50 minutes

This is the 6th (final) module of a 2-day 6-module workshop on experimental designs for building optimal adaptive health interventions.

By now, you know what an ATS is. You have discussed why they are important in terms of managing chronic disorders (indeed, an ATS formalizes the type of clinical practice taking place today). You have been introduced to the SMART clinical trial design, the rationale for SMARTs, and some important SMART design principles. Also, you have been introduced to typical primary aims and their associated data analysis methods.

In this module, we are going discuss Q-learning, a new type of data analysis method used as a secondary research aims using data arising from SMART studies.
Secondary Aims, Outline

- Discuss what is a “more deeply-tailored ATS”
- Review of auxiliary data typical of SMARTs
- Q-Learning: Using backward-induction logic
  - S(a): Analysis of time-varying tailoring variables of second-stage treatment
    - There is one step to this: Step (i)
  - S(b): Analysis of baseline tailoring variables of first-stage (assuming an optimal second-stage treatment)
    - There are two steps to this: Steps (ii) and (iii)
What is a more deeply-tailored ATS?

• To understand this, we first review what are the “embedded ATSs” within the ADHD SMART study

• Recall there are 4 SMART-design embedded ATSs.
You were introduced to this SMART in modules 1, 2, 3, 4 and 5.
These are the 2 embedded ATS that employ the “add other treatment” tactic.
These are the 2 embedded ATSSs that employ the “intensify initial treatment” tactic

Medication
- Responders → Continue Medication
- Non-Responders → Increase Medication Dose → Add Behavioral Intervention

Behavioral Intervention
- Responders → Continue Behavioral Intervention
- Non-Responders → Increase Behavioral Intervention → Add Medication
So why consider a more deeply-tailored ATS?

- First, it may be that some participants (e.g., those who have used MED in the past) may benefit more from MED vs BMOD.
- Second, certain types of non-responders (e.g., those who do not adhere to initial treatment) may benefit more from INTENSIFY vs ADD.
Notice this is more individualized than the 4 embedded ATS that were part of the SMART study design.

Also: note the similarities between this ATS and the Drug Court ATS proposed and implemented by Doug Marlowe (see Module 1 on Adaptive Treatment Strategies)
The remaining slides in this Module are devoted to understanding how to use auxiliary data arising from a SMART with a regression method known as Q-learning to develop/learn/discover a more deeply-tailored ATS such as the one shown on the previous slide.

But first, what do we mean by auxiliary data?
In addition to standard outcomes scales/measures, many other things could be measured during initial treatment (in this SMART study) that could be used in secondary analyses to more deeply tailor/individualize subsequent treatment, including:

- Allegiance/rapport of individual with the psychologist/psychiatrist,
- Environmental outcomes (parent outcomes, ...),
- Ecological momentary assessments (daily/weekly substance use patterns, rituals, etc.)

Notice that some of the O2 measures may be available for non-responders, but not available (e.g., “structurally missing”) for responders: an example of this in this ADHD study is the time until non-response!
How are O1 and O2 used?

- We can use the auxiliary data O1 to help decide which is best MED v BMOD for certain individuals
- We can use the auxiliary data O1 and O2 to help decide which is best INTENSIFY v ADD for certain individuals
The remaining slides in this Module are devoted to understanding how to use auxiliary data arising from a SMART with a regression method known as Q-learning to develop/learn/discover a more deeply-tailored ATS such as the one shown on the previous slide.
Using Q-Learning to develop a more richly-individualized ATS

Q-Learning is an extension of regression to sequential treatments.

- Q-Learning results in a proposal for an adaptive treatment strategy with greater individualization.
- A subsequent trial would evaluate the proposed adaptive treatment strategy versus usual care.

This is likely the most interesting & fun aim that one can investigate using data from a SMART design.

The name “Q-learning” refers to learning more about the “Quality” of an adaptive treatment strategy.
This is an idea borrowed from computer scientists.
Three Steps in Q-Learning Regression

*Work backwards (as you would for a project time-line)*

i. Do a **regression** to learn how to more deeply-tailor second-stage treatment using O1 and O2 (in ADHD SMART, this is only with non-responders)

ii. Assign each non-responder the value \( \hat{Y}_i \), an estimate of the outcome under the second-line treatment that yields best outcome for that person \( i \). Responders get their observed \( Y_i \).

iii. Using \( \hat{Y}_i \), do a **regression** to learn how to more deeply-tailor first-stage treatment using O1

3 Steps:

Step (i): Do a regression at stage 2 to learn about the optimal second-line treatment given characteristics of the participant at baseline and outcome during first-line treatment

Steps (ii) and (iii): Do regression using an outcome that already has taken into account future optimal treatment to learn about the optimal first-line treatment.

