



A “SMART” Design for Building Individualized Treatment Sequences

H. Lei,¹ I. Nahum-Shani,² K. Lynch,³ D. Oslin,⁴
and S.A. Murphy⁵

¹Department of Statistics, University of Michigan, Ann Arbor, Michigan 48109;
email: ehlei@umich.edu

²Institute for Social Research, University of Michigan, Ann Arbor, Michigan 48106;
email: inbal@isr.umich.edu

³Treatment Research Center and Center for Studies of Addictions, Department of
Psychiatry, University of Pennsylvania, Philadelphia, Pennsylvania 19104;
email: lynch_k@mail.trc.upenn.edu

⁴Philadelphia Veterans Administration Medical Center, Philadelphia, Pennsylvania 19104,
and Treatment Research Center and Center for Studies of Addictions, Department of
Psychiatry, University of Pennsylvania, Philadelphia, Pennsylvania 19104;
email: oslin@upenn.edu

⁵Department of Statistics, Institute for Social Research, and Department of Psychiatry,
University of Michigan, Ann Arbor, Michigan 48109; email: samurphy@umich.edu

Annu. Rev. Clin. Psychol. 2012. 8:14.1–14.28

The *Annual Review of Clinical Psychology* is online
at clinpsy.annualreviews.org

This article's doi:
10.1146/annurev-clinpsy-032511-143152

Copyright © 2012 by Annual Reviews.
All rights reserved

1548-5943/12/0427-0001\$20.00

Keywords

adaptive intervention, Sequential Multiple Assignment Randomized
Trial, individualized treatment sequences

Abstract

Interventions often involve a sequence of decisions. For example, clinicians frequently adapt the intervention to an individual's outcomes. Altering the intensity and type of intervention over time is crucial for many reasons, such as to obtain improvement if the individual is not responding or to reduce costs and burden when intensive treatment is no longer necessary. Adaptive interventions utilize individual variables (severity, preferences) to adapt the intervention and then dynamically utilize individual outcomes (response to treatment, adherence) to readapt the intervention. The Sequential Multiple Assignment Randomized Trial (SMART) provides high-quality data that can be used to construct adaptive interventions. We review the SMART and highlight its advantages in constructing and revising adaptive interventions as compared to alternative experimental designs. Selected examples of SMART studies are described and compared. A data analysis method is provided and illustrated using data from the Extending Treatment Effectiveness of Naltrexone SMART study.

Contents

OVERVIEW	14.2
ADAPTIVE INTERVENTIONS	14.3
What Are Adaptive Interventions? ..	14.3
Why Consider Adaptive Interventions?	14.4
Example: The Adaptive Drug Court Program	14.4
THE SEQUENTIAL MULTIPLE ASSIGNMENT RANDOMIZED TRIAL	14.5
What Is the SMART Design?	14.5
An Example of the SMART	14.6
Why Run a SMART Trial?	14.7
Comparison to Alternative Experimental Designs	14.8
The SMART Experimental Design Versus the Adaptive Experimental Design	14.10
EXAMPLES OF SMART STUDIES	14.11
Adaptive CCNIA Developmental and Augmented Intervention ...	14.11
Adaptive Pharmacological and Behavioral Treatments for Children with ADHD Trial	14.13
Adaptive Reinforcement-Based Treatment for Pregnant Drug Abusers	14.15
Extending Treatment Effectiveness of Naltrexone	14.16
Comparison of SMART Studies ...	14.18
USING DATA FROM A SMART TO INFORM THE CONSTRUCTION OF AN OPTIMIZED ADAPTIVE INTERVENTION	14.20
THE EXTEND TRIAL EXAMPLE	14.20
Sample	14.20
Measures	14.21
Illustrative Data Analyses	14.22
Results	14.22
SUMMARY	14.24

OVERVIEW

Adaptive interventions are interventions in which the type or the dosage of the intervention offered to patients is individualized on the basis of patients' characteristics or clinical presentation and then are repeatedly adjusted over time in response to their ongoing performance (see, for example, Bierman et al. 2006, Marlowe et al. 2008, McKay 2005). This approach is based on the notion that patients differ in their responses to interventions: In order for an intervention to be most effective, it should be individualized and repeatedly adapted over time to individual progress. An adaptive intervention is a multi-stage process that can be operationalized via a sequence of decision rules that recommend when and how the intervention should be modified in order to maximize long-term primary outcomes. These recommendations are based not only on patients' characteristics but also on intermediate outcomes collected during the intervention, such as the patient's response and adherence. Adaptive interventions are also known as dynamic treatment regimes (Murphy et al. 2001, Robins 1986), adaptive treatment strategies (Lavori & Dawson 2000, Murphy 2005), multi-stage treatment strategies (Thall et al. 2002, Thall & Wathen 2005) and treatment policies (Lunceford et al. 2002; Wahed & Tsiatis 2004, 2006).

The conceptual advantages of adaptive interventions over a fixed-intervention approach (in which all patients are offered the same type or dosage of the intervention) have long been recognized by behavioral scientists. For example, Weisz et al. (2004, pp. 302–303) presented a discussion of dissemination and evidence-based practice in clinical psychology, suggesting that evidence-based practice should ideally consist of much more than simply obtaining an initial diagnosis and choosing a matching treatment:

Evidence-based practice... is not a specific treatment or a set of treatments, but rather an orientation or a value system that relies on evidence to guide the entire treatment process. Thus, a critical element of evidence-based care is periodic assessment to gauge whether the

treatment selected initially is in fact proving helpful. If it is not, adjustments in procedures will be necessary, perhaps several times over the course of the treatment.

In recent years, intervention scientists have become increasingly interested in experimental designs that explicitly inform the adaptive aspects of the intervention. The general aim is to obtain data useful for addressing questions concerning the individualization and adaptation of intervention options so as to construct high-quality (i.e., highly efficacious) adaptive interventions. Here, we discuss an experimental design useful for developing (i.e., for constructing or revising) adaptive interventions in the area of **behavioral health**. First we review how decision rules can be used to operationalize the individualization of intervention options (i.e., intervention type and dosage) on the basis of **patients' characteristics** and the repeated adaptation of intervention options over time in response to the **ongoing performance of the patient**. Next we discuss how the **Sequential Multiple Assignment Randomized Trial (SMART)** (Murphy 2005; Murphy et al. 2007a,b; for precursors and related ideas, see also Lavori & Dawson 2000, 2004) can be used to develop adaptive interventions. More specifically, we discuss **how data from SMART designs are useful in addressing key research questions that inform the construction of a high-quality adaptive intervention**. We discuss the advantages of SMART relative to other experimental approaches and describe four selected examples of SMART studies that are ongoing or completed. Finally, we provide a data analysis method for addressing several interesting research questions concerning adaptive interventions. We use the Extending Treatment Effectiveness of Naltrexone (ExTEND) SMART study for illustration.

ADAPTIVE INTERVENTIONS

What Are Adaptive Interventions?

An adaptive intervention consists of **four key elements**. The first element is a **sequence of**

critical decisions in a patient's care. Critical decisions might concern which intervention to provide first and, if the initial intervention is unsuccessful, which intervention to provide second. In many settings, the risk of relapse or exacerbations is high; thus, critical decisions must be made even after an acute response has been achieved. These decisions may concern which maintenance intervention should be offered and whether and how signs of impending relapse should be monitored (McKay 2009). The second element is **a set of possible intervention options at each critical decision point**. Possible intervention options include different types of behavioral and pharmacological interventions, different modes of delivery, different combinations of interventions, different approaches to enhance engagement and adherence to the intervention, and different intervention timelines. The third element is **a set of tailoring variables that is used to pinpoint when the intervention should be altered and to identify which intervention option is best for whom**. These variables usually include information that is useful in detecting early signs that the intervention is insufficiently effective (e.g., early signs of nonresponse to intervention, adherence, side effects, and burden), but it can also include contextual information (e.g., individual, family, social, and environmental characteristics) as well as information concerning the intervention options already received. The logic is that the best intervention option for patients varies according to different values of the tailoring variables. The fourth ingredient is **a sequence of decision rules, one rule per critical decision**. The decision rule links individuals' characteristics and ongoing performance with specific intervention options. The aim of these decision rules is to guide practitioners in deciding which intervention options to use at each stage of the adaptive intervention based on available information relating to the characteristics and/or ongoing performance of the patient. Each decision rule inputs the tailoring variables and outputs one or more recommended intervention options (Collins et al. 2004;

Lavori & Dawson 2000, 2004; Lavori et al. 2000).

Why Consider Adaptive Interventions?

