Preparing for a SMART Study

Getting SMART About Developing Individualized Sequences of Adaptive Health Interventions

Association for Behavioral and Cognitive Therapies
November 10, 2011

Susan A. Murphy & Daniel Almirall

Outline

- Briefly, discuss some preliminary data analyses that could help justify a SMART
- We discuss scientific, logistical, and statistical issues specific to executing a SMART that should be considered when planning a SMART (e.g., in a SMART pilot study)
  - Sample size calculation for SMART pilots

We discuss issues pertinent to SMART study preparation only.

Of course, investigators should also rely on standard RCT design principles/preparation in addition to what we discuss.

Preliminary Data Analyses

- Suppose you observed that once a patient had 2 missed clinic visits, the chances of them coming back to treatment or responding in the future were lowest (closest to zero)?
- Consider appropriate framework for analyzing time-varying treatments
  - Effects of sequences of treatment
  - Effect of naturalistic switching
  - Time-varying moderators

Pilot Studies

Primary Aim of Pilot Studies

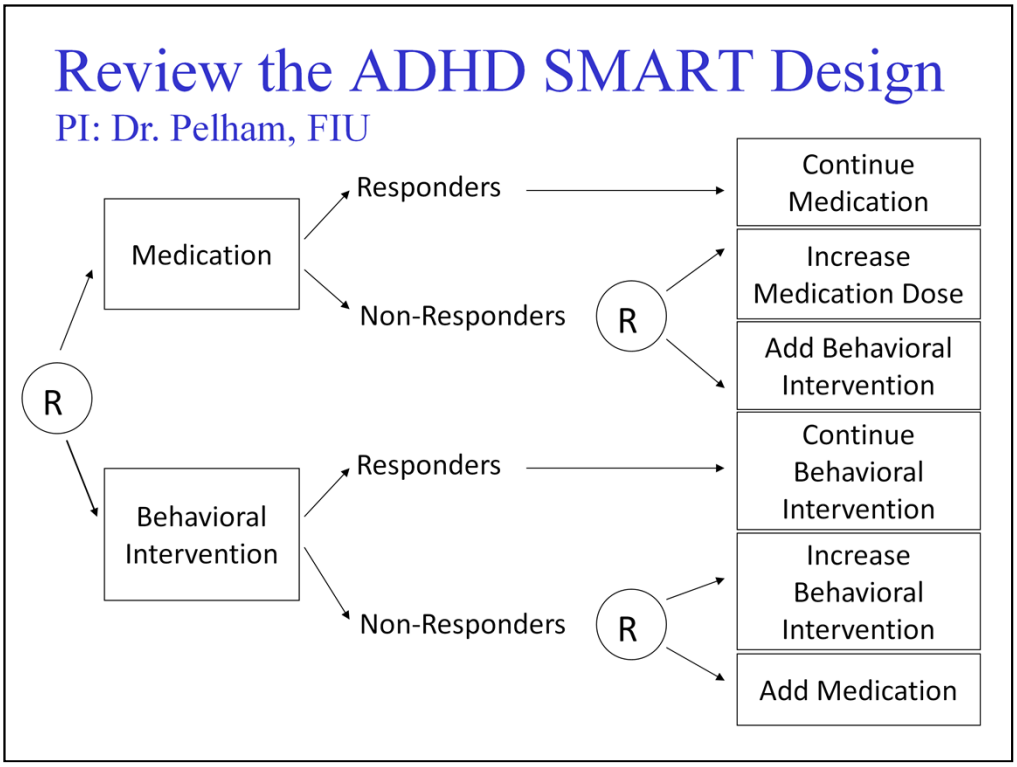
- Is to examine feasibility of full-scale trial: e.g.,
  - Can investigator execute the trial design?
  - Will participants tolerate treatment?
  - Do co-investigators buy-in to study protocol?
  - To manualize treatment(s)
  - To devise trial protocol quality control measures
- Is not to obtain preliminary evidence about efficacy of treatment/strategy, nor ES to power.
  - Rather, in the design of the full-scale SMART, the min. detectable effect size comes from the science.

First review the general principles behind pilot studies.