Conduct the regressions in backwards order! E.g. Stage 2 first, then Stage 1. Why?
   - Stage 1 dependent variable must control for Stage 2 treatment.
   - Stage 1 dependent variable is a predictor of \( Y \) under optimal treatment in stage 2.
   - Stage 2 analysis is used to construct \( \hat{Y} \)

We are going to demonstrate the Q-learning algorithm results with adherence as the candidate stage 2 tailoring variable; and the presence and acceptability of medication in the year prior to beginning the adaptive intervention as the candidate stage 1 tailoring variable. We are first going to go through all the steps in Q-Learning to gain intuition. Then, later, we will show you how to use SAS to implement Q-learning.
Why is this such an interesting question? Because if adherence to initial treatment strongly moderates the impact of increase/add on Y, then it can help decide which second stage treatment is best for which non-responder. In this example, adherence is a tailoring variable because it helps tailor stage 2 treatment.

(Or such an analysis may suggest a tailoring variable that we investigate in more detail in the next randomized trial).
Step (i) of Q-Learning: Second-stage treatment tailoring?

To accomplish Step (i) of Q-Learning, you may fit a regression model such as this:

\[ Y = b_1 + b_2 o_{11}c + b_3 o_{12}c + b_4 o_{13}c + b_5 o_{14}c + b_6 o_{21}c + b_7 a_1 + b_8 o_{22} + b_9 a_2*a_1 + b_{10} a_2*o_{22} + b_{11} a_2 + \text{error} \]

From such a model we would learn if, for example, \( o_{22} \) (adherence to first stage txt) is a strong moderator of the effect of \( a_2 \).

The parameters \( b_9, b_{10}, \) and \( b_{11} \) help us understand the effect of INTENSIFY vs ADD at stage 2 within levels of \( A_1 \) (previous treatment) and \( O_{22} \) (adherence to previous treatment). If \( A_1 \) and \( O_{22} \) are “strong moderators” of the effect of \( A_2 \), then we can use \( A_1 \) and \( O_{22} \) as tailoring variables.

That is, by “strongly moderate” we mean, for example, that adherence can be used as a tailoring variable in a decision rule for whether to intensify versus add txt.

Later, it will become clear why we put the “main effect” of \( a_2 \) as the last term in the regression. We recognize this is non-standard way of showing the model.
Step (i) of Q-Learning: Learn best second-stage treatment for non-responders

This is what we might learn from a regression such as the one shown on the previous slide.
Step (i) of Q-Learning: Learn best second-stage treatment for non-responders

Among adherers to MED, better to Intensify MED.

Higher Score is Better

This analysis is with simulated data.
Step (i) of Q-Learning: Learn best second-stage treatment for non-responders

Among adherers to BMOD, better to intensify but not by much (NS).

Higher Score is Better

This analysis is with simulated data.
Three Steps in Q-Learning Regression

Work backwards (as you would for a project time-line)

i. Do a regression to learn how to more deeply-tailor second-stage treatment using O1 and O2 (in ADHD SMART, this is only with non-responders)

ii. Assign each non-responder the value \( \hat{Y}_i \), an estimate of the outcome under the second-line treatment that yields best outcome for that person \( i \). Responders get their observed \( Y_i \).

iii. Using \( \hat{Y}_i \) do a regression to learn how to more deeply-tailor first-stage treatment using O1

3 Steps:

Step (i): Do a regression at stage 2 to learn about the optimal second-line treatment given characteristics of the participant at baseline and outcome during first-line treatment.

Steps (ii) and (iii): Do regression using an outcome that already has taken into account future optimal treatment to learn about the optimal first-line treatment.

Conduct the regressions in backwards order! E.g. Stage 2 first, then Stage 1. Why?