There are many reasons for considering adaptive interventions for the prevention and treatment of mental health disorders. First, patients may vary in their response to treatment. Some patients may respond whereas others may not respond to the same intervention type, intensity, or duration. For example, in a clinical trial of four psychosocial interventions for cocaine dependence (Crits-Christoph et al. 1999, Stulz et al. 2010), 35% to 60% of patients were continuing to use cocaine at the end of the six-month intervention. This motivates the provision of subsequent intervention options for patients who do not respond adequately to the initial intervention. These options often include switching to an alternative intervention or increasing the dose/intensity of the intervention. Second, the effectiveness of an intervention may change over time due to dynamically evolving risk and resiliency or waxing and waning of the disorder (Angst et al. 2009, Hser et al. 2007). Hence, it is important to decide if and when the intervention is no longer working and, accordingly, which intervention should follow. A third reason for considering an adaptive intervention is the presence of, or evolving, comorbidities (e.g., comorbid HIV infection and substance use, depression and alcoholism). Comorbid conditions require decisions regarding which disorder should be treated first or whether multiple disorders should be treated simultaneously. A fourth reason is that relapse may be common. For example, in the treatment of substance abuse, patients are at high risk for relapse even after achieving an acute response (McKay 2009, McKay et al. 2010, McLellan 2002, McLellan et al. 2000). Fifth, the high cost of intensive interventions combined with possible burden and/or side effects motivate the development of interventions in which the intensity of the treatment is reduced when

possible. Finally, difficulties in maintaining adherence to interventions are another important reason to consider adaptive interventions. For example, adherence to treatment with antipsychotic medications by patients with schizophrenia ranges from 45% to 63% (Liu-Seifert et al. 2011). Low adherence motivates the development of interventions in which the type or intensity of the treatment is tailored to encourage adherence. To clarify these ideas, consider the following abbreviated description of the Adaptive Drug Court Program implemented by Marlowe et al. (2008).

Example: The Adaptive Drug Court Program

The Adaptive Drug Court Program was implemented by Marlowe and colleagues (2008) for individuals charged with a drug-related misdemeanor offense. At entry into the program, offenders were classified as high risk for failure if they met the criteria in the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*, for antisocial personality disorder or had previously attended at least one formal drug abuse intervention, excluding self-help groups; otherwise offenders were classified as low risk. Offenders who were classified as high risk were assigned to biweekly court hearings, whereas offenders classified as low risk were assigned to as-needed court hearings. In addition to the court hearings, all offenders were required to attend weekly group substance abuse counseling sessions and to provide weekly urine specimens.

Each offender's progress in the program was assessed monthly. If at any monthly assessment an offender had missed two or more counseling sessions without an excuse or failed to provide two or more scheduled urine specimens, he/she was classified as noncompliant, and the level of court supervision was increased. In particular, offenders who were assigned to as-needed court hearings but did not comply moved to biweekly court hearings; offenders who were assigned to biweekly court hearings but did not comply were placed on a jeopardy contract in which further violation of the rules of the program

Table 1 Decision rules in the adaptive drug court program in if-then-else language

Decision rules in the adaptive drug court program
IF baseline risk assessment = {low risk}
THEN first-stage intervention = {as-needed court hearings + standard counseling}
IF monthly assessment = {noncompliant}
THEN second-stage intervention = {biweekly court hearings + standard counseling}
ELSE IF monthly assessment = {nonresponsive}
THEN second-stage intervention = {as-needed court hearings + ICM}
ELSE IF monthly assessment = {compliant and responsive}
THEN second-stage intervention = {as-needed court hearings + standard counseling}
ELSE IF baseline risk assessment = {high risk}
THEN first-stage intervention = {biweekly court hearings + standard counseling}
IF monthly assessment = {noncompliant}
THEN second-stage intervention = {jeopardy contract}
ELSE IF monthly assessment = {nonresponsive}
THEN second-stage intervention = {biweekly court hearings + ICM}
ELSE IF monthly assessment = {compliant and responsive}
THEN second-stage intervention = {biweekly court hearings + standard counseling}

Abbreviation: ICM, intensive case management.

resulted in moving into a regular court system. If an offender attended the scheduled counseling sessions, provided the urine specimens and did not commit new infractions, but two or more urine specimens were drug-positive, then the offender was classified as nonresponsive. In this case, the intensity and scope of the drug abuse treatment was altered. More specifically, these offenders entered an intensive case management program in which they were provided twice-weekly individual substance abuse counseling sessions. Offenders who were both compliant and responsive continued on the initially assigned program. **Table 1** and **Figure 1** describe the decision rules that operationalize the Adaptive Drug Court Program.

In this adaptive intervention, two critical decisions must be made regarding (a) the initial intensity of intervention (court hearings) and (b) how to alter the intervention in response to the ongoing performance of the offender. When making the first critical decision, available intervention options include as-needed court hearings and biweekly court hearings. The tailoring variable for the initial intervention is the risk level of the offender (high risk or low risk). The subsequent intervention

options are referral to an intensive clinical case management program, instituting a jeopardy contract, instituting a more intensive court hearing schedule, or continuing the previous court hearing schedule. The tailoring variable for the subsequent intervention is how well offenders were progressing in the program, in particular whether the offender is noncompliant, nonresponsive but compliant, or both compliant and responsive.

THE SEQUENTIAL MULTIPLE ASSIGNMENT RANDOMIZED TRIAL

What Is the SMART Design?

Sequential Multiple Assignment Randomized Trials (Dawson & Lavori 2011, Lavori & Dawson 2000, Murphy 2005) are used to inform the development of adaptive interventions. A SMART involves multiple intervention stages; each stage corresponds to one of the critical decisions involved in the adaptive intervention. Each participant moves through the multiple stages, and at each stage the participant is randomly (re)assigned

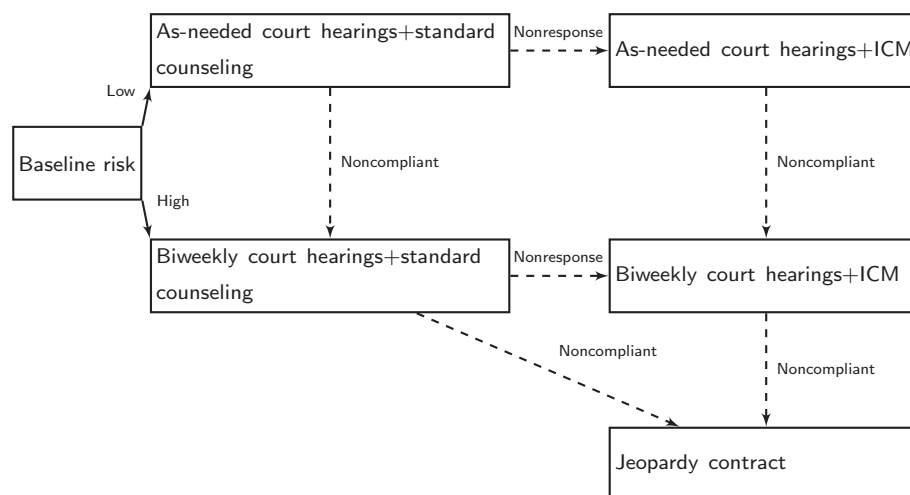


Figure 1

Flow chart of the Adaptive Drug Court Program. ICM, intensive case management.

to one of several intervention options. As in standard intervention trials, the randomization allows scientists to make valid causal inferences concerning the relative effectiveness of the intervention options without having to make unverifiable assumptions (Rubin 1974, 2007). Because SMARTs are used to develop adaptive interventions as opposed to confirming that a particular adaptive intervention is better than control, SMARTs should be followed by a randomized confirmatory trial in which the developed adaptive intervention is compared to an appropriate alternative. It is worth noting that although SMARTs involve randomization, once an adaptive intervention has been constructed, the delivery of the adaptive intervention does not involve randomization.

Trials in which participants are randomized multiple times have been implemented in a variety of fields, likely beginning in cancer research (Joss et al. 1994, Stone et al. 1995, Tummarello et al. 1997). Other precursors of SMART include the CATIE trial of treatments for patients with Alzheimer's disease (Schneider et al. 2003) and STAR*D for treatment of depression (Rush et al. 2004, 2006). Recently concluded SMART trials include the CATIE trial of treatments for schizophrenia (Stroup et al. 2003),

Phase II trials at University of Texas MD Anderson Cancer Center for treating cancer (Thall et al. 2007a), the ExTEND trial of treatments for alcohol dependence by Oslin (2005) at the University of Pennsylvania, and a SMART study of treatments for children with attention deficit/hyperactivity disorder (ADHD) by W. Pelham at the University at Buffalo, SUNY (personal communication). Trials currently in the field include a SMART for nonverbal children with autism by Kasari (2009a,b) at UCLA, a SMART concerning behavioral interventions for drug-addicted pregnant women by Jones (2010) at the Johns Hopkins University, and a SMART for the treatment of metastatic renal cell carcinoma by Tannir (2010) at the University of Texas MD Anderson Cancer Center.