The key point here is that pilot studies are NOT used to obtain "preliminary evidence of efficacy". Citations for this follow:

Citations for Role of Pilot Studies

- Leon AC, Davis LL, Kraemer HC. (2011) The role and interpretation of pilot studies in clinical research. *Journal of Psychiatry Research.*
- Kraemer HC et al. (2006). Caution regarding the use of pilot studies to guide power calculations for study proposals. *Arch Gen Psychiatry.*
- Thabane L, Ma J, et al. (2010). A tutorial on pilot studies: the what, why, and how. *BMC Medical Research Methodology*.
- Lancaster GA, et al. (2004). Design and analysis of pilot studies: recommendations for good practice. *Journal of Evaluation in Clinical Practice.*



You were introduced to this SMART in modules 1 and 2.

Let's review.

Primary/Design Tailoring Variable

- Explicitly/clearly define early non/response
- We recommend binary measure
  - Theory, prior research, conventions, and/or preliminary data can be used to find a cut-off.
- Need estimate of the non/response rate
  - Using data from prior trials; or maybe in a pilot
- Should be associated with long-term response
  - Surrogate marker or mediation theories
- Should be easily assessed/measured in practice

All of these can also be investigated in a SMART pilot

How often should early non/response be evaluated? Only At the end of 8 weeks? Or throughout? If responder status is measured at the end of 8 weeks, say, should non-responders move to second stage if they fail to respond at ANY point between week 1 and week 8?

Need estimate of the non/response rate because (1) could be used in sample size calculations, and (2) can be used to justify the need to sequencing or second-stage treatments.

Protocol for Missing Primary Tailoring Variable

- Suppose participant misses clinic visit when the primary tailoring variable is assessed
  - How do we assign second stage treatment if/when participant returns?
- This is a non-standard missing data issue
- Need a fixed, pre-specified protocol for determining responder status based on whether/why primary tailoring variable is missing. Guided by actual clinical practice.

Non-standard missing data issue in the sense that this is not a problem about how we deal with missing data in data analysis (this issue is also important but is outside scope of this seminar).

Rather this is a problem with how we deal w missing data during the execution of the trial (e.g., participant misses the tailoring variable which guides entry to second stage, but returns to the study).

Of course, investigators wil have a retention plan and may have a way they assess the missing tailoring variable by letter, phone, email, or text message, say. But beware: the primary tailoring variable may have different considerations because it is not a variable used in assessment/evaluation of the treatment (strategies). Rather, it is used to guide subsequent treatment. More on this below.

Example Protocol for Missing Primary Tailoring Variable

- Need a fixed, pre-specified protocol for determining responder status based on whether/why primary tailoring variable is missing. Guided by actual clinical practice.
- Example 1: Classify all participants with missing response as non-responders.
- Example 2: Classify all participants with missing response as responders.
- Ex3: Need a third category for those missing?

Note: the above examples may also include "windows" of time where it is still ok to measure the primary tailoring much later even if participant was not around at the appropriate/ideal clinic visit. This flexibility is often exercised in RCTs. But this has to be operationalized ahead of time. This is the critical point here.

The actual protocol used will depend on particular research questions and types of disorders and treatments being studied. The protocol has to be well-justified and make sense scientifically. Above we give "non-responding until proven responding" and "responding until proven non-responding". The actual protocol could also be somewhere in between based on additional data about the missingness. For example, Investigators may need to differentiate between excused and unexcused missed assessment . Another example: may depend on how long the participant has been missing.

Manualizing Treatment Strategies

- Recall: SMART participants move through stages of treatment as part of embedded ATSs
- Treatment strategies are manualized
  - Not just the treatment options by themselves
  - Includes transitions between treatment options
- Treatment has an expanded definition here
  - Example: stepping down is a treatment decision
- Recall: randomization is not part of treatment

The point here is that manualizing treatments for a SMART may require additional thinking beyond what we typicaly do when we manualized treatments in standard RCTs. It may not be enough to know that each treatment option at each decision point, by itself, is manualized. The key here is to also manualize the transitions between treatment options, etc. Consideration of the fact that within a SMART, there may be X number of ATSs and usually 2*X number of treatment sequences is critical to this endevour.

