

SMARTer Discontinuation Trial Designs for Developing an Adaptive Treatment Strategy

Daniel Almirall, Ph.D.,¹ Scott N. Compton, Ph.D.,² Moira A. Rynn, M.D.,³
John T. Walkup, M.D.,⁴ and Susan A. Murphy, Ph.D.⁵

Abstract

Objective: Developing evidenced-based practices for the management of childhood psychiatric disorders requires research studies that address how to treat children during both the acute phase of the disorder and beyond. Given the selection of a medication for acute treatment, discontinuation trials are used to evaluate the effects of treatment duration (e.g., time on medication) and/or maintenance strategies following successful acute-phase treatment. Recently, sequential multiple assignment randomized trials (SMART) have been proposed for use in informing sequences of critical clinical decisions such as those mentioned. The objective of this article is to illustrate how a SMART study is related to the standard discontinuation trial design, while addressing additional clinically important questions with similar trial resources.

Method: The recently completed Child/Adolescent Anxiety Multimodal Study (CAMS), a randomized trial that examined the relative efficacy of three acute-phase treatments for pediatric anxiety disorders, along with a next logical step, a standard discontinuation trial design, is used to clarify the ideas. This example is used to compare the discontinuation trial design relative to the SMART design.

Results: We find that the standard discontinuation trial can be modified slightly using a SMART design to yield high-quality data that can be used to address a wider variety of questions in addition to the impact of treatment duration. We discuss how this innovative trial design is ultimately more efficient and less costly than the standard discontinuation trial, and may result in more representative comparisons between treatments.

Conclusions: Mental health researchers who are interested in addressing questions concerning the effects of continued treatment (for different durations) following successful acute-phase treatment should consider SMART designs in place of discontinuation trial designs in their research. SMART designs can be used to address these and other questions concerning individualized sequences of treatment, such as the choice of a rescue treatment in case of postacute phase relapse.

Introduction

DEVELOPING EVIDENCED BASED PRACTICES for the management of childhood psychiatric disorders requires research studies that address how to treat children during both the acute phase of the disorder and beyond. There is currently a strong evidence base for the efficacy/effectiveness of acute-phase treatments for most common childhood and adolescent psychiatric disorders. Studies are now needed that examine intervention approaches for achieving continued response, remission, and sustained recovery following successful acute treatment.

Discontinuation trials in psychiatry are commonly used to evaluate the impact of treatments following successful acute-phase treatment. In the standard discontinuation trial, all patients recruited into the study are initially treated with an acute-phase treatment. At the end of the acute period, treatment responders are randomized to either discontinue treatment or continue with treatment for specified durations. In the majority of discontinuation trials, the primary outcome is the time to first relapse in the post-acute phase. For example, Emslie et al. (2008) used a discontinuation trial to answer the following question: “Among depressed children and adolescents who acutely respond to 12 weeks of

¹Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, Michigan.

²Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham, North Carolina.

³New York State Psychiatric Institute–Columbia University Medical Center, New York, New York.

⁴Weill Cornell Medical College and New York Presbyterian Hospital, New York, New York.

⁵Department of Statistics and Institute for Social Research, University of Michigan, Ann Arbor, Michigan.

Portions of this article were presented at the NIMH–New Clinical Drug Evaluation Unit (NCDEU) Conference in Phoenix, Arizona on May 28, 2008.

Funding was provided by National Institute of Mental Health (NIMH) grants R01-MH-080015 (Dr. Murphy) and K23-MH-075843-04 (Dr. Compton), and National Institute on Drug Abuse (NIDA) grant P50-DA-010075 (Dr. Murphy).

fluoxetine, what is the effect of immediately discontinuing medication versus discontinuation after an additional 12 weeks of treatment on time to first relapse?"

Assessing the impact of different treatment durations is often the explicit and primary rationale for discontinuation trials. However, embedded within this rationale is the implicit assumption that sequential, individualized decision making is essential for the long-term management of chronic psychiatric disorders. This is evident by the fact that discontinuation trials inform what to do following acute-phase treatment (treatment sequencing), and inform how to treat participants who respond to both acute-phase treatment and postacute phase treatment (a form of adaptive, treatment individualization).

In childhood mental health treatment, sequential, individualized, decision making is critical as some children respond well to a particular treatment or length of treatment, whereas others do not. It is often necessary to alter or modify ongoing treatment to achieve maximum benefit or minimize risk/burden (Murphy et al. 2007a). Moreover, not all children who respond initially will need to remain on treatment, especially when considering the potential for adverse events during chronic treatment. Adaptive treatment strategies (ATs) have been used to formalize the sequential, individualized decision making that clinicians often use when providing treatment in actual clinical practice (Lavori et al. 2000; Lavori and Dawson 2007, 2008; Murphy 2005; Murphy et al. 2007a, b, c; Murphy and Almirall 2009). Discontinuation trials appear to provide the first basic step for developing ATs across the acute and postacute treatment phases. However, the standard discontinuation trial does not continue to follow participants who relapse, and, therefore, results in a missed opportunity in terms of 1) understanding the effect of treatment duration ATs on long-term outcomes, and 2) informing how to rescue relapsing participants.

Sequential Multiple Assignment Randomized Trials (SMARTs), which will be discussed in more detail, were developed explicitly for the purpose of developing ATs (Murphy 2005; Nahum-Shani et al. 2012). At the time of this writing, there are a number of SMARTs currently being conducted in a variety of fields, such as to develop ATs for children with attention-deficit/hyperactivity disorder (ADHD) (Pelham, personal communication; Nahum-Shani et al. to appear), alcohol abuse (McKay, personal communication), and autism (Kasari, personal communication) and substance abuse among pregnant women (Jones, personal communication), among others. Lei et al. (2011) describe the similarities and differences among the four above-mentioned SMART designs. The SMART designed in Thall et al. (2000), reported in Thall et al. (2007), and later re-analyzed by Wang et al. (in press) considers 12 ATs in the treatment of prostate cancer. These ATs begin with one of four chemotherapies; participants who respond to the initial chemotherapy remain on it, whereas participants who fail to respond to the initial chemotherapy are switched to one of the other three. Among the interesting findings from this SMART is that the rescue treatment in one of the better performing ATs was a treatment that performed poorly if provided initially.

This article examines the "missed opportunity" of the standard discontinuation trial and proposes the use of the SMART as an enhanced alternative to the standard discontinuation trials. Specifically, we argue that SMARTs are ultimately less costly and more efficient than standard discontinuation trials by maximizing the amount of useful scientific information obtained in a single trial, while still addressing the primary scientific aims of standard discontinuation trials.

To illustrate these concepts, we utilize the recently completed Child/Adolescent Anxiety Multimodal Study (CAMS; Walkup et al. 2008). First, we review the results of CAMS and for illustration we provide an example of a standard discontinuation trial that could follow CAMS. Then, we present ATs and SMARTs, and describe their relation to the standard discontinuation trial. Then, we describe two ways to enhance the standard discontinuation trial using a SMART. The final section offers a discussion.