An Example of the SMART

To make the discussion more concrete, consider the SMART concerning behavioral interventions for drug-addicted pregnant women. This trial is designed to provide data regarding how the intensity and scope of reinforcement-based treatment (RBT) might be adapted to a pregnant woman's progress in treatment.

There are four types of RBT: abbreviated RBT (aRBT), reduced RBT (rRBT), treatment-as-usual RBT (tRBT), and enhanced RBT (eRBT), listed in order of increasing intensity and scope. In this trial, each participant is randomized to one of the two possible initial interventions (tRBT or rRBT). After two weeks, participants are classified as either nonresponders or responders. A participant is classified as a nonresponder if she misses an intervention day with no excuse, provides a positive opioid or cocaine urine specimen, or self-reports using either drug; otherwise the participant is classified as a responder. Then each nonresponder is rerandomized to one of two subsequent interventions that are equally as intensive as or are more intensive (in terms of dose and scope) than the initial intervention. Specifically, nonresponders to tRBT are rerandomized to either tRBT or eRBT, and nonresponders to rRBT are rerandomized to either rRBT or tRBT. Each responder is randomized to one of two subsequent interventions that are equally as intensive as or are less intensive (in terms of dose and scope) than the initial intervention that was offered to the participant. Responders to rRBT are rerandomized to either rRBT or aRBT, and responders to tRBT are rerandomized to either rRBT or tRBT. The SMART study design is illustrated in **Figure 2**. Eight adaptive interventions are embedded in this design (**Table 2**). In the examples section we provide more details concerning this SMART study.

Why Run a SMART Trial?

As stated above, the goal of the SMART design is to inform the development of adaptive interventions. Data from a SMART design can be used to address many interesting questions that are useful for the construction of high-quality adaptive intervention. These questions may concern the comparison of different intervention options at different stages of the intervention. For example, comparisons may be made of the difference between the first-stage intervention options (rRBT and tRBT) in the context of the specified second-stage intervention

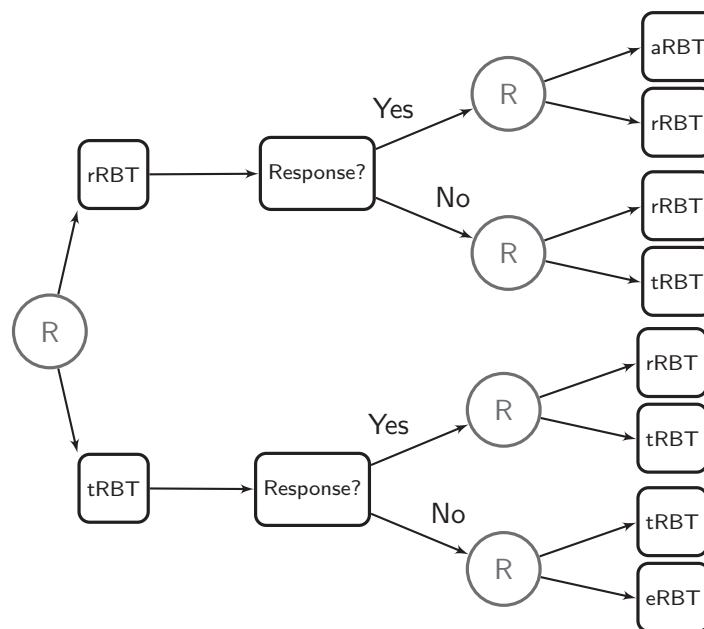


Figure 2

Design of the adaptive Reinforcement-Based Treatment (RBT) Trial (H. Jones, P.I.). aRBT, abbreviated RBT; eRBT, enhanced RBT; rRBT, reduced RBT; tRBT, treatment-as-usual RBT.

options or of the difference between the second-stage intervention options for nonresponding or responding participants. Other questions may concern the comparison of adaptive interventions that are embedded in the SMART trial, such as the differences between the eight adaptive interventions specified in **Table 2**.

Data from a SMART design can also be used to address questions that concern the construction of more deeply tailored adaptive interventions that go beyond those explicitly embedded as part of the SMART design. Specifically, investigators can collect information on candidate tailoring variables that are outcomes of the initial intervention and assess their usefulness in matching subsequent intervention options to participants. For example, in the secondary analysis of data from the RBT SMART study described above, one interesting question might be, “Does the effect of the second-stage intervention options (i.e., the difference between rRBT and tRBT) for nonresponders to rRBT vary depending on the participant’s

Table 2 Embedded adaptive interventions in the Reinforcement-Based Treatment (RBT) Trial

Adaptive intervention	Decision rule
First offer rRBT, then switch to aRBT for responders and continue rRBT for nonresponders	First-stage intervention option = {rRBT} IF two-week evaluation = {response} THEN second-stage intervention option = {aRBT} ELSE second-stage intervention option = {rRBT}
First offer rRBT, then continue rRBT for both responders and nonresponders	First-stage intervention option = {rRBT} IF two-week evaluation = {response} THEN second-stage intervention option = {rRBT} ELSE second-stage intervention option = {rRBT}
First offer rRBT, then switch to aRBT for responders and switch to tRBT for nonresponders	First-stage intervention option = {rRBT} IF two-week evaluation = {response} THEN second-stage intervention option = {aRBT} ELSE second-stage intervention option = {tRBT}
First offer rRBT, then continue rRBT for responders and switch to tRBT for nonresponders	First-stage intervention option = {rRBT} IF two-week evaluation = {response} THEN second-stage intervention option = {rRBT} ELSE second-stage intervention option = {tRBT}
First offer tRBT, then switch to rRBT for responders and continue tRBT for nonresponders	First-stage intervention option = {tRBT} IF two-week evaluation = {response} THEN second-stage intervention option = {rRBT} ELSE second-stage intervention option = {tRBT}
First offer tRBT, then continue tRBT for both responders and nonresponders	First-stage intervention option = {tRBT} IF two-week evaluation = {response} THEN second-stage intervention option = {tRBT} ELSE second-stage intervention option = {tRBT}
First offer tRBT, then switch to rRBT for responders and switch to eRBT for nonresponders	First-stage intervention option = {tRBT} IF two-week evaluation = {response} THEN second-stage intervention option = {rRBT} ELSE second-stage intervention option = {eRBT}
First offer tRBT, then continue tRBT for responders and switch to eRBT for nonresponders	First-stage intervention option = {tRBT} IF two-week evaluation = {response} THEN second-stage intervention option = {tRBT} ELSE second-stage intervention option = {eRBT}

Abbreviations: aRBT, abbreviated RBT; eRBT, enhanced RBT; rRBT, reduced RBT; tRBT, treatment-as-usual RBT.

RBT skill comprehension measured during the first-stage of the intervention?” Addressing this question may improve the primary outcome through **greater individualization**.

Comparison to Alternative Experimental Designs

Recall that **the SMART experimental design was developed to aid in the construction and refinement of adaptive interventions**. With this goal in mind, at least two other plausible

experimental alternatives to the SMART are possible: (a) the use of multiple, one-stage-at-a-time, randomized trials and (b) a randomized trial in which a fully formed adaptive intervention is compared to an appropriate alternative. Advantages and disadvantages of the SMART design relative to these approaches are discussed below.

Consider the first plausible alternative to the SMART design, that of employing multiple, one-stage-at-a-time, randomized trials. This approach focuses separately on each

intervention stage composing the sequence as opposed to viewing the intervention sequence as a whole. For clarity, suppose the critical decisions are (a) which intervention to provide initially, (b) which intervention to provide to responders, and (c) which intervention to provide to nonresponders. The one-stage-at-a-time approach would utilize three randomized trials, one per critical decision. This is a very attractive option because often the initial interventions have been evaluated in prior studies. The intervention that performed best in these prior studies is selected as the first-stage intervention. Thus, a natural next step is to conduct two studies: a study of intervention responders and a study of intervention nonresponders. The adaptive intervention obtained this way consists of the selected first-stage intervention followed by offering the intervention that was found to perform best in the study involving responders, and offering the intervention option that was found to perform best in the study of nonresponders.

The SMART design has at least three advantages as compared to multiple, one-stage-at-a-time, randomized trials in developing adaptive interventions. These advantages are (a) increased validity of analyses aimed at discovering when the effect of one intervention is enhanced by subsequent or prior interventions, (b) increased validity of analyses aimed at discovering useful tailoring variables, and (c) increased ability to reduce the impact of cohort effects. All three advantages enhance the ability of scientists to construct high-quality adaptive interventions. In the multiple, one-stage-at-a-time, randomized trials approach, scientists can attempt to detect synergistic effects by collecting retrospective information about the type and intensity of the interventions participants received prior to nonresponse (or prior to response). In contrast, in a SMART study the clinical team can prospectively manipulate and record the type and intensity of the intervention during the initial stage.