Prepare to Collect Other Potential Tailoring Variables

- Additional variables used in secondary aims that could be useful in tailoring treatment
- Pilot new scales, instruments, or items that could be used as tailoring variables in practice
- Have protocols for discovering additional unanticipated tailoring variables:
  - Process measures (e.g., allegiance with therapist)
  - Use focus groups during and at end of pilot
  - Use exit interviews during and at end of pilot

Be clear here that the "unanticipated tailoring variables" I am referring to are additional variables NOT part of the study design. That is, there is a primary tailoring variable in SMARTs (usually early non/response) that is used to restrict subsequent randomizations. These are not used as part of the study design, but could be used later to further individualize treatment. These are candidate tailoring variables.

Evaluation Assessment versus Treatment Assessment

- Use (blinded) independent evaluators to collect outcomes measures used to evaluate effectiveness of embedded ATS
- But acceptable to use treating clinicians to measure the primary tailoring variable used to move to second-stage of treatment
  - Why? Because this is part of the intervention!
- SMART Pilot study can be used to practice protocols to keep these distinct

The first point is no different in SMARTs as in any other trial.

The key here is the distinction between treatment assesssments which could be used to guide treatment decisions vs evaluation assessments used in the data analysis (much later) to build an optimal ATS and/or to evaluate treatment effectiveness/efficacy.

Staff Acceptability to Changes in Treatment

- Challenges in a SMART:
  - Researchers maybe not accustomed to protocolized treatment *sequences/strategies*
  - SMART may limit use of clinical judgement
- Use a pilot SMART to identify concerns by staff and co-investigators about
  - Assessment of early non/response
  - Sequences of treatment provided
- Ex: clinician wants to classify early-nonrspder

For instance, in our example SMART, consider a child randomized to receive MED as first stage treatment and that prior to week 8, say at week 6, the treating clinician is concerned that the child is worsening and insists that the child be immediately moved to the next stage of treatment (that is, prior to reaching the protocolized 8 week mark).

Is this an indication that the definition and timing of non-response should be revised prior to the full-scale SMART? Do staff members need training in how to manage these emergent clinical situations in a consistent manner? Can something be learned from this situation that will refine and improve the treatment protocol? A SMART pilot can be used to identify when and where staff flexibility is warranted, to develop fidelity measures for its continued assessment, and to receive staff feedback about the timing of protocol-specified treatment switches and augmentations.

In this setting, suggest that regular staff/investigator group meetings should occur during which a clinical staff member who wants to deviate from the protocol for a patient must formulate a cogent argument for deviating from the rule. Deviations should only occur with consensus by group and should be documented.

Participant Adherence/concerns about Changes in Treatment

- Use the pilot SMART to identify concerns by participants using
  - Focus groups, exit interviews, or additional survey items
- May ask participants about
  - Experience transitioning between treatments
  - Was rationale for treatment changes adequate?
  - Was appropriate information you shared with clinician(s) in stage 1 understood by stage 2 clinician(s)?

Randomization Procedure

- A SMART pilot will allow investigators to practice re-randomization procedures
- We are referring to actual "coin flipping" here
  - Patient meets inclusion criteria, consent/assent, and we randomize him/her
  - In the typical 2-arm RCT we do this by blocked, stratified randomization
- Before we go on: Let's review what it means to block and stratify randomizations.

Use ADHD SMART design to explain real-time vs up-front randomization

Explain what I mean by randomization procedure. This refers to the actual coin-flip of which there are 3 (each patient gets randomized twice) in our example SMART.

Real-time may require "clinical trial support" software, now becoming standard in clinical trial design. SMART pilot will allow investigator to practice using this.

In the real-time approach: Will you use a minimization scheme or a permuted, stratified blocking scheme? Need to assess feasibility of doing real time approach. Assess feasibility of using the clinical trial support systems/software, if any.

Need to remind audience why we do stratified randomizations in the first place so everybody can be on the same page with the logic of all this here

Randomization Procedure

- A SMART pilot will allow investigators to practice re-randomization procedures
- Up-front versus real-time randomization
  - Up-front: After baseline, randomize participants to the embedded ATSs
  - Real-time: Randomize sequentially
- We recommend real-time because we can balance randomized second stage options based on responses to initial treatment.