CAMS and a Standard Discontinuation Trial

CAMS (Walkup et al. 2008) was a multisite, randomized placebo-controlled trial that examined the relative efficacy of cognitive behavioral therapy (CBT), sertraline (SRT), and their combination (COMB) against pill placebo (PBO) for the treatment of social anxiety disorder (SAD), generalized anxiety disorder (GAD), and social phobia in children and adolescents between the ages of 7 and 17. Four hundred eighty-eight participants were randomly assigned to one of the four acute-phase treatment conditions. At the end of 12 weeks of acute-phase treatment, investigators reported an 80.7% response rate among participants treated with COMB, a 59.7% among participants treated with CBT only, 54.9% among participants treated with SRT only, and a 23.7% response rate among participants treated with PBO. In CAMS, response was defined as a Clinical Global Impressions-Improvement rating scale value of ≤ 2 . The results of CAMS showed that combination therapy (CBT+SRT), as well as the monotherapies (CBT alone or SRT alone), constituted three effective treatments (relative to no treatment) for a diverse population of children and adolescents with three of the most common and disabling pediatric and adolescent anxiety disorders.

Despite the success of the three active treatments reported in CAMS, important public health questions remain unanswered about long-term treatment of pediatric anxiety disorders. For example, given the short duration of medication treatment in most clinical trials (e.g., 12 weeks in CAMS) relative to the duration of medication treatment reported in community settings (e.g., 25% of 9–18-year-old new antidepressant users may complete at least 6 months of medication treatment; Shireman et al. 2002), concerns about side effects resulting from long-term medication exposure (Homberg et al. 2010), reservations on behalf of parents about medication use (Young et al. 2005, 2006), and the genuine need on behalf of clinicians and parents alike about how best to maintain symptom improvement, achieve remission, and prevent relapse, lead to the following important questions:

1. "What is the impact of discontinuing SRT treatment versus continuing SRT in children who respond to 12 weeks of COMB treatment?"
2. "Should the previous decision depend upon characteristics of the child?"

These, and similar questions, were described as significant public health concerns even before the completion of CAMS (Pine 2002). Discontinuation trials are intended primarily to address the timing question (1), or from a parent's perspective, "How long does my child need to continue on medication now that she is better?" Further, secondary data analyses of discontinuation trials can investigate whether the impact on relapse rates of one discontinuation strategy versus another depends upon child characteristics or child outcomes collected during acute-phase treatment (2).

Figure 1A depicts a standard discontinuation trial that could be viewed as the next logical step following the outcomes observed

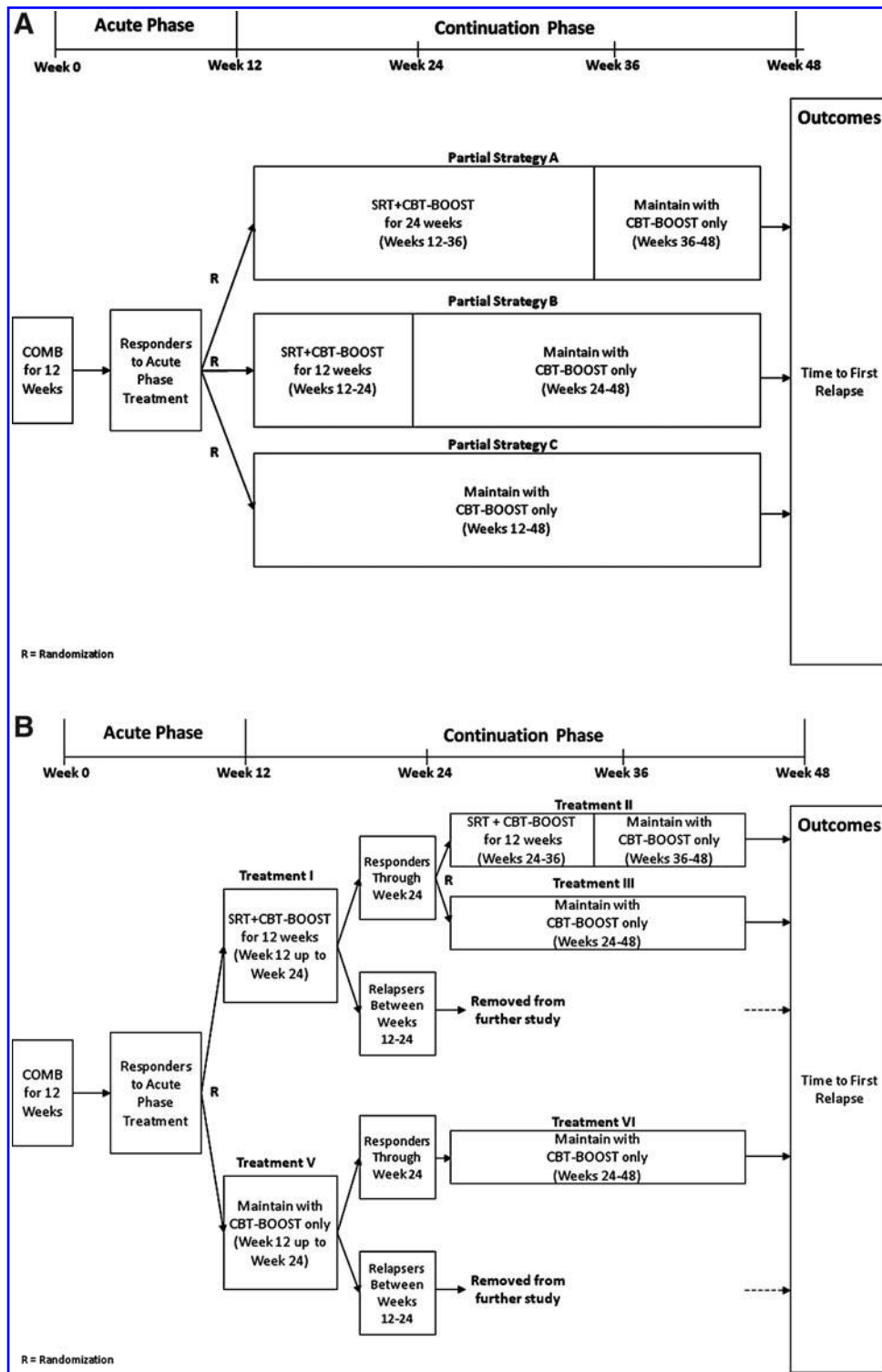


FIG. 1. (A) Example of a standard discontinuation trial. In standard discontinuation trials, relapsers (at any point) are removed from further study. Partial Strategy A and Partial Strategy B are identical up to week 24. The primary aim is a comparison of Partial Strategy A versus Partial Strategy B. This is equivalent to a comparison of time until relapse from week 24 onward, among responders up to week 24. (B) An alternate presentation of the standard discontinuation trial shown in (A). In standard discontinuation trials, relapsers (at any point) are removed from further study. The primary aim discussed is a comparison of Treatment II versus Treatment III in terms of time until first relapse. This is a comparison of time until relapse from week 24 onward, among responders up to week 24. SRT = sertraline medication; CBT = cognitive behavioral therapy; COMB = SRT + CBT; CBT-BOOST = CBT booster sessions.

in CAMS. This trial is designed to address the timing of SRT discontinuation following successful acute-phase response to COMB therapy in the context of once-monthly continued CBT booster sessions (CBT-BOOST). CBT-BOOST is designed to help patients manage minor stressors as medication is removed, thereby potentially preventing the unnecessary restart of medication, and also identifying patients who are likely medication dependent. In the treatment of obsessive-compulsive disorders, expert guidelines recommend once-monthly CBT booster sessions during medication discontinuation. (March et al. 1997). The options to which participants are randomized in a standard discontinuation trial are only “partial” treatment strategies, because they do not formally operationalize what to do for participants who relapse. In this example trial, all participants receive 12 weeks of COMB treatment at baseline. Following 12 weeks, responders are then randomly assigned to one of three continuation-phase treatment arms:

Partial Strategy A: Remain on SRT+CBT-BOOST for an additional 24 weeks, then taper SRT while remaining on CBT-BOOST;

Partial Strategy B: Remain on SRT+CBT-BOOST for an additional 12 weeks, then taper SRT while remaining on CBT-BOOST; or

Partial Strategy C: Taper SRT immediately while remaining on CBT-BOOST.