The second advantage of the SMART design as compared to multiple, one-stage-at-a-time, randomized trials is that the SMART

design enhances the ability of researchers to assess whether outcomes to the initial intervention might be useful in individualizing the subsequent intervention. For example, in secondary data analyses of the ADHD study, Nahum-Shani et al. (2011a,b) found evidence that children's adherence to the first-stage intervention could be used to individualize the second-stage intervention. Among children who do not respond to the first-stage intervention (either low-dose medication or low-dose behavioral modification), those with low adherence performed better when the first-stage intervention was augmented with the other type of intervention, relative to increasing the dose or intensity of the first-stage intervention. One might try to obtain information concerning adherence to the first-stage intervention in a nonresponders study by asking participants to retrospectively report their level of adherence to the intervention provided to them before they entered the nonresponders study. However, the quality and validity of these recall measures might be lower relative to the more proximally measured adherence collected during the SMART study.

The third advantage of the SMART design as compared to multiple, one-stage-at-a-time, randomized trials is that it facilitates the avoidance of deleterious cohort effects (for more details, see Murphy et al. 2007b and Nahum-Shani et al. 2011a,b) because participants who remain in a SMART study may be more representative of the population than are participants who enroll and remain in single-stage trials. There are two reasons why this may be the case. The first is related to the fact many single-stage trials are used to compare a set of interventions that are fixed over time; thus, participants who do not respond well to their assigned intervention have no options other than nonadherence or drop-out from the study. The lack of options for nonimproving participants within the context of the trial may lead to study attrition or nonadherence. Since the SMART design alters the intervention of nonresponding participants, these participants may be more likely to remain in the study than

would be the case if the intervention were not altered. Alternately, if participants do well, then the intensity or burden of the intervention may be reduced. The second reason why one can expect to recruit a more representative group of participants who remain in SMART studies than would otherwise be the case is because the group of nonresponders to the initial treatment can be expected to be more representative of treatment nonresponders in general. That is, to the extent that nonresponders who are willing to enroll in a nonresponders study are more motivated to adhere and are more receptive to the intervention than the general population of nonresponders, the single-stage nonresponders study will lead to nonrepresentative results.

As mentioned above, the second plausible alternative to a SMART design is a randomized trial in which participants are randomized between a fully formed adaptive intervention and an appropriate alternative (standard or usual care, attention control, etc.). In this case researchers might utilize available literature, clinical experience, results from prior randomized trials, and a variety of well-established theoretical principles [e.g., the stepped care model by Sobell & Sobell (2000) advocates starting patients on the least intensive intervention and stepping up only when necessary] to completely formulate the adaptive intervention. This includes selecting the tailoring variables, operationalizing the decision rules, and deciding which interventions will be employed in each stage.

The advantages of SMART designs over a randomized comparison of the fully formed adaptive intervention with a control are related to the fact that the SMART design enables clinical scientists to open the “black box” using randomized comparisons. In contrast, the randomized two-group comparison approach is designed with a very different goal, that is, to provide information on whether the adaptive intervention outperforms the best alternative. SMART designs can be used to proactively compare intervention options (e.g., the dose, the type of the intervention, the timing of the treatment) at different stages of the adaptive

intervention. To address these issues in a randomized two-group trial setting, investigators would have to rely on nonexperimental variation in intervention options received (e.g., variation in dose due to nonadherence or implementation infidelity). Secondary analysis of data from SMART studies may also assess the usefulness of candidate tailoring variables at each stage. This enables us to address questions of clinical and theoretical interest such as whether the intervention provided to nonresponding patients should differ depending on the patient’s level of adherence to the initial intervention. In a randomized two-group trial, there is no randomization to different second-stage interventions, and thus again analyses must depend on variation due to nonadherence and implementation infidelity. It is well known that untestable, strong assumptions must hold to ensure that findings based on such nonrandomized comparisons are not biased by confounding (Hernán et al. 2000, 2002; Robins et al. 2000).

The SMART Experimental Design Versus the Adaptive Experimental Design

The **adaptive experimental design** is defined by Dragalin (2006) as a multistage study design in which accumulating data are used to decide how to modify aspects of the study without undermining the validity and integrity of the trial. Chow & Chang (2008) provide a thorough review of various types of adaptive experimental designs. **The adjective “adaptive” comes from the fact that these experimental designs are adapted in that the randomization probabilities for future participants are biased toward the best-performing interventions.** In one type of adaptive experimental design, called a group sequential design, decisions concerning whether to stop the trial early and whether to adjust the target sample size are based on outcome data from the participants enrolled earlier in the trial (Pocock 1977; for examples, see Bandyopadhyay & Dragalin 2007, Honigsmann et al. 2007).

The SMART experimental design and the adaptive experimental design are quite

different experimental approaches. Although both approaches involve multiple stages, the word “stage” has a different meaning in SMARTs as opposed to its meaning in adaptive experimental designs. In SMART designs, each stage is an intervention stage that corresponds to a critical decision point in the care for a patient, whereas in adaptive experimental designs, each stage corresponds to an experimental stage that involves a new set of participants. The word “adaptive” also has a different meaning in the context of SMARTs as opposed to adaptive experimental designs. In SMART designs, the intervention offered to a specific participant is adapted on the basis of information obtained on this participant (e.g., the participant’s signs of early response or level of adherence to the initial interventions), whereas in adaptive experimental designs, the randomization probabilities are adapted on the basis of information obtained from other participants in prior stages (e.g., to the primary outcomes obtained from other participants). Recently, methodologists have proposed adaptive experimental designs for use in developing individualized or personalized one-stage interventions (Zhou et al. 2008). Furthermore, Thall and colleagues (2007b) have proposed combining the adaptive experimental design with the SMART design for use in constructing adaptive interventions in cancer treatment.

EXAMPLES OF SMART STUDIES

Below we provide examples of four SMART studies that are completed or currently ongoing: (a) the Adaptive Characterizing Cognition in Nonverbal Individuals with Autism (CCNIA) Developmental and Augmented Intervention (Kasari 2009a,b) for school-age, nonverbal, children with autism spectrum disorders; (b) the Adaptive Pharmacological and Behavioral Treatments for children with ADHD; (c) the Adaptive Reinforcement-Based Treatment for Pregnant Drug Abusers (RBT) (Jones 2010); and (d) the ExTEND study for alcohol-dependent individuals (Oslin 2005). After reviewing the studies, we discuss

similarities and differences between the designs. We use the ExTEND study to illustrate how one can analyze the data resulting from SMART studies in the section titled Using Data from a SMART to Inform the Construction of an Optimized Adaptive Intervention.

Adaptive CCNIA Developmental and Augmented Intervention

Rationale. The Adaptive CCNIA Developmental and Augmented Intervention Study (C. Kasari, P.I.) focuses on nonverbal children with autism spectrum disorders who have not made satisfactory progress by age 5 even though they have received traditional intensive interventions. These children experience poorer long-term outcomes (Lord 2000) than those of verbal children with autism spectrum disorders. When this trial was designed, no known efficacious interventions existed for these children, even though they represent 25% to 30% of children with autism spectrum disorders (Anderson et al. 2007). Given the lack of known efficacious interventions combined with the poor outcomes of traditional interventions, the research team was interested in planning a “rescue” intervention that would be used if the initial intervention was insufficiently successful. Thus, the critical questions that motivated this study include which intervention to provide initially and, if the child does not improve his/her expressive language skills, which rescue intervention to provide.

Interventions. The joint attention/joint engagement (JAE) intervention was combined with two interventions, enhanced milieu teaching (EMT) and augmentative and alternative communication (AAC). JAE (Adamson et al. 2004; Kasari et al. 2006, 2008) was developed to facilitate a state of supported or coordinated joint engagement between the child and a social partner. Both EMT and AAC were developed to facilitate expressive language in young children with developmental disabilities. EMT (Hancock & Kaiser 2006) is a naturalistic language intervention that promotes functional

use of new language forms in the context of everyday interactions with parents and other social partners. The AAC intervention utilizes a developmentally chosen augmentative communication device (Cafiero 2005) to facilitate communicative exchanges within play routines and daily activities. Both EMT and AAC were adapted for 5- to 8-year-old children and integrated with JAE to form two interventions, JAE+EMT and JAE+AAC. More intensive versions of both JAE+EMT and JAE+AAC included additional sessions provided by a skilled child therapist and additional training with the parent to promote parent and child generalization. Parents are also encouraged to implement the strategies they learned in the intervention across home routines, thereby potentially increasing the hours of interventions children receive. Overall, four intervention options are considered: JAE+EMT, JAE+AAC, intensified JAE+EMT, and intensified JAE+AAC.