- Stage 1 dependent variable must control for Stage 2 treatment.
- Stage 1 dependent variable is a predictor of \( Y \) under optimal treatment in stage 2.
- Stage 2 analysis is used to construct hat\( \{Y\} \)

We are going to demonstrate the Q-learning algorithm results with adherence as the candidate stage 2 tailoring variable; and the presence and acceptability of medication in the year prior to beginning the adaptive intervention as the candidate stage 1 tailoring variable. We are first going to go through all the steps in Q-Learning to gain intuition. Then, later, we will show you how to use SAS to implement Q-learning.
Whereas, responders get the observed $Y$.

Step (iii) would then run a regression using this “new outcome” to determine the best first-stage treatment at different levels of baseline covariates.

We have SAS software that implements Steps 2 and 3 automatically and provides correct statistical inference. More on this next.
Three Steps in Q-Learning Regression

*Work backwards (as you would for a project time-line)*

i. Do a regression to learn how to more deeply-tailor the second-stage treatment using O1 and O2 (in ADHD SMART, this is only with non-responders)

ii. Assign each non-responder the value $\hat{Y}_i$, an estimate of the outcome under the second-line treatment that yields best outcome for that person $i$. Responders get their observed $Y_i$.

iii. Using $\hat{Y}_i$ do a regression to learn how to more deeply-tailor first-stage treatment using O1

3 Steps:

Step (i): Do a regression at stage 2 to learn about the optimal second-line treatment given characteristics of the participant at baseline and outcome during first-line treatment.

Steps (ii) and (iii): Do regression using an outcome that already has taken into account future optimal treatment to learn about the optimal first-line treatment.

Conduct the regressions in backwards order! E.g. Stage 2 first, then Stage 1. Why?

Stage 1 dependent variable must control for Stage 2 treatment.

Stage 1 dependent variable is a predictor of Y under optimal treatment in stage 2.

Stage 2 analysis is used to construct hat{Y}

We are going to demonstrate the Q-learning algorithm results with adherence as the candidate stage 2 tailoring variable; and the presence and acceptability of medication in the year prior to beginning the adaptive intervention as the candidate stage 1 tailoring variable. We are first going to go through all the steps in Q-Learning to gain intuition. Then, later, we will show you how to use SAS to implement Q-learning.
Step (iii) of Q-Learning: First-stage treatment tailoring?

To accomplish Step (iii) of Q-Learning, you may fit a regression model such as this:

\[
\hat{Y} = b_1 + b_2 \, o_{11}c + b_3 \, o_{12}c + b_4 \, o_{14}c + b_5 \, o_{13} + b_6 \, o_{13} \ast a_1 + b_7 \, a_1 + \text{error}
\]

From such a model we would learn if, for example, \( o_{13} \) (prior MED) is a strong moderator of the effect of \( a_1 \) such that it would make a good tailoring variable.

The parameters \( b_6 \) and \( b_7 \) help us understand the effect of BMOD vs MED at stage 1 within levels of prior medication use.
So, we should assign MED to kids with MED in prior year.
So, we should assign BMOD to kids who did not have MED in the prior year.
SAS software to perform Q-Learning

- We next show you how to do
- Step (i) using regression
- Steps (ii) and (iii) using a SAS add-on known as PROC QLEARN
- We will use the two example regression models shown previously.
Contrast 1: Among non-responders who adhere to MED, it is better to INTENSIFY treatment rather than ADD a different treatment (positive effect).

Contrast 2: Among non-responders who do not adhere to MED, it is better to ADD than to INTENSIFY (negative effect).

Contrast 3: Among non-responders who adhere to BMOD, it is better to INTENSIFY than ADD (but this effect is not significant at 0.05; p-value=0.45).

Contrast 4: Among non-responders who do not adhere to BMOD, it is better to ADD than to intensify (negative effect).
Try it yourself in SAS

• Go to the file:
  sas_code_modules_4_5_and_6_ADHD.doc
• **Copy the SAS code on Page 11**
• Paste into SAS Enhanced Editor window
• Press F8 or click the Submit button (the little running guy)
Contrast 1: Among non-responders who adhere to MED, it is better to INTENSIFY treatment rather than ADD a different treatment (positive effect).

Contrast 2: Among non-responders who do not adhere to MED, it is better to ADD than to INTENSIFY (negative effect).

Contrast 3: Among non-responders who adhere to BMOD, it is better to INTENSIFY than ADD (but this effect is not significant at 0.05; p-value=0.45).