This trial is intended primarily to compare time to first relapse between different SRT duration times in the continuation phase. In this example of a standard discontinuation trial (and in the alternative study designs we consider subsequently), we assume that all participants can be followed for a maximum of 48 weeks (12 weeks acute phase + 36 weeks continuation phase). In the standard discontinuation trial, once the primary outcome has been observed (e.g., first relapse), participants are removed from further study.

ATSs and Discontinuation Trials

What is an ATS?

ATSs operationalize sequential, clinical decision making via a sequence of decision rules, one per clinical decision. In the context of ATSs, because the aim is to inform actual clinical practice, “treatment” is broadly defined to include all treatment management or strategy decisions. Hence, the decision to “discontinue

SRT,” for example, is thought to be a “treatment.” Or, as another example, if “SRT discontinuation” involves referral to the community where the patient might receive other concomitant treatment, then the “treatment” is “SRT discontinuation with the possibility of receiving community-based care.”

The decision rules utilize patient characteristics and observed patient outcomes to recommend whether, how, or when to alter the intensity, type, or delivery of treatment (Lavori et al. 2000; Collins et al. 2005; Murphy 2005; Murphy et al. 2007b; Lavori and Dawson 2007, 2008;). An example of a two-decision point, continuation-phase ATS among patients with anxiety disorders following response to 12 weeks of acute phase COMB treatment is shown in Figure 2.

In this simple ATS, responders to 12 weeks acute-phase COMB treatment (SRT+CBT) continued on the medication sertraline (SRT) while tapering CBT to monthly CBT booster sessions for an additional 12 weeks. At the end of 24 weeks, those participants who continued to show a positive treatment response were tapered off SRT while continuing to receive monthly CBT booster sessions for an additional 12 weeks. However, participants who relapsed at any point during weeks 12–24, restarted COMB treatment for an additional 12 weeks. This ATS addressed two critical decision points during the continuation phase: First, it addressed which treatment to provide during weeks 12–24; and second, given either continued response, or relapse, it addressed which treatment to provide next (during weeks 24–36 for responders; and up to week 36 for relapsers, depending upon the time of relapse).

From the perspective of the clinician, this ATS is a guide (i.e., a decision tree, as depicted in Fig. 2) for decision making at two critical decision points in the continuation phase. The assessment used to measure response/relapse during weeks 12–24 in Figure 2 is known as a *tailoring variable*, because future treatment is tailored (e.g., individualized) depending upon its value. In this ATS, relapse (or early signs of nonresponse) can be defined as meeting a criterion on the Clinical Global Impressions (CGI) scale (e.g., three or more on the CGI) or some other symptom measure collected during or after the initial continuation-phase treatment. From the perspective of the patient, an ATS is just a sequence of treatments. For example, a child who responds to acute-phase 12 week COMB treatment and continues to respond to SRT + CBT-BOOST during weeks 13–24, that child’s complete assigned continuation-phase treatment sequence is (SRT+CBT-BOOST weeks 13–24 followed by CBT-BOOST only weeks 25–36).

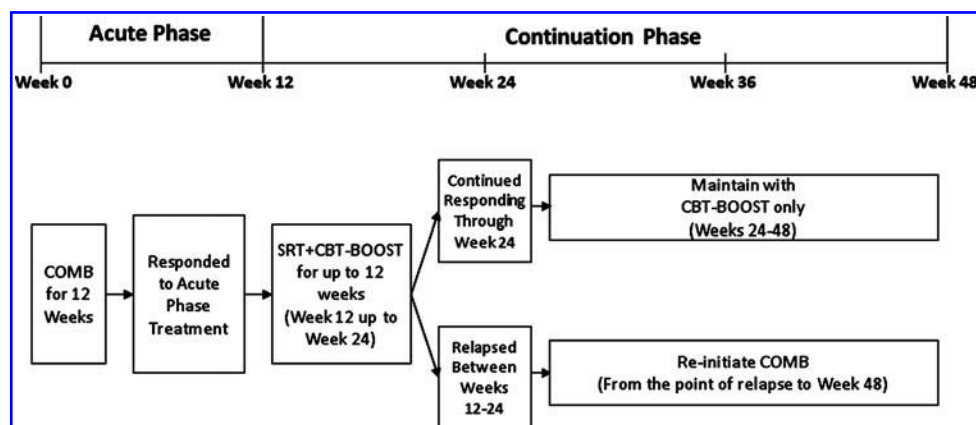


FIG. 2. A continuation-phase adaptive treatment strategy among responders to acute-phase treatment. SRT = sertraline medication; CBT = cognitive behavioral therapy; COMB = SRT + CBT; CBT-BOOST = CBT booster sessions.

Two characterizing features of the standard discontinuation trial

Although rarely stated explicitly, developing ATSS is one of the main reasons for conducting discontinuation trials in mental health research. This is evident from three characterizing features of the standard discontinuation trial design and analysis. First, the discontinuation trial is fundamentally concerned with how treatments should be sequenced, for example, acute-phase COMB followed by continuation-phase CBT-BOOST only (i.e., discontinue SRT). Second, in standard discontinuation trials, subjects who respond to the acute-phase treatment are randomized to two or more continuation-phase treatment alternatives. This study design choice reflects the fact that discontinuation trials are designed to inform the first critical decision about what to do following a successful initial, acute-phase treatment response. Third, in standard discontinuation trials, the continuation-phase treatment alternatives to which participants are randomized usually constitute different durations of the same treatment. This study design choice reflects the fact that discontinuation trials inform whether treatment should continue among those patients who are responding earlier in the continuation phase.

To better appreciate the third point, it is instructive to view the standard discontinuation trial reconfigured as a trial with a sequence of randomizations, one per critical decision. In Figure 1B, there are conceptually two randomizations: The first, for responding participants at week 13, to either continue treatment or discontinue treatment; and the second, for responding participants at week 25, to continue or discontinue treatment.