Tailoring variables. To determine whether a child is exhibiting early signs of response to the initial intervention, the investigators used assessments of the child's improvement in (a) total social communicative utterances, (b) percentage communicative utterances, (c) number different word roots, (d) mean length of utterance in words, (e) number of utterances where the function is to comment (rather than request), (f) words per minute, and (g) unique word combinations (included only if the child's target talk consists of more than two words). These assessments address the range of changes in spoken and augmented language that might be expected in three months and that represent meaningful language change. Using these assessments, the investigators constructed 14 variables that represent change in spoken and augmented language, two per each of the seven assessments. More specifically, for each assessment, the first variable was calculated as the difference in the average assessment between the first two intervention sessions and the last two intervention sessions during the first stage of the intervention; the second variable was calculated as the difference between the assessment at the screening visit and the month-three visit. The above measures are collected via videotapes of the child and therapist sessions.

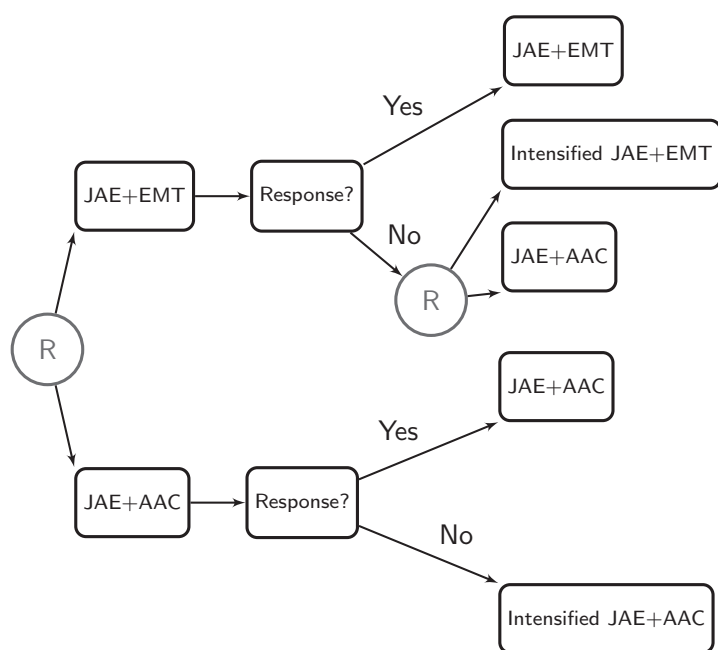


Figure 3

The Characterizing Cognition in Nonverbal Individuals with Autism Trial design (C. Kasari, P.I.). AAC, augmentative and alternative communication; EMT, enhanced milieu teaching; JAE, joint attention/joint engagement.

Design. As shown in **Figure 3**, at the first stage of the intervention, children are randomized to JAE+EMT or JAE+AAC. This stage lasts three months, during which children in both intervention arms receive biweekly intervention sessions. After three months, children are classified as responders or nonresponders to the initial intervention. The criterion for response is 25% or more improvement on at least 50% of the 14 measures described above. Children who do not meet this criterion are classified as nonresponders. The second stage is also of three months' duration. At this stage, children who respond to the initial interventions continue with the same initial intervention. Children who do not respond to initial JAE+EMT are randomly assigned either to

receive intensified JAE+EMT or to switch to JAE+AAC. Because taking the augmentative communication device from children who do not respond to JAE+AAC was deemed inappropriate, intensified JAE+AAC is offered to children who are classified as nonresponders to JAE+AAC. The total study duration is six months, with a follow-up assessment at nine months.

Embedded adaptive interventions. Three adaptive interventions are embedded in this SMART design: (a) Begin with JAE+EMT, and after three months, continue JAE+EMT if the child adequately responds, or switch to JAE+AAC if the child does not respond; (b) begin with JAE+EMT, and after three months, continue JAE+EMT if the child adequately responds, or offer more intensive JAE+EMT if the child does not respond; and (c) begin with JAE+AAC, and after three months, continue JAE+AAC if the child adequately responds, or offer more intensive JAE+AAC if the child does not respond.

Outcome measures and specific aims.

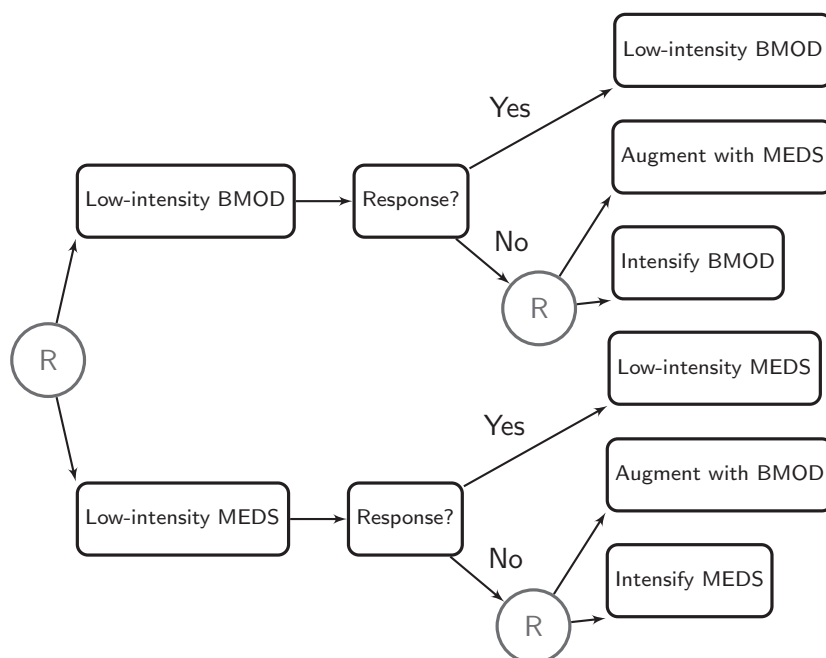
Outcomes include three time-varying variables that measure communicative and expressive language abilities: (a) the number of words used spontaneously during parent-child interaction, (b) the number of communicative functions used for each word during parent-child interactions, and (c) the generalization of spontaneous words to express multiple communication functions. These variables are obtained from videotapes of child-mother interactions. The primary aim of this study is to test the main effect of beginning with JAE+AAC versus beginning with JAE+EMT on the above time-varying outcomes. A secondary aim of the study is to examine potential baseline moderators of the difference between the three embedded adaptive interventions. The baseline variables included severity of repetitive/compulsive behaviors, degree of apraxia, and developmental variables (based on cognitive and language test results). In particular, the research team hypothesized that children with greater severity of

apraxia would do better on JAE+AAC than on JAE+EMT because the communication device would better provide a means to communicate.

Adaptive Pharmacological and Behavioral Treatments for Children with ADHD Trial

Rationale. In contrast to the setting for the CCNIA study, a number of acutely effective pharmacological and behavioral interventions are presently available for treating children with ADHD (e.g., Fabiano et al. 2009, Hammerness et al. 2009, Pelham & Fabiano 2008, Wigal et al. 2011). There has been considerable debate, however, on whether the first-line intervention should be pharmacological or behavioral; for example, a task force of the American Psychological Association recommends psychosocial first (Brown et al. 2007), whereas the guidelines of the American Academy of Child and Adolescent Psychiatry (2007) recommend using medication first. Thus a critical decision is whether to initially use a behavioral or pharmacological intervention. Because anywhere from 20% to 50% of children can be expected to insufficiently respond to the initial intervention, a second critical decision concerns the choice of “rescue” intervention for these children.

Interventions. The interventions include differing doses of methylphenidate (a psychostimulant drug) and differing intensities of behavioral modification (consisting of a school-based component with the teacher, a Saturday treatment component involving social skills development, and a parent-training component targeted at helping parents to identify problematic behaviors with the relevant child-functioning domains). The higher-dose option for methylphenidate includes late-afternoon doses, if needed. The higher-intensity option for the behavioral modification includes more intensive training in social skills in the school-based component and, if needed, both additional individual parent training sessions that target specific behavior management issues and practice sessions with children.

**Figure 4**

Design of the Adaptive Pharmacological and Behavioral Treatments for Children with ADHD Trial (W. Pelham, P.I.). BMOD, behavior modification; MEDS, low-dose medication.

Tailoring variables. Two well-known instruments were employed to assess whether children were responding to the first-stage intervention: the Impairment Rating Scale (IRS) (Fabiano et al. 2006) and an individualized list of target behaviors (ITB) (e.g., Pelham et al. 1992). The IRS provides a comprehensive index of a child's impairment in various domains such as peer relationships, classroom behavior, family functioning, and academic achievement. The ITB was used to assess improvement on child-specific behavior goals. Both were reported by the regular classroom teachers.