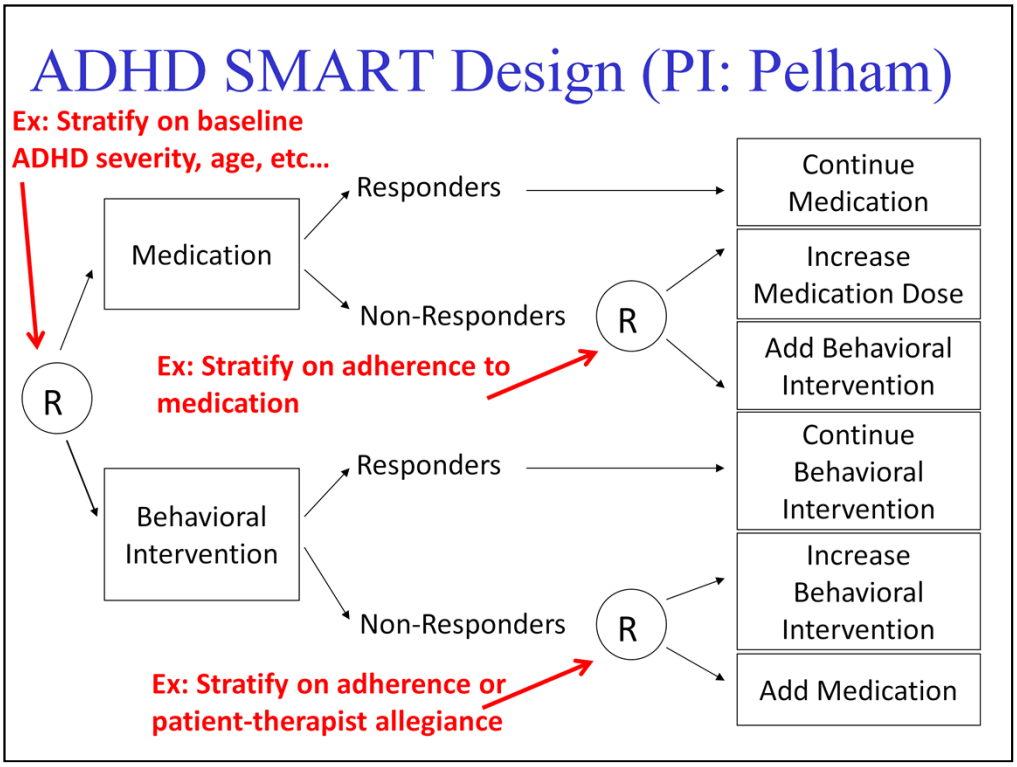
Use ADHD SMART design to explain real-time vs up-front randomization

Explain what I mean by randomization procedure. This refers to the actual coin-flip of which there are 3 (each patient gets randomized twice) in our example SMART.

Real-time may require “clinical trial support” software, now becoming standard in clinical trial design. SMART pilot will allow investigator to practice using this.

In the real-time approach: Will you use a minimization scheme or a permuted, stratified blocking scheme? Need to assess feasibility of doing real time approach. Assess feasibility of using the clinical trial support systems/software, if any.

Need to remind audience why we do stratified randomizations in the first place so everybody can be on the same page with the logic of all this here