The discontinuation trial in Figure 1B, although it appears more complicated, is essentially identical to the trial depicted in Figure 1A. For example, a comparison of the time to first relapse between participants randomized to Partial Strategies A and B in Figure 1A is effectively a comparison of the time to first relapse between participants randomized to Treatment I followed by II versus those randomized to treatment I followed by treatment III in Figure 1B. This follows because in the trial in Figure 1A, participants in the two arms, Partial Strategy A and Partial Strategy B, are on the same treatment during the first 12 weeks.

The missed opportunity in the standard discontinuation trial

In the standard discontinuation trial, participants are removed from further study immediately after their first relapse; this can occur at any stage in the course of the experiment. From our perspective, this is a critical methodological weakness of this design, as it leads to a missed opportunity by prohibiting investigators to study how to clinically manage those who relapse while on postacute-phase treatment.

To better appreciate this issue, consider the following. First, discarding participants who relapse implicitly assumes that the *only* outcome of interest is time to first relapse, as all other outcomes are missing (by design) once a participant fails to respond. This eliminates the ability of investigators to examine the impact of treatment discontinuation on other longitudinal outcomes (for example, trajectory of longitudinal symptom counts, side effects, possible rescue treatments). This shortcoming is problematic in the context of childhood mental health in which symptoms often wax and wane (Witcher et al. 2002), and a variety of contextual factors often contribute to a relapse (e.g., stress of entering a new school year). It is important to understand the trajectory of outcomes beyond the initial relapse.

Second, once participants are removed, the opportunity to characterize how those who relapse might respond longitudinally on rescue treatment(s) is lost. For example, a longer duration of the acute treatment might make it easier to rescue patients who have relapsed than a shorter duration of acute treatment; however this question cannot be addressed in the traditional discontinuation design because investigators “disenroll” participants who relapse. That is, *by design* the standard discontinuation trial does not provide the data to assess the cumulative (or synergistic) effects of sequenced treatment decisions. This is a critical shortcoming, given that the development and evaluation of effective interventions for children who are treatment refractory is a public health priority.

Third, the practice of removing participants who relapse is not resource effective. This study design choice necessitates the execution of separate trials to investigate treatment options among participants who relapse in the continuation phase. However, investing additional resources to recruit an entirely new sample of relapsed participants to answer these questions is not an efficient use of limited research resources.

Fourth, and perhaps most important, the conduct of a separate trial for participants who relapse may lead to misleading results. The recruited sample of relapsed participants may differ from the population of patients who relapse after successful acute treatment. For example, the recruited sample may include a higher fraction of children with parents that are highly motivated and less likely to be discouraged by the relapse than the population at large. Higher motivation/lower levels of discouragement might lead to higher adherence to the rescue treatment. In this example, to the extent that family motivation influences child outcomes, the results of a separate trial on relapsed participants will not reflect the population effect of the available rescue treatments. Moreover, even when separate trials are conducted among those who relapse in the continuation phase, they do not allow investigators to observe the total effect of an initial continuation treatment, including the delayed (or “downstream”) effects mentioned previously.

Moving Beyond the Standard Discontinuation Trial Design

Improving the discontinuation trial using SMARTer designs

A discontinuation trial design that follows all participants (including those who relapse) and/or evaluates potential rescue treatments among postacute-phase relapsers would make better use of trial resources and offer enhanced scientific opportunities. Such a trial would allow investigators to fully inform the development of continuation-phase ATSS. Further, it would increase the number and types of scientific questions that could be answered in a single study design while still permitting the investigators to examine the original treatment discontinuation question(s), without altering the primary outcome analysis (time to first relapse).

We propose two improvements to the standard discontinuation trial. Both proposals constitute a Sequential Multiple Assignment Randomized Trial (SMART; Lavori and Dawson 2000; Murphy 2005; Lavori and Dawson 2007; Murphy et al. 2007c; Oetting et al. 2007). Both SMART proposals have these features: 1) participants are randomized multiple times over the course of a trial—that is, at each critical decision point—in order to experimentally evaluate which is the most appropriate treatment at each critical decision point, and 2) all participants are followed through the end of the trial to provide investigators the opportunity to evaluate and compare a variety of longitudinal outcomes (not just time to relapse)

under different ATSS. The standard discontinuation trial, exemplified by Figure 1, possesses feature 1 but not feature 2.

SMARTer Design A

In this design, participants who relapse at any point during the continuation phase are followed by the investigative team and undergo regularly scheduled outcome assessments until the end of the study. A rescue treatment for participants who relapse is delivered in a protocol-driven way (i.e., manualized) or if participants are referred out to community providers (i.e., treatment provided by a non-study clinician), the referral process is manualized and the treatment received in the community is documented. Figure 3A is an example of this SMART. Here, participants who relapse in the continuation phase return to COMB (a treatment known to have worked in the acute phase for this population). Note that this proposal involves only a minimal change from the standard discontinuation trial (Fig. 1B). Those participants who fail to respond are offered COMB treatment and continue to be followed in the trial. Regularly scheduled assessments should include detailed information about all treatments received over the full course of the trial, rather than solely up to the point of initial nonresponse, as is common in standard discontinuation trials.

Unlike the standard discontinuation trial (Fig. 1), the proposed SMARTer Design A (Fig. 3A) would allow investigators to address, for example, the following question: “Among responders through week 24 on continued SRT+CBT-BOOST, do the symptoms over the succeeding 12 weeks (weeks 24–36) vary between participants randomized to continue SRT (Treatment II) versus those randomized to discontinue SRT (Treatment III)?”

In addition to evaluating longitudinal symptoms, another important benefit of following SMARTer Design A (Fig. 3A) is that it provides investigators the ability to evaluate longitudinal outcomes for each of several ATSS. For example, the proposed design in Figure 3A includes three embedded continuation-phase ATSS. The three ATSS are:

ATS 1: Continue SRT+CBT-BOOST for 12 additional weeks. If the child continues to respond through week 24, then *maintain* SRT+CBT-BOOST for yet another 12 weeks; otherwise, if the child relapses anytime between week 12 and week 24, then immediately *re-initiate* COMB (SRT+full-protocol CBT).

ATS 2: Continue SRT+CBT-BOOST for 12 additional weeks. If the child continues to respond through week 24, then *step down* treatment to CBT-BOOST alone (by discontinuing SRT) for an additional 12 weeks; otherwise, if the child relapses anytime between week 12 and week 24, then *re-initiate* COMB.

ATS 3: Step down treatment to CBT-BOOST only (by discontinuing SRT medication) for first 12 weeks in the continuation phase. If the child continues to respond through week 24, then *maintain* CBT-BOOST only for an additional 12 weeks; otherwise, if the child relapses anytime between week 12 and week 24 after SRT discontinuation (with CBT-BOOST), then *re-initiate* COMB.

Note that **ATS 2** is the example ATS shown in Figure 2. In contrast, the standard discontinuation trial (Fig. 1B) does not include *any one* of these ATSS, because participants who relapse are “disenrolled” from the trial. Further, note that because the standard discontinuation trial design is embedded in the SMART in Figure 3A, it allows investigators to examine the original discontinuation question.