Design. As shown in **Figure 4**, children were randomly assigned to begin with low-intensity behavioral modification or with low-dose medication. This stage lasts two months, after which the ratings from IRS and ITB were used monthly to assess the child's response to the initial intervention. The criterion for nonresponse at each month was an average performance of less than 75% on the ITB and

a rating of impairment in at least one domain on the IRS. Children who did not meet the criteria for nonresponse continued with the initial intervention assigned to them, whereas nonresponders were rerandomized. There are two ways to conceptualize the second-stage intervention options that were offered to nonresponders. One way would be to focus on the type of the intervention offered. Children who did not respond adequately to low-intensity behavioral modification were rerandomized to either an intensified behavioral modification or to behavioral modification augmented with medication, whereas children who did not respond adequately to low-dose medication were rerandomized to either increased dose of medication or to medication augmented with behavioral modification. Another way to frame the second-stage options for nonresponders would be to focus on the intervention tactic. That is, all nonresponders were rerandomized either to receive an intensified version of the initial intervention or to augment the initial

intervention with the alternative intervention option. The study duration was eight months during a school year, with a follow-up assessment during November of the following school year.

Embedded adaptive interventions. Four adaptive interventions are embedded in this design. For example, low-intensity behavioral modification is offered at the start; if the child does not respond adequately, the intervention is augmented with a low dose of medication; otherwise, behavioral modification of the same intensity is continued.

Outcome measures and specific aims. Outcome measures included time-varying teacher and parent ratings of child behavior as well as the ending dosages of the behavioral modification and medication. The primary aim of this study was to test the main effect of beginning with low-dose medication versus beginning with low-intensity behavioral modification on the rate of nonresponse and ending dosages. Secondary aims included assessing how baseline variables (e.g., severity of impairment, comorbid child psychopathology, and prior medication history) influence (a) the difference between the two initial interventions, (b) the difference between the second-stage intervention options for nonresponding children, and (c) the difference between the four adaptive interventions embedded in the design.

Adaptive Reinforcement-Based Treatment for Pregnant Drug Abusers

Rationale. Prenatal drug use is often associated with an increased risk of pregnancy complications and adverse neonatal outcomes such as pre-eclampsia, stillbirth, and premature labor. RBT has been shown to be an efficacious intervention (Jones et al. 2005, 2011); however, two obstacles remain. First, RBT is time-consuming on the part of the participant, and second, about 40% of participants do not respond as well as desired. Thus, the aim of this study is to address three critical questions:

(a) Whether the frontline version of RBT can be reduced in intensity and scope; (b) whether a woman who does not respond quickly to the frontline version of RBT should continue on the same version or be moved to a more-intensive, larger-scope version of RBT; and (c) whether the intensity and scope of RBT can be reduced if a woman responds quickly to the frontline version. The general aim of this study is to develop an RBT-based adaptive intervention that is more cost-effective.

Interventions. In this design, four variants of one successful multicomponent intervention, RBT, are being considered. RBT focuses on areas of change that are likely to reduce drug use or chances of relapse by facilitating participants with nondrug-related reinforcers such as life skills training, recreational therapy, and employment. The four RBT variants, ordered in increasing intensity and scope, are aRBT, rRBT, tRBT, and eRBT. tRBT includes 11 key elements:

1. Functional assessment of drug/alcohol use
2. Use of behavioral contracts
3. Motivational interviewing style of therapy
4. Graphing and monitoring of critical identified behaviors to sustain abstinence
5. Abstinence-contingent access to elements 8–11 below as well as other tangible reinforcers
6. Outreach upon first noncompliant event
7. Individual therapy
8. Recreation paid for by the program
9. Job club
10. Social club including free lunch
11. Skills-building modules

rRBT consists of all elements in tRBT, with some of the key elements provided in a reduced level. For example, social and job club are provided once a week compared with three times a week in tRBT. In aRBT, the least-intense version of RBT, the scope and frequencies of some of the RBT components are further reduced. For example, behavioral

graphing focuses only on drug abstinence, and no tangible reinforcers are provided. Individual therapy sessions are provided once a week compared with three times per week in tRBT, and recreational activities and job club are held less frequently. eRBT is the most-intense version of RBT. It provides participants with additional intervention sessions, such as home visits due to nonresponse or relapse. For example, individual therapy sessions in eRBT are held twice as often as in tRBT.

Tailoring variables. Since early drug use lapse or relapse and adherence are known predictors of poor intervention response (Ciraulo et al. 2003), response is assessed via self-reported drug use, results of urine tests, and attendance on intervention days.

Design. As shown in **Figure 2**, at the first stage of the intervention, participants are randomly assigned to either tRBT or rRBT. After two weeks, participants are classified as responders or nonresponders to the first-stage intervention. The criterion for nonresponse is missing an intervention day with no excuse, or a positive opioid/cocaine urine specimen, or self-report use of either drug. At the second stage, nonresponders are rerandomized to one of the two possible subsequent interventions that are as intensive or more intensive in dose and scope as the initial intervention. More specifically, nonresponders to tRBT are rerandomized to either tRBT or eRBT, whereas nonresponders to rRBT are rerandomized to either rRBT or tRBT. Responders are randomized to one of the two subsequent interventions that are as intensive or less intensive in dose and scope as the initial intervention. More specifically, responders to rRBT are rerandomized to either rRBT or aRBT, whereas responders to tRBT are rerandomized to either tRBT or rRBT. The study duration is the duration of the pregnancy plus six weeks postpartum.

Embedded adaptive interventions. Because both responders and nonresponders are rerandomized in this study, eight adaptive interventions are embedded in this SMART design. An

example of an embedded adaptive intervention is to begin by offering rRBT; after two weeks, continue rRBT if the patient is showing adequate response; otherwise, switch to tRBT.

Outcome measures and specific aims.

Primary outcomes include treatment completion (delivery while in program) and heroin and cocaine use. A primary aim of the study is to compare always tRBT (i.e., offering tRBT throughout the pregnancy) with always rRBT (i.e., offering rRBT throughout the pregnancy). This primary aim concerns the comparison of two nonadaptive interventions (always tRBT versus always rRBT) that are embedded in the design. Secondary aims involve assessing the usefulness of candidate tailoring variables, such as the amount of illegal activity (e.g., prostitution).

Extending Treatment Effectiveness of Naltrexone

Rationale. NTX is an opioid receptor antagonist used in the prevention of relapse to alcoholism. Even though NTX has been shown to be efficacious, its use by clinicians has been limited. One reason is that it is difficult to adhere to NTX because this drug diminishes the pleasurable effects of alcohol use. Furthermore, from a clinical perspective, there is little guidance in answering critical questions such as (a) the extent of drinking behavior that reflects nonresponse to NTX, (b) the type of treatment that would be useful for participants who do not respond adequately to NTX, and (c) the type of treatment that would be useful in reducing the chance of relapse among individuals who respond adequately to NTX. The ExTEND SMART study (D. Oslin, P.I.) was designed to address these three critical questions. Notice that the first critical question in the ExTEND study concerns the definition of nonresponse, that is, the number of heavy drinking days that should be used as the criteria for classifying a participant as an early nonresponder to NTX.

Interventions. Four interventions are considered. These are NTX, medical management

(MM), combined behavioral intervention (CBI), and telephone disease management (TDM). MM is a face-to-face, basic, minimal clinical support for the use of effective pharmacotherapy and reduction in drinking (Pettinati et al. 2004, 2005). CBI is a multicomponent intervention that includes components targeting adherence to pharmacotherapy and enhancement of participant motivation for change. This intervention includes family involvement when possible and emphasizes the utilization of the participant's social/community context to reinforce abstinence (Longabaugh et al. 2005, Miller et al. 2003). TDM includes the same content as MM, but it is delivered via telephone.

Tailoring variable. The number of self-reported heavy drinking days was used to assess early nonresponse to NTX.