Contrast 4: Among non-responders who do not adhere to BMOD, it is better to ADD than to intensify (negative effect).
SAS add-on: PROC QLEARN
What do we provide PROC QLEARN?

1. Data set with O1, A1, R, O2, A2, Y
2. The second stage regression model
   - Y ~ O1, A1, O2, A2
   - Specify among who to do this regression (e.g., non-responders in ADHD SMART)
3. The first stage regression model
   - ῦ ~ O1, A1
In Step (i) PROC QLEARN just reproduces the second stage regression you did by hand.

Step (iii) implements a special bootstrap procedure (Laber and Murphy, 2012; JASA) to produce appropriate statistical inferences (confidence intervals) concerning the first-stage regression parameters.

Laber and Murphy (2012; JASA) call these “adaptive confidence intervals”.

---

SAS add-on: PROC QLEARN
What does PROC QLEARN do?

1. It implements Step (i): Second-stage Regression
   • Provides us with second-stage regression parameter estimates
2. It implements Step (ii): Obtains \( \hat{Y} \)
   • It assigns non-responders the outcome under the best second-stage treatment based on Step (i).
   • It assigns responders their observed outcome.
3. It implements Step (iii): First-stage Regression
   • Provides first-stage regression parameter estimates
   • Provides appropriate confidence intervals for the first-stage regression parameter estimates
**SAS add-on: PROC QLEARN**

What does the syntax look like?

```sas
PROC QLEARN <options for input> ;
  MAIN1 variables ;
  TAILOR1 variables ;
  MAIN2 variables ;
  TAILOR2 variables ;
  RESPONSE variable ;
  STG1TRT variable ;
  STG2TRT variable ;
  STG2SAMPLE variable ;
  ALPHA value ;
RUN;
```

MAIN1 and TAILOR1 are used to specify the first-stage regression (next slide explains)
MAIN2 and TAILOR2 are used to specify the second-stage regression (next slide explains)
RESPONSE specifies the outcome variable Y
STG1TRT gives SAS the name of the A1 variable, must be coded -1/+1
STG2TRT gives SAS the name of the A2 variable, must be coded -1/+1
STG2SAMPLE is a 0/1 indicator variable specifying (if equal to 1) the sample used for the second-stage regression
ALPHA specifies the Type-I error used to calculate the confidence intervals for the first-stage regression
MAIN1 and TAILOR1 are used to specify the first-stage regression (next slide explains)
MAIN2 and TAILOR2 are used to specify the second-stage regression (next slide explains)
RESPONSE specifies the outcome variable Y
STG1TRT gives SAS the name of the A1 variable
STG2TRT gives SAS the name of the A2 variable
STG2SAMPLE specifies the sample used for the second-stage regression
ALPHA specifies the Type-I error used to calculate the confidence intervals for the first-stage regression
Recall our second-stage regression model

data dat10; set dat1; if R=0; run; * use only non-responders;
proc genmod data = dat10;
    model y = o11c o12c o13c o14c a1 o21c o22 a2 a2*a1 a2*o22;
run;

This is how that same model is specified in PROC QLEARN

data dat11; set dat1; S = 1-R; run; ^ use only non-responders;
proc qlearn data=dat11;
    ...[next slide]
    main2   o11c o12c o13c o14c o21c;
    tailor2 a1 o22;
    stg2sample s;
    response y;
    stg2trt a2; ...[next slide]
run;

Y = b1 + b2 o11c + b3 o12c + b4 o13c
    + b5 o14c + b6 o21c + b7 a1 + b8 o22
    + b9 a1*a2 + b10 o22*a2 + b11 a2 + error

Notice this tells PROC QLEARN to do the second stage regression with only the S=1 participants which are the R=0 participants, which are the non-responders who were re-randomized.

In the next slide we show how to specify the first-stage regression by showing a complete specification for PROC QLEARN.
There are other options (optional) that we do not describe in the slide. The User’s Guide explains these in more detail.