SMARTer Design B

This proposal builds upon the SMARTer Design A. It can be used to provide evidence for which treatment to provide to participants who relapse in the continuation phase. Here, participants who relapse after having discontinued SRT in the continuation phase are re-randomized between two rescue treatment options (Fig. 3B). This SMART study enjoys all of the advantages of the SMART in Figure 3A, plus it allows the investigator to experimentally address the following rescue treatment question: “Among children who respond to 12 weeks of acute-phase COMB treatment, yet relapse by week 24 after stepping down to CBT-BOOST alone (by discontinuing SRT), what is the difference in longitudinal outcomes between children assigned to *re-initiate* to COMB (Treatment VII) versus those children assigned to *switch* to SRT alone (Treatment VIII)?” This question asks whether it is necessary to continue CBT or switch to medication as part of the rescue therapy. Further, the trial shown in Figure 3B has the added benefit of providing data on yet a fourth ATS, namely:

ATS 4: Step-down treatment by discontinuing medication while continuing CBT booster sessions for the first 12 weeks of the continuation phase. If the child continues to respond through week 24, then *maintain* CBT booster sessions only (no medication) for an additional 12 weeks. Otherwise, if the child relapses anytime between week 12 and week 24 after medication discontinuation, then *switch* to SRT alone (without CBT).

The design described in Figure 3B is used to illustrate and highlight the important point that subsequent randomizations in a SMART do not have to be only among responders (as in the standard discontinuation trial). SMARTs allow investigators to examine the efficacy/effectiveness of treatment options among responders and nonresponders alike.

The decision between the trial in Figure 3A versus the trial in Figure 3B will depend upon a variety of factors, including feasibility and cost. For example, in the context of our example, one important factor is the number of participants expected to relapse under the initial SRT discontinuation (CBT-BOOST alone after the acute phase). If an insufficient number of children are expected to relapse during the continuation phase, then re-randomizing nonresponders will not provide sufficient data to compare rescue treatments for nonresponders; in this case, the trial in Figure 3A is more appropriate. Subsequently, we discuss sample size considerations.

Additional Advantages of the Proposed SMARTer Designs over the Standard Discontinuation Trial

Conducting a SMART as opposed to a standard discontinuation trial has a variety of other important advantages.

Is ultimately more resource effective

It may appear that the proposed SMARTs are more costly relative to the standard discontinuation trial because of the added expense of 1) following all participants over the full course of the trial, and 2) potentially delivering multiple treatments for relapsing participants. With respect to 1), standard discontinuation trial personnel costs already account for a follow-up over the full course of the trial, as it is not possible to know ahead of time at what point during the study participants will relapse. Indeed, this is the primary outcome of the standard discontinuation trial!

With respect to 2), yes, there are additional costs incurred by providing additional treatments. However, these additional costs

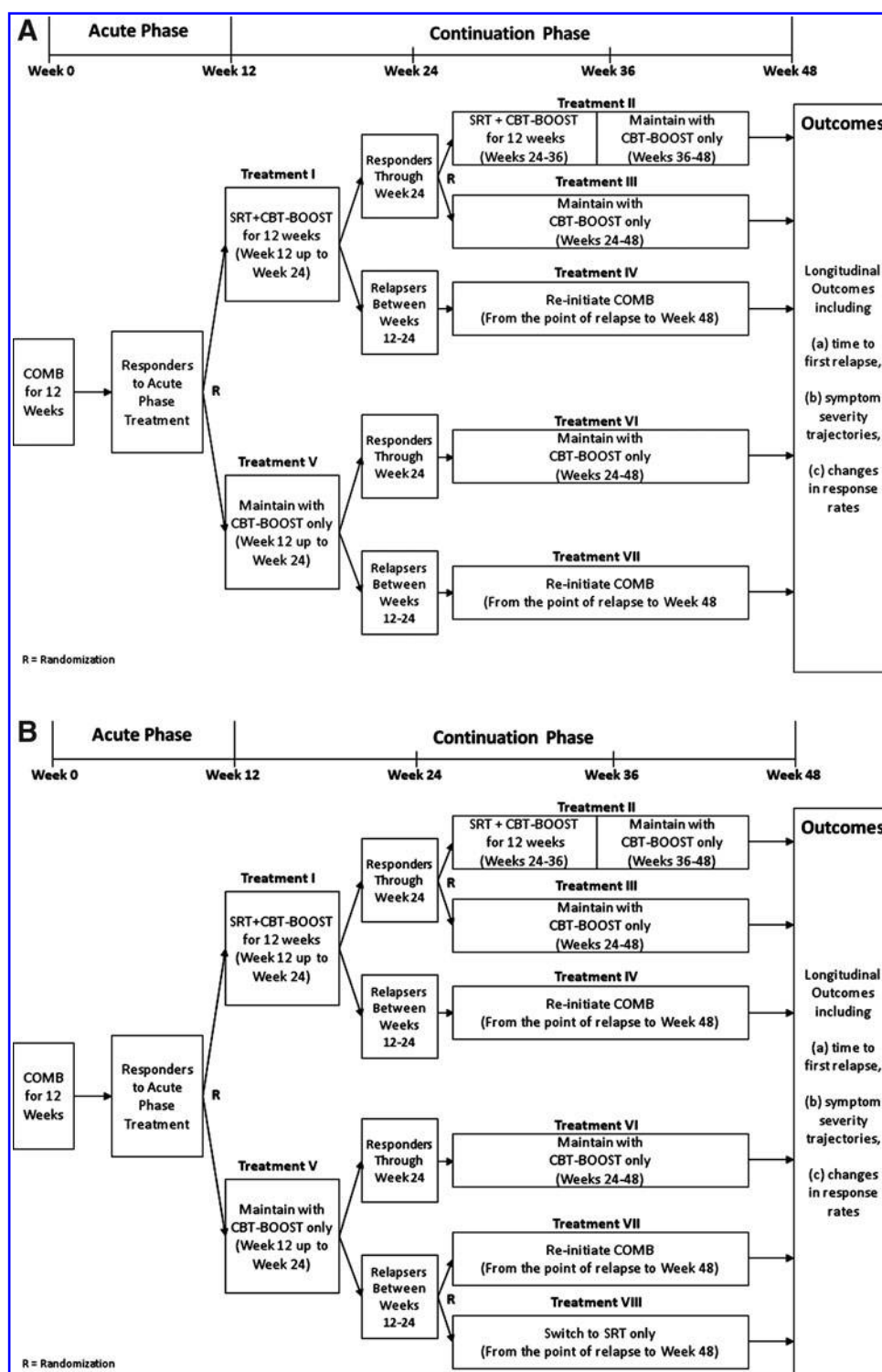


FIG. 3. (A) A SMART that builds on the standard discontinuation trial. The primary aim is a comparison of Treatment II versus Treatment III in terms of time until first relapse. This is a comparison of time until relapse from week 24 onward, among responders up to week 24. (B) A SMART that improves the standard discontinuation trial. The primary aim is a comparison of Treatment II versus Treatment III in terms of time until first relapse. This is a comparison of time until relapse from week 24 onward, among responders up to week 24. SMART, sequential multiple assignment randomized trials; SRT = sertraline medication; CBT = cognitive behavioral therapy; COMB = SRT + CBT; CBT-BOOST = CBT booster sessions.

must be placed within the context of the additional knowledge that will be gained in a trial such as the ones proposed, in particular by avoiding the design and execution of separate trials. The current approach to conducting clinical trials typically involves one trial to answer a discontinuation question and then a separate trial, often with a new sample of participants, to investigate which rescue treatment is most effective for patients who relapse. Therefore, the total cost associated with the traditional way may be significantly higher.

