENT)	3.79	3.21	4.38 0.30
Mean Y: AI#3 (MED, INTENSIFY)	1.78	1.02	2.54 0.39
Mean Y: AI#4 (BMOD, INTENSIFY)	2.61	1.95	3.26 0.33



Implementation

- Q-Learning software on d3c's website
 - R package qlaci
 - SAS procedure PROC GENMOD
- This software incorporates statistical adjustments that are necessary for obtaining the correct confidence intervals
- During vModule 3, I am going to show you how to use the qlaci package in R



References

Nahum-Shani, I., Qian, M., Almirall, D., Pelham, W. E., Gnagy, B., Fabiano, G. A., ... & Murphy, S. A. (2012). Q-learning: A data analysis method for constructing adaptive interventions. Psychological methods, 17(4), 478.

Ertefaie, A., Deng, K., Wagner, A. T., & Murphy, S. A. (2014). *qlaci R package for using q-learning to construct adaptive interventions using data from a SMART (Version 1.0)*. University Park: The Methodology Center, Penn State. Available from methodology.psu.edu.



Q&A

4 10 min