The PROC QLEARN Full Example, ADHD Data:

```sas
data dat1; set dat1; S = 1-R; run;
proc qlearn data=dat1
   contrast1=contrast1
derive1;
   main1 o11c o12c o14c;
tailor1 o13;
   main2 o11c o12c o13c o14c o21c;
tailor2 a1 o22;
   stg2sample s;
   response y;
   str1trt a1;
   str2trt a2;
run;
```

This will ask SAS to fit the following two regressions:

**Stage 2:**

\[ Y = b1 + b2 \ o11c + b3 \ o12c + b4 \ o13c + b5 \ o14c + b6 \ o21c + b7 \ a1 + b8 \ o22 + b9 \ a1*a2 + b10 \ o22*a2 + b11 \ a2 + \text{error} \]  

    (fit only with non-responders, S=1)

**Stage 1:**

\[ \hat{Y} = b1 + b2 \ o11c + b3 \ o12c + b4 \ o14c + b5 \ o13 + b6 \ o13*a1 + b7 \ a1 + \text{error} \]
This is the example contrast matrix that we use with the ADHD data.
Try it yourself in SAS

- Go to the file:
  sas_code_modules_4_5_and_6_ADHD.doc
- **Copy the SAS code on Page 14**
  - This code defines the contrast matrix and runs PROC QLEARN
- Paste into SAS Enhanced Editor window
- Press F8 or click the Submit button (the little running guy)
My results for the estimates will be identical to yours. My results for the confidence intervals will be different from yours. This is because the confidence intervals by Laber and Murphy (2012) are based on a bootstrapping procedure that re-samples the data. Not shown on this slide (but it will show on your output screen are the parameter estimates for the stage 2 model). PROC QLEARN provides these so you can be sure you implemented the correct stage 2 model that you implemented previously. You should check you got the same answers as for the model you ran on Slide 20.

Interpretations:
You can see that medication in the prior year is, indeed, a significant tailoring variable. That is, the sign for BMOD vs MED changes (contrast 3 vs contrast 6) depending on the level of medication in the prior year. Since contrast 3 is borderline, it would not be surprising if some of you see that this interval covers zero.
What did we learn with Q-learning?

*Adaptive Treatment Strategy Proposal*

- If the child used MED in prior year, then begin with MED; otherwise, begin with BMOD.

- If the child is non-responsive and non-adherent to either first-line treatment, then AUGMENT with the other treatment option.

- If the child is non-responsive but adherent to either first-line treatment, then it is better to INTENSIFY first-line treatment.

- If the child is responsive to first-line treatment, then CONTINUE first-line treatment.

Note that this is not 1 of the 4 embedded ATS as part of the trial design. This is a more deeply individualized ATS that is a function of much more than just early response status. Indeed, this ATS is a function of prior MED use, first-line txt, response status, adherence to first-line txt, and second-line txt.
The SAS code you received with this workshop shows you code for how to get the mean under the more deeply-tailored ATS discovered by Q-Learning. We do not have time to go over this in this Module.
Citations to Technical Reports


Practicum

*Autism Exercises:* As before, we will go through the Autism Starter File to continue practicing/working through these primary data analyses using the Autism data set.
Extra Slides: For Statistical Aficionados!

In the next two slides, we actually do Steps (ii) and (iii) of Q-Learning by hand. This is what the PROC QLEARN software automatically does. The issue, however, is that the standard errors here are incorrect. PROC QLEARN calculates the appropriate standard errors.

The SAS code for this is on Pages 12 and 13 of the sas_code....doc file.
This is Step (ii) of Q-learning (done manually): This SAS code manually assigns yhat to prepare for the Step (iii) regression. The coefficients in the definition of yhat are from the Step (i) model results. The absolute value does algebraically what was shown graphically in the previous slide: that is, it assigns the outcome had the child been assigned their optimal treatment at stage 2 (i.e., assigning ADD vs INTENSITY based on adherence to first-stage treatment). As you can see from the PROC MEANS output, yhat has larger mean than y. This is expected if we did it right!!

Note:
We are doing all of this manually here, but in forthcoming slides we describe SAS software (PROC QLEARN) that does all of this automatically!
Step (iii) of Q-learning done manually.

Contrast 1: Among children who had medication (and found it acceptable) in the prior year, then starting with MED is better (negative effect).

Contrast 2: Among children who did not have medication (or did and found it unacceptable) in the prior year, then starting with BMOD is better (positive effect).

HOWEVER, WE CANNOT TRUST THE 95% CONFIDENCE INTERVALS PROVIDED BY THE REGRESSION PROCEDURE HERE !!!