Permits a variety of interesting hypothesis-generating analyses

The multiple randomizations in the SMART support a variety of interesting secondary analyses. Considering Figure 3B, it is plausible that continuing SRT+CBT-BOOST in the postacute phase (Treatment I) results in lower relapse rates than for CBT-BOOST alone (Treatment V) in the short term (e.g., by week 24). However, in the context of the provided maintenance and rescue treatments, participants administered CBT-BOOST alone initially (Treatment V) may have an overall better symptom management at week 48. That is, it is possible that the movement of early relapsing participants from initial CBT-BOOST alone to COMB (full CBT + medication) enhances the durability of the skills acquired during acute-phase CBT, and thus leads to improved symptom scores at 48 weeks. On the other hand, participants who were initially assigned SRT+CBT-BOOST (Treatment I) and either did not relapse or relapsed very late in the study (i.e., did better in the shorter term), might never receive the opportunity to consolidate their CBT skills. The standard discontinuation trial does not allow investigators to investigate these delayed (or downstream) effects because it is not designed to evaluate treatments in sequence.

As a second example, suppose that some initial treatments elicit outcomes in the intermediate stages that may be used to better individualize future treatments and, in turn, produce more positive effects over the longer term. Considering Figure 3B, removing SRT at 12 weeks postacute phase may serve to diagnose participants who are “medication dependent” (i.e., show a quick rise in symptoms caused by SRT removal). This information (quicker rise versus slower rise in symptoms following SRT removal) may then be used in secondary analyses to determine the types of participants who should restart medication only, versus re-initiating COMB. The result is that the act of discontinuing medications may not have had therapeutic effect in the short term, *per se*, but helped identify individuals of a certain type (in this case, the medication-dependent type) whose future treatment could then be better individualized. SMARTs allows investigators to study these prescriptive effects, whereas the discontinuation trial does not.

Ensures balance to subsequent treatment randomizations

The standard discontinuation trial randomizes participants up front to three or more discontinuation regimes, even when some of these regimes differ only in the later weeks (for example, Partial Strategies A and B in Figure 1A are identical over the first 12 weeks postacute-phase, i.e., through week 24). One benefit of the sequential randomization is that it allows investigators to incorporate both baseline and intermediate outcome measures (up to the point of the subsequent critical decision point) to reduce chance imbalance in subsequent treatment offerings. For example, as in Figure 3A, the randomization to Treatment II or III at week 24 can be

stratified by adherence and other outcomes collected over the course of the trial through week 24.

Enhances generalizability and recruitment

Because participants understand that treatment will be adapted to their needs over the course of the full trial—akin to what happens in actual clinical practice—a more representative group of participants is likely to participate in a SMART. This enhances the heterogeneity of the sample, which improves the trial’s external validity and its ability to identify predictors (both participant characteristics and time-varying measures) of treatment response. SMART designs are likely (although we currently have only anecdotal data to support this assertion) to attract a greater variety of participants given that families are informed at the outset that their child will be treated sequentially: that is, if the first assigned treatment is not effective, then another treatment will be offered.

SMARTs Do Not Require Larger Sample Sizes

SMART designs do not necessarily require larger sample sizes than do discontinuation trials. As in all randomized clinical trials, the choice of the primary research question (aim) determines sample size. The SMART alternatives given in Figure 3 can be sized/powered to address the *very same* “time to first relapse” primary question typically used to size/power the standard discontinuation trial design (Fig. 1).

As a concrete example, suppose the primary aim of the standard discontinuation trial in Figure 1 is to examine, in terms of time to first relapse, the impact of continuing SRT+CBT-BOOST (the SRT continuation group) at week 24 versus beginning CBT-BOOST only (the SRT discontinuation group) at week 24 among children who continue to respond to 24 weeks of COMB treatment. This is a two-group comparison of Partial Strategy A versus Partial Strategy B in Figure 1A. (In addition to being of scientific interest, this is a natural primary comparison because it ensures sufficient power to detect differences between Partial Strategies A versus C and between B versus C, which usually have larger effect sizes.) In Figures 1B and 3A and B, this primary aim is a two-group randomized comparison of time until relapse (beginning at week 24, and onward) between participants who are assigned Treatment II versus those assigned Treatment III. The sample size calculation for a two-group comparison of time to relapse can be based on the log-rank test (Bland and Altman 2004), e.g., using SAS PROC POWER (SAS Institute 2008) with the TWOSAMPLESURVIVAL and test=logrank options.

Under standard assumptions, the total sample size for the SMART in Figure 3A or B is 239 children and adolescents. The number 239 is derived as follows: First, 56 responding participants are required at week 24 for a comparison of Treatment II versus Treatment III under the assumptions of a 1:1 randomization, exponentially distributed relapse distributions for both groups with an 80% continued response rate for the SRT+CBT-BOOST group and a 60% continued response rate for the CBT-BOOST only group at the end of week 36 (a 20% difference in survival at the end of week 36, or hazard ratio of 2.29), a type I error rate of 5%, and 80% power. Additional assumptions include a uniform participant accrual period of 160 weeks (3 years of recruitment), and a follow-up period of 24 weeks after the week 24 randomization, for a total of 48 weeks of study per participant as shown in the figures.

To obtain the 56 participants at week 24, 150 participants are needed to enter the continuation phase at week 12; this calculation assumes a 75% continued response rate at the end of week 24 and a

1:1 randomization at week 12 to Treatment I versus Treatment IV, that is, $2 \times 56 / 0.75 = 2 \times 75 = 150$. Next, to obtain the 150 participants at week 12, 215 participants need to enter the acute phase (COMB); this calculation assumes a 70% response rate at the end of the acute-phase period (a conservative assumption based on the 80% acute-phase response rate for COMB, as reported in CAMS), that is, $150 / 0.70 = 215$. Further inflating this sample size to account for 10% study dropout, a total of $n = 215 / 0.90 = 239$ children and adolescents would have to be recruited. This is the total sample size required for any of the three studies (Figs. 1 and 3A or B) in order to address the primary question mentioned previously, recalling that for any of the three studies considered, the data analysis for this primary question is the same.

Although the sample size for a clinical trial is always determined by the primary aim, it is often of interest to examine the resulting/implied power to test secondary hypotheses/aims. For purposes of the additional calculations given subsequently, we use the effective sample size (i.e., 215 entering the acute phase, or 150 entering the continuation phase), which already accounts for dropout. In a standard discontinuation trial such as the one shown in Figure 1A, a typical secondary question may be “Among children and adolescents who acutely respond to 12 weeks of COMB, what is the effect of discontinuing medication (CBT-BOOST only) versus continuing medication (SRT + CBT-BOOST) at week 12 on time to first relapse?” All three studies (Figs. 1 and 3A and B) are able to address this question: In Figure 1A, this is a comparison of Partial Strategies A + B versus Partial Strategy C, whereas in Figures 1B and 3A and B, this is a comparison of all participants beginning on Treatment I versus all participants beginning on Treatment IV. Based on the sample calculations provided previously, with 150 participants randomized at week 12 (entering the continuation phase), we can detect a difference of at least 12% in the survival rate (continued response) at week 24 (e.g., a 75% response in the SRT + CBT-BOOST continuation group versus $\leq 63\%$ in the CBT-BOOST only discontinuation group, or a hazard ratio of ≥ 1.62) with 80% power.