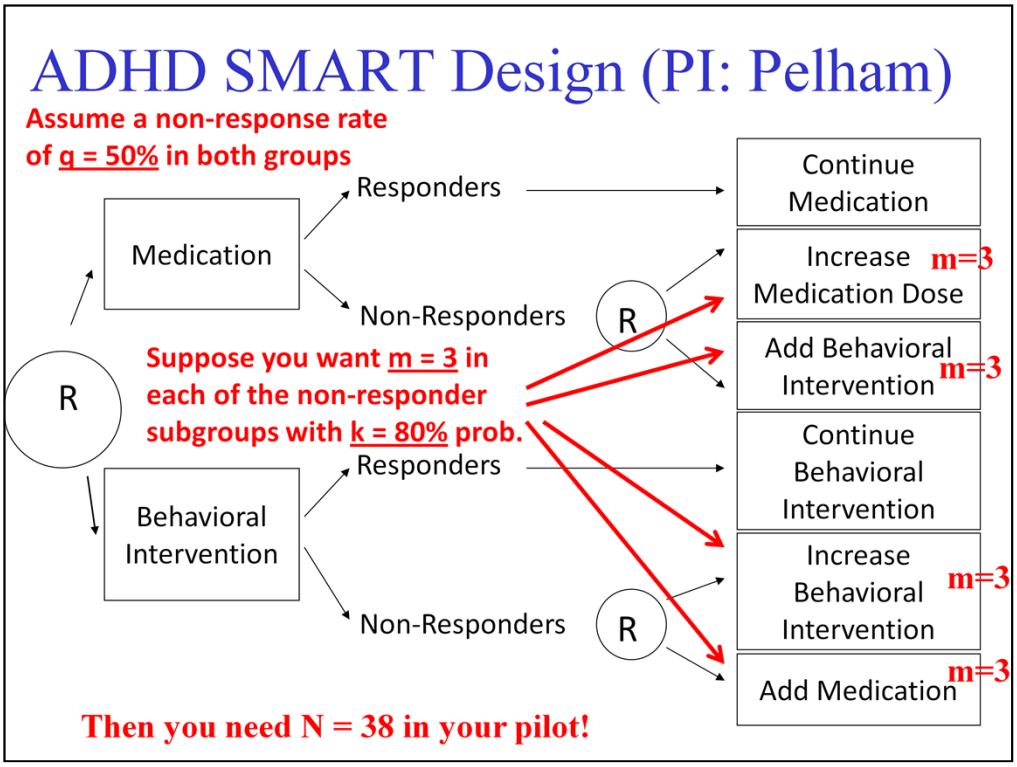


You were introduced to this SMART in modules 1 and 2.

Let's review.

Sample Size for a SMART Pilot

- Sample size calculation based on feasibility aims, not treatment effect detection/evaluation
- Approach 1: Primary feasibility aim is to ensure investigative team has opportunity to implement protocol from start to finish with sufficient numbers
- Choose pilot sample size so that with probability *k*, at least *m* participants fall into non-responder sub-groups (the "small cells")
  - Investigator chooses *k* (say 80%) and *m* (say 3)



How many participants are needed in the SMART pilot ?

N required	q = anticipated non-response rate						
	0.35	0.40	0.45	0.50	0.55	0.60	0.65
$k = 0.80$							
$m = 2$	42	36	32	28	26	22	20
$m = 3$	56	48	42	38	34	30	28
$m = 4$	70	60	52	46	42	38	34
$m = 5$	82	72	62	56	50	46	42
$k = 0.85$							
$m = 2$	44	38	34	30	26	24	22
$m = 3$	58	50	44	40	36	32	28
$m = 4$	72	62	54	48	44	40	36
$m = 5$	86	74	66	58	52	48	42
$k = 0.90$							
$m = 2$	48	40	36	32	28	26	22
$m = 3$	62	54	46	42	38	34	30
$m = 4$	76	66	58	52	46	42	38
$m = 5$	90	78	68	60	54	50	44

Sample Size for a SMART Pilot

- Approach 2: To obtain estimate of overall non/response rate with a given margin of error
  - Point estimation with precision
  - Usually requires larger sample than Approach 1
  - Use this approach if there is very poor information available about non/response rate
- 95% MOE = 2*SQRT(p (1-p) / N)
- Example 1: p=0.35, MOE=0.15 requires N=41
- Example 2: p=0.50, MOE=0.10 requires N=100

p is guess at the unknown response rate we are trying to estimate !

In example 2, we assume p=0.5, which is the most conservative guess, ie, leading to largest N.

This approach is interesting, but we think Approach 1 is more easily justified. This approach could be used as additional justification.

Citations

- Almirall D, Compton SN, Gunlicks-Stoessel M, Duan N, Murphy SA (under review). Designing a Pilot SMART for Developing an Adaptive Treatment Strategy.
  - Available as Technical Report at The Methodology Center!
- Leon AC, Davis LL, Kraemer HC. The role and interpretation of pilot studies in clinical research. *Journal of Psychiatry Research.*
- Kraemer HC et al. (2006). Caution regarding the use of pilot studies to guide power calculations for study proposals. *Arch Gen Psychiatry.*

Practice Exercise

Exercise: *Write down data sources available to you that you could use as preliminary data for a SMART. If you would like to do a SMART pilot, what is the primary feasibility aim?*

Next up: Primary Aims Using Data Arising from a SMART