Design. The design of the trial is shown in **Figure 5**. The same initial intervention

(NTX+MM) was provided to all participants. The randomization at the first stage was to one of two criteria for classifying a participant as an early nonresponder: (a) the stringent criterion for nonresponse, in which two or more heavy drinking days during the first eight weeks of NTX treatment is defined as nonresponse, or (b) the lenient criterion for nonresponse, in which five or more heavy drinking days during the first eight weeks of NTX treatment is defined as nonresponse. Participants were assessed weekly for drinking behavior. If a participant met his/her assigned criteria for nonresponse during the first eight weeks of NTX treatment, he/she was immediately rerandomized to one of the two rescue intervention options for nonresponders: CBI+MM+placebo or CBI+MM+NTX. If the participant did not meet his/her assigned nonresponse criteria during the first eight weeks of NTX treatment, he/she was rerandomized at the end of week eight to one of the two "maintenance" intervention options for responders: to either

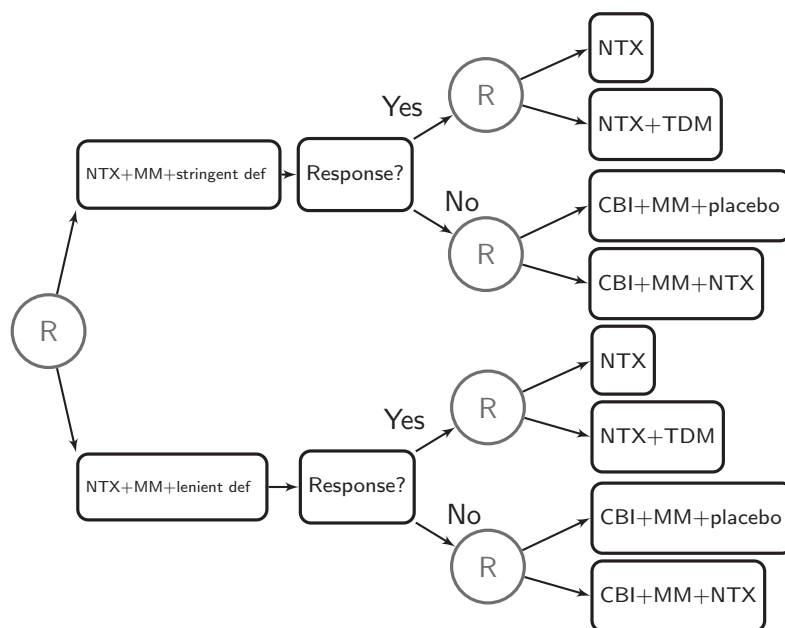


Figure 5

The Extending Treatment Effectiveness of Naltrexone (ExTEND) Trial design (D. Oslin, P.I.). CBI, combined behavioral intervention; MM, medical management; NTX, naltrexone; TDM, telephone disease management.

open-label NTX+TDM or open-label NTX. The study duration was six months.

Embedded adaptive interventions. Eight adaptive interventions are embedded in this SMART design. One example is to begin with NTX+MM for up to eight weeks; if the patient experiences two or more heavy drinking days during these eight weeks, immediately offer CBI+MM+NTX; otherwise, after eight weeks provide NTX+TDM.

Outcome measures and specific aims. The primary outcomes are the percent of heavy drinking days and percent drinking days over the last two months of the study. A primary aim was to test the main effect of the rescue intervention options for nonresponders, that is, whether there is a difference in drinking outcomes for nonresponding participants treated with CBI+MM+NTX versus CBI+MM+placebo. A secondary aim of the study was to assess candidate-tailoring variables for the rescue intervention options for nonresponders. One candidate-tailoring variable was the average daily pill counts during the first stage of the intervention, measured by pill counts from blister cards. Note that this variable is an outcome of the first-stage intervention (NTX+MM) as opposed to a baseline variable.

Comparison of SMART Studies

The four examples of SMART studies (i.e., the CCNIA, ADHD, RBT, and ExTENd) discussed above can be compared along four key dimensions: (a) the extent of multiple randomizations used in the design; (b) the nature of the randomizations used, or in other words, the type of the critical decisions addressed by the randomizations; (c) the extent of treatment modalities underlying the intervention options that are offered; and (d) the primary questions the study was sized to address.

Concerning the first dimension, the four studies described above are ordered to reflect the extent to which multiple randomizations

were used in the design. In the CCNIA study, only a subset of the nonresponding children is rerandomized. Specifically, children who do not respond to JAE+EMT are randomly assigned either to receive intensified JAE+EMT or to switch to JAE+AAC. In this study, the decision not to rerandomize nonresponders to JAE+AAC was made for ethical considerations. Switching from JAE+AAC to JAE+EMT would mean that the communication device that was given to children as part of JAE+AAC had to be taken away. The researchers felt that this was ethically inappropriate and hence do not rerandomize children who do not respond to JAE+AAC, but rather offer an intensified version of JAE+AAC. The ADHD study, on the other hand, involved rerandomization of nonresponders in both arms of the initial intervention (i.e., medication and behavioral modification). Guided by past studies, the investigators expected a nonnegligible rate of nonresponse to both frontline low-dose interventions. Hence, they were interested in finding the best rescue intervention for children who showed early signs of nonresponse. Children who responded to the frontline intervention were not rerandomized in this study. The use of randomization is even greater in the RBT study, in which responders and nonresponders to both initial interventions are rerandomized. Acknowledging the burden associated with the frontline intervention, the investigators were interested not only in finding the best rescue intervention for nonresponding participants but also in assessing whether the intensity of the frontline intervention could be reduced for responding participants. Similarly, in the ExTENd design, both responders and nonresponders were rerandomized.

Notice that the number of adaptive interventions embedded in the SMART design increases to the extent that the design involves more randomizations, that is, to the extent that the design addresses more critical decisions. Only three adaptive interventions are embedded in the CCNIA SMART design, in which only two critical decisions were addressed by randomization: one concerning the frontline

intervention and the other concerning the rescue intervention for nonresponders to one of the frontline interventions. Four adaptive interventions are embedded in the ADHD SMART design, in which three critical decisions were addressed by the randomizations: The first concerns the frontline intervention, the second concerns the rescue intervention for nonresponders to one of the initial intervention options (i.e., medication), and the third concerns the rescue intervention for nonresponders to the other initial intervention option (i.e., behavioral modification). Finally, eight adaptive interventions are embedded in the RBT and the ExTEND SMART designs, in which five critical decisions were addressed by randomizations. For example, in the RBT study, the five critical decisions concern (a) the frontline intervention options, (b) the rescue intervention options for responders to one of the frontline intervention options (e.g., tRBT), (c) the rescue intervention option for responders to the other frontline intervention option (e.g., rRBT), (d) the maintenance intervention option for nonresponders to one of the frontline intervention options (e.g., tRBT), and (e) the maintenance intervention option for nonresponders to the other frontline intervention option (e.g., rRBT).

Concerning the second dimension, the randomizations in SMART designs may correspond to various types of critical decisions. In the ADHD and the CCNIA trials, the first-stage randomization corresponds to a decision concerning the frontline treatment, in which two different treatment options are compared (medication versus behavioral modification in the case of the ADHD study; JAE+EMT versus JAE+AAC in the case of the CCNIA study). The second-stage randomization in these trials corresponds to a decision concerning the rescue tactics for nonresponding children, where two treatment tactics are compared (intensify versus augment with the alternative intervention options in the case of the ADHD study; intensify versus switch to JAE+AAC in the case of the CCNIA). In the ExTEND study, on the other hand, the first-stage

randomization corresponds to a decision concerning the definition of nonresponse or in other words the criteria for classifying participants as nonresponders to NTX (stringent versus lenient definition of nonresponse). The randomizations in SMART designs may correspond to other critical decisions, such as decisions concerning the individual's participation in treatment (e.g., who should set health-related goals, the participant or the care provider?), the location of the intervention offered (e.g., is it better to offer treatment at home or at the clinic?), the provider of the intervention (e.g., should the parent or the teacher intervene?), the mode of delivery (e.g., is face-to-face delivery better than Internet-based delivery?), or the timing of treatment (e.g., is it better to intervene immediately or at some later point?)

Concerning the third dimension, the number of treatment modalities that are involved in SMART studies may vary. For example, the ADHD SMART study contrasts intervention options that are based on two different treatment modalities, behavioral and pharmacological. The RBT SMART study, on the other hand, contrasts variants of the same intervention (RBT) and hence involves only one treatment modality. The number of treatment modalities utilized in SMART studies depends on the scientific evidence and the research questions at hand.

Concerning the fourth dimension, SMART designs may vary in the primary research questions they were sized to address. Most SMART studies are sized to address primary research questions concerning the first-stage randomization. Secondary research questions typically concern (a) the comparison of second-stage intervention options, (b) the comparison of embedded adaptive interventions, and (c) the usefulness of candidate tailoring variables (beyond the tailoring variable already embedded in the design). For example, the ADHD data were sized to address a primary research question concerning the comparison between the two frontline intervention options (i.e., medication and behavioral modification). The secondary research questions concerned the comparison

of rescue tactics for nonresponding children, the comparisons of the four adaptive interventions that are embedded in the design, and the usefulness of candidate tailoring variables such as baseline severity of ADHD symptoms and adherence to frontline intervention. Other SMART designs are sized to address a primary research question concerning the second-stage intervention options. For example, in the ExTEND study, the primary aim was to test the main effect of the rescue intervention options for nonresponders to the first-stage intervention (NTX+MM), that is, to assess whether there is a difference in drinking outcomes for nonresponding participants treated with CBI+MM+NTX versus CBI+MM+placebo. SMART designs can also be sized to address primary research questions concerning the adaptive interventions that are embedded in the design. For example, in the RBT study, the primary aim concerns the comparison of two nonadaptive interventions (always tRBT versus always rRBT) that are embedded in the design, and the secondary aims concern the usefulness of candidate tailoring variables such as the amount of illegal activity (e.g., prostitution).