Several interesting secondary questions that cannot be addressed by the standard discontinuation trial design, but can be addressed by the SMART design are as follows: First, consider the secondary question “Among children and adolescents who acutely respond to 12 weeks of COMB, what is the effect of discontinuing medication (CBT-BOOST only) versus continuing medication (SRT + CBT-BOOST) at week 12 on the longitudinal change in anxiety symptoms between week 12 and week 48?” This question cannot be answered by the standard design because not all participants are followed for the full 48 weeks, whereas it can be answered by the two SMART designs shown in Figure 3A and B. In Figure 3A and B, this is a two-group longitudinal comparison of all participants initially assigned to Treatment I versus all participants initially assigned to Treatment V, in terms of longitudinal outcomes from week 12 onward. A common outcome measure for this type of question in child anxiety research is the continuous Pediatric Anxiety Rating Scale (PARS; RUPP Anxiety Group 2002). Based on the calculations given previously, with 150 participants randomized at week 12 (entering the continuation phase), assuming a 0.30 within-person correlation in PARS (based on the CAMS data), Type I error rate of 5%, and 80% power, we can detect a small-to-moderate effect size (Cohen 1988) of 0.44, which translates to a clinically significant difference of 2.9 units on the PARS (based on $SD = 6.6$ using the CAMS data).

Second, suppose that investigators are interested in documenting the value of rescue treatments among children and adolescents who

acutely respond to 12 weeks of COMB and who discontinue medication at week 12 (CBT-BOOST only), but subsequently relapse by week 24. This type of question cannot be addressed by the standard discontinuation trial design because these children are removed from the study. However, it can be examined (albeit in two different ways) by the two proposed SMART designs. Investigators may choose between the SMARTer Figure 3A design versus the SMARTer Figure 3B design. Using the former design (Fig. 3A), investigators can estimate the fraction of relapsing children who are rescued by re-initiation of COMB. Using the latter design (Fig. 3B), investigators can make a randomized comparison of the two rescue treatments (VII: re-initiate COMB or VIII: re-initiate medication only) among children who had discontinued medication at week 12 and relapsed by week 24. The data analysis here would be a between-groups (VII vs. VIII) comparison of change in PARS from week 24 to week 48. To decide between the design in Figure 3A and the design in Figure 3B, consideration of the anticipated relapse rate by week 24 the minimum detectable effect size between the two rescue treatments evaluated in the Figure 3B design is useful. In general, the higher the rate of relapse and larger the minimum detectable effect size, the easier it is to justify using the Figure 3B design over the Figure 3A design. Based on the calculations given previously, $150/2 = 75$ participants will be randomized initially to CBT-BOOST only (Treatment V in Figs. 3A and B). As mentioned, assuming a 0.30 within-person correlation in PARS (based on the CAMS data), type I error rate of 5%, 80% power, and relapse rates ranging from 40% (15 randomized to each rescue treatment in Fig. 3B) to 70% (26 randomized to each rescue treatment), the minimum detectable effect size ranges from 0.52 (moderate) to 0.70 (large). These effect sizes translate into between rescue treatment group changes of 3.47–4.63 units on the PARS. Typically, a change of at least 4 units on the PARS is considered clinically meaningful. Therefore, if relapse rates of $\geq 60\%$ are expected by week 24 on initial medication discontinuation (Treatment V), then the design in Figure 3B may be justifiable; otherwise, the design in Figure 3A is preferred.

Third, a further secondary aim might be to select the best among the ATSS embedded in the SMART (Oetting et al. 2007). As discussed, the design in Figure 3A has three ATSSs; whereas the design in Figure 3B has four ATSSs. Based on the results in Oetting et al. (2007), a sample size of 150 is sufficient for choosing the best ATS with high probability ($\sim 80\%$) when the best two ATSSs differ by moderate effect sizes or larger. Again, this is an example of a secondary question that cannot be examined in the standard discontinuation trial, but can be addressed by the proposed SMARTer alternatives.

Discussion

Other variations of the SMART design besides the two discussed here are possible. In this article, we have focused on these two choices because they represent two simple adaptations of the standard discontinuation trial, yet provide additional hypothesis-generating analyses concerning the individualization and sequencing of treatments, with relatively little additional logistical or cost considerations relative to the amount of scientific knowledge to be gained.

After examining the primary aims of a SMART, investigators may use extensions of modern regression methods such as Q-Learning (Nahum-Shani et al. to appear), to develop more richly tailored ATSSs. The resulting ATSSs exhibit greater individualization than do the ATSSs embedded in the SMART. For example,

investigators may learn in subsequent data analyses that adherence during the postacute treatment continuation phase is important for deciding both whether or not to discontinue treatment after week 24 and which rescue treatment to provide if the participant relapses.

SMARTs constitute an important part of the developmental, or phased, approach to experimentation (Collins et al. 2005). This means that after examining the primary and secondary scientific questions in a SMART, and after developing a more richly tailored ATS using data arising from a SMART, a “fully optimized ATS” can be compared against usual care, standard care, or other promising treatments in a follow-up two-group (or more) randomized clinical trial.

Conclusions

Discontinuation trials have been used in psychiatry to answer important clinical questions, such as how long to provide a given treatment. In this article, we suggest that the implicit motivation for executing a discontinuation trial is to inform individualized sequences of treatment decisions, that is, ATSS, for use in the management of mental health disorders. A weakness of the standard discontinuation trial is that it does not allow investigators to develop, examine, or compare sequences of treatment decisions over the course of the postacute phase of treatment. In essence, by removing nonresponders from further study, the standard discontinuation trial fails to yield data that can be used to shed light on how to treat this population. We showed how the standard discontinuation trial can be modified and significantly improved to yield data that can better construct ATSS. In particular, we showed how the primary questions addressed in a standard discontinuation trial can also be addressed by embedding the standard discontinuation trial within a SMART.

The innovative SMART design is ultimately more efficient and less costly than the standard discontinuation trial, and it is more attractive scientifically because of the additional scientific questions that it allows the investigator to address.

Clinical Significance

SMARTs are broadly applicable in any settings in which investigators are interested in developing an ATS for optimizing and individualizing (i.e., personalizing) treatment over time. This article will be of interest to clinical investigators who are interested in deriving more information out of a standard discontinuation trial. In this context, SMARTs can be used to address clinically significant questions such as “When to discontinue treatment following acute-phase response?” and “Which treatment to provide given a relapse following treatment discontinuation?”

Acknowledgments

We are grateful to Joel T. Sherrill for insightful conversations and guidance concerning this work.

Disclosures

Drs. Almirall and Murphy have no disclosures. Dr. Compton receives consultancy fees from Shire. He has received research or grant support from the Tourette Syndrome Association. Dr. Rynn receives or has received grants, research support and/or consultancy fees from Eli Lilly, Pfizer, Merck, and Shire. She receives royalties from American Psychiatric Publishing. Dr. Walkup receives or has received grant, travel, honoraria, and/or research support from, and serves on the advisory board for, the Tourette

Syndrome Association. He received free drugs and placebos from Eli Lilly, Pfizer, and Abbott for National Institutes of Health (NIH)-funded studies. He receives consultancy fees from Shire, and he receives royalties from Guilford Press and Oxford Press.

References

- Bland JM, Altman DG: The logrank test. *BMJ* 328:1073, 2004.
- Cohen J: *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: L. Erlbaum Associates; 1988.
- Collins LM, Murphy SA, Nair VN, Strehler VJ: A strategy for optimizing and evaluating behavioral interventions. *Ann Behav Med* 30:65–73, 2005.
- Emslie GJ, Kennard BD, Mayes TL, Nightingale–Teresi J, Carmody T, Hughes CW, Rush AJ, Tao RR, Rintelmann JW: Fluoxetine versus placebo in preventing relapse of major depression in children and adolescents. *Am J Psychiatry* 165:459–467, 2008.
- Homberg JR, Schubert D, Gaspar P: New perspectives on the neurodevelopmental effects of SSRIs. *Trends Pharmacol Sci* 31:60–65, 2010.
- Lavori PW, Dawson R: Adaptive treatment strategies in chronic disease. *Annu Rev Med* 59:443–453, 2008.
- Lavori PW, Dawson R: A design for testing clinical strategies: Biased adaptive within-subject randomization. *J Roy Statist Soc Ser A* 163:29–38, 2000.
- Lavori PW, Dawson R: Improving the efficiency of estimation in randomized trials of adaptive treatment strategies. *Clin Trials* 4: 297–308, 2007.
- Lavori PW, Dawson R, Rush AJ: Flexible treatment strategies in chronic disease: Clinical and research implications. *Biol Psychiatry* 48:605–614, 2000.
- Lei H, Nahum–Shani I, Lynch K, Oslin D, Murphy SA: A SMART design for building individualized treatment sequences. *Ann Rev Clin Psychol* 8:21–48, 2011.
- March JS, Frances A, Kahn DA, Carpenter D: Expert consensus guidelines: Treatment of obsessive-compulsive disorder. *J Clin Psychiatry* 1997:2–72, 1997.
- Murphy SA: An experimental design for the development of adaptive treatment strategies. *Stat Med* 24:1455–1481, 2005.
- Murphy SA, Almirall D: Dynamic treatment regimens. In: *Encyclopedia of Medical Decision Making*. Edited by Kattan M. Thousand Oaks, CA: Sage Publications; 2009; pp. 419–422.
- Murphy SA, Collins LM, Rush AJ: Customizing treatment to the patient: Adaptive treatment strategies. *Drug Alcohol Depend* 88: S1–S3, 2007a.
- Murphy SA, Lynch KG, Oslin D, McKay JR, TenHave T: Developing adaptive treatment strategies in substance abuse. *Drug Alcohol Depend* 88:S24–S30, 2007b.
- Murphy SA, Oslin DW, Rush AJ, Zhu J, MCATS: Methodological challenges in constructing effective treatment sequences for chronic psychiatric disorders. *Neuropsychopharmacology* 32:257–262, 2007c.
- Nahum–Shani I, Qian M, Almirall D, Pelham WE, Gnagy B, Fabiano G, Waxmonsky J, Murphy SA: Experimental design and primary data analysis methods for comparing adaptive interventions. *Psychol Methods*, 2012. Epub ahead of print.
- Nahum–Shani I, Qian M, Almirall D, Pelham WE, Gnagy B, Fabiano G, Waxmonsky J and Murphy SA: Q-Learning: A data analysis method for constructing adaptive interventions. *Psychol Methods*, to appear.
- Oetting A, Levy J, Weiss R, Murphy SA: Statistical methodology for a SMART design in the development of adaptive treatment strategies. In: *Causality and Psychopathology: Finding the Determinants of Disorders and their Cures*. Edited by Shrout PE. Arlington, VA: American Psychiatric Publishing, Inc.; 2007; pp. 179–205.

- Pine DS: Treating children and adolescents with selective serotonin reuptake inhibitors: How long is appropriate? *J Child Adolesc Psychopharmacol* 12:189–203, 2002.
- RUPP Anxiety Study Group: The Pediatric Anxiety Rating Scale (PARS): Development and psychometric properties. *J Am Acad Child Adolesc Psychiatry* 41:1061–1069, 2002.
- SAS Institute I: SAS/STAT SAS User's Guide. Cary, NC: SAS Institute; 2008.
- Shireman TI, Olson BM, Dewan NA: Patterns of antidepressant use among children and adolescents. *Psychiatr Serv* 53:1444–1450, 2002.
- Thall PF, Logothetis C, Pagliaro LC, Wen SJ, Brown MA, Williams D, Millikan RE: Adaptive therapy for androgen-independent prostate cancer: A randomized selection trial of four regimens. *J Natl Cancer Inst* 99:1613–1622, 2007.
- Thall PF, Millikan RE, Sung HG: Evaluating multiple treatment courses in clinical trials. *Stat Med* 19:1011–1028, 2000.
- Walkup JT, Albano AM, Piacentini J, Birmaher B, Compton SN, Sherrill JT, Ginsburg GS, Rynn MA, McCracken J, Waslick B, Iyengar S, March JS, Kendall PC: Cognitive behavioral therapy, sertraline, or a combination in childhood anxiety. *N Engl J Med* 359:2753–2766, 2008.
- Wang L, Rotnitzky A, Lin X, Millikan RE, Thall PF: Evaluation of viable dynamic treatment regimes in a sequentially randomized trial of advanced prostate cancer. *J Amer Statist Assoc*, in press.
- Wittchen HU, Kessler RC, Beesdo K, Krause P, Hofler M, Hoyer J: Generalized anxiety and depression in primary care: Prevalence, recognition, and management. *J Clin Psychiatry* 63:24–34, 2002.
- Young BJ, Beidel DC, Turner SM, Ammerman RT, McGraw K, Coaston SC: Pretreatment attrition and childhood social phobia: Parental concerns about medication. *J Anxiety Disord* 20:1133–1147, 2006.
- Young BJ, Beidel DC, Turner SM, McGraw K, Ammerman RT, Coaston S: Pretreatment attrition in a randomized clinical trial: The role of parental fears and attitudes toward medication. *J Child Adolesc Psychopharmacol* 15:855–856, 2005.

Address correspondence to:

Daniel Almirall, Ph.D.

2204 Institute for Social Research

426 Thompson Street

University of Michigan

Ann Arbor, MI 48104-2321

E-mail: dalmiral@umich.edu

This article has been cited by:

1. Neal D. Ryan, Jae-Won Kim, Bennett L. Leventhal. 2012. What To Do When It Is not so Clear What To Do: New Opportunities for Studies. *Journal of Child and Adolescent Psychopharmacology* 22:6, 463-465. [[Citation](#)] [[Full Text HTML](#)] [[Full Text PDF](#)] [[Full Text PDF with Links](#)]