USING DATA FROM A SMART TO INFORM THE CONSTRUCTION OF AN OPTIMIZED ADAPTIVE INTERVENTION

A variety of interesting questions can be addressed using SMART study data. Frequently the primary analysis concerns the main effect of the first-stage intervention options. In the ADHD example, this is the main effect of starting with a low-intensity behavioral modification versus low-dose medication. Additionally, secondary data analyses can be used to assess moderators of the first-stage intervention options, such as baseline participant characteristics. Other data analyses may concern the main effect of the second-stage intervention options. In the ExTEND example, this would be the main effect of NTX+TDM versus NTX alone for responding participants and the main effect of CBI+MM+NTX versus

CBI+MM+placebo for nonresponding participants. Secondary data analyses can be used to assess moderators of the effect of the second-stage intervention options such as baseline measures and outcomes observed during the first-stage treatment. Data analyses may also concern the comparison of embedded adaptive interventions. Recall that in the ExTEND study there are eight embedded adaptive interventions. Nahum-Shani et al. (2011a,b) provide analysis methods for addressing primary and secondary research questions using data from SMART studies. Below, we use data from the ExTEND SMART study to illustrate the comparison of the eight embedded adaptive interventions.

THE EXTEND TRIAL EXAMPLE

Sample

The ExTEND study included 302 participants (70% white, 86% male, and 28% above the age of 55). Forty-nine participants dropped out prior to experiencing two heavy drinking days and thus experienced no effect of the first-stage randomization. The effective sample size for the randomized study was 253 participants. Only three additional participants dropped out during the first-stage treatment; the data from these participants were deleted, leaving a sample size of 250 participants. At the first stage, 123 participants were randomized to the stringent definition of nonresponse (two or more heavy drinking days), and 127 participants were randomized to the lenient definition of nonresponse (five or more heavy drinking days). Sixty-seven participants met their assigned definition of nonresponse; of these participants, 34 were rerandomized to CBI+MM+NTX, and the remaining 33 were rerandomized to CBI+MM+placebo. Of the 183 responding participants, 92 were rerandomized to NTX+TDM, and the remaining 91 responders were rerandomized to NTX only. Forty-one participants dropped out at the second stage. To accommodate the resulting missing data, multiple imputation was used in the analyses below.

Table 3 Estimated probability of drinking at the end of stage two, for each of the eight embedded adaptive interventions

Adaptive intervention	Sample size	Number of responders	Estimated probability of drinking between second and third visit, stage two
(Lenient, NTX, CBI+MM+placebo)	60	50	0.49
(Stringent, NTX, CBI+MM+placebo)	64	41	0.76
(Lenient, NTX+TDM, CBI+MM+placebo)	63	53	0.40
(Stringent, NTX+TDM, CBI+MM+placebo)	62	39	0.64
(Lenient, NTX, CBI+MM+NTX)	64	50	0.57
(Stringent, NTX, CBI+MM+NTX)	61	41	0.71
(Lenient, NTX+TDM, CBI+MM+NTX)	67	53	0.48
(Stringent, NTX+TDM, CBI+MM+NTX)	59	39	0.58

Abbreviations: CBI, combined behavioral intervention; MM, medical management; NTX, naltrexone; TDM, telephone disease management.

Recall that the ExTEND study provides data on eight embedded adaptive interventions. The number of participants who contribute data that inform the estimated average of each adaptive intervention is given in **Table 3**. Each participant is consistent with two adaptive interventions (i.e., the observed primary outcome of each participant is used in estimating the average primary outcome of two adaptive interventions; for more details, see Nahum-Shani et al. 2011a). For example, outcome data that inform the estimated average outcome of two adaptive interventions are provided by a responder who is assigned the stringent definition of nonresponse (two or more heavy drinking days), is responding at the end of eight weeks, and then is assigned to continue on NTX: (a) use a stringent definition of nonresponse, then offer NTX to responders and CBI+MM+NTX to nonresponders; and (b) use a stringent definition of nonresponse, then offer NTX to responders and CBI+MM+placebo to nonresponders.

Measures

Data for a participant in the ExTEND trial is denoted by $(O_1, A_1, O_2, A_{2R}, A_{2NR}, Y)$. Here O_1 is the vector of pretreatment information described below. A_1 denotes the nonresponse definitions at the first stage, coded $A_1 = 1$ for the stringent nonresponse definition (two or more heavy drinking days) and $A_1 = -1$ for the lenient nonresponder definition (five or more heavy drinking days). O_2 is the vector of intermediate outcomes, which includes response to the first-stage intervention, drinking behaviors, and adherence to medication. A_{2R} denotes the second-stage intervention options for responding participants, coded $A_{2R} = 1$ for NTX+TDM and $A_{2R} = -1$ for NTX. A_{2NR} denotes the second-stage intervention options for nonresponding participants, coded $A_{2NR} = 1$ for CBI+MM+NTX and $A_{2NR} = -1$ for CBI+MM+placebo. Y is the primary outcome. For these illustrative analyses, Y is an indicator of whether the participant was ($Y = 0$)

Table 4 Baseline summary statistics: median/mode values

Parameter	Type	Median/mode
Baseline: PDHD	Continuous	0.15
Baseline: PDNH	Continuous	0
Baseline: ethnic	Class (level = 2)	0 (non-Hispanic/Latino)
Baseline: age	Continuous	49
Baseline: gender	Class (level = 2)	0 (male)
Baseline: education years	Continuous	14

Abbreviations: PDHD, percent days of heavy drinking prior to entering the study; PDNH, percent days of drinking but not heavy drinking prior to entering the study.

or was not ($Y = 1$) abstinent between the second and the third visit occurring at the end of the second stage. Baseline measures include the following: (a) percent days of heavy drinking prior to entering the study (PDHD), (b) percent days of drinking but not heavy drinking prior to entering the study (PDNH), (c) gender, (d) ethnicity, (e) age, and (f) education (in years). In coding the two classification variables—gender and ethnicity—male and non-Hispanic/Latino were chosen as the reference categories, respectively. Continuous baseline variables were centered at their median values. The median/mode values are given below (Table 4).

Illustrative Data Analyses

Here we present an illustrative comparison of the eight adaptive interventions embedded in the ExTEND SMART design. We use standard software (e.g., the SAS GENMOD procedure: SAS Inst. 2008) with modifications involving weighting (Nahum-Shani et al. 2011a,b; Orellana et al. 2010; Robins et al. 2008). These modifications are necessary because, as noted above, the data for each participant are consistent with more than one adaptive intervention. The logistic regression model below can be used to compare the eight adaptive interventions embedded in the ExTEND trial:

$$\begin{aligned}
 \logit P(Y = 1 | A_1, A_{2R}, A_{2NR}, O_1) \\
 = \beta_0 + \beta_1 * A_1 + \beta_2 * A_{2R} + \beta_3 * A_{2NR} \\
 + \beta_4 * A_1 * A_{2R} + \beta_5 * A_1 * A_{2NR} + \gamma^T O_1.
 \end{aligned}
 \tag{1}$$

Here O_1 is a vector of selected baseline measures, β_0 is the intercept; β_1 – β_3 are the regression coefficients representing the effects of the nonresponse definitions and the second-stage interventions for responders and nonresponders, respectively; β_4 , β_5 are regression coefficients of the interactions between the nonresponse definitions and the second-stage intervention options for responders and nonresponders, respectively; γ is the regression coefficient for the baseline variables. In general, additional interaction terms could be included in Model 1, such as $A_1 * O_1$, $A_{2R} * A_{2NR}$, $O_1 * A_{2R}$, $A_1 * A_{2NR} * A_{2R}$.

Results

The missing data in this data set are handled with multiple imputation using the R mice procedure (van Buuren & Groothuis-Oudshoorn 2010). The estimated coefficients in Model 1 are given in Table 5.¹

The results reported in Table 5 indicate that the coefficients of A_1 , and $A_1 * A_{2NR}$ are significantly different from zero ($\hat{\beta}_1 = 0.39$; $\hat{\beta}_5 = -0.14$, $p < 0.05$, respectively), and the

¹The coefficients in Model 1 were estimated using the SAS GENMOD procedure (SAS Inst. 2008) for each of 10 imputed datasets. Because each participant's observations are consistent with more than one adaptive intervention, we had the SAS GENMOD procedure (SAS Inst. 2008) provide robust standard errors (for a detailed justification, see Nahum-Shani et al. 2011a,b), which are then used to form the test statistics for Model 1. These are the inputs of the SAS MIANALYZE procedure, which produces estimates and inferences adjusted for the missing data.

Table 5 Results (parameter estimates) for Model 1

Parameter	Estimate	Robust SE	95% Confidence limits		P-value
Intercept	0.34	0.21	−0.07	0.74	0.1021
Baseline: PDHD	2.90	0.85	1.23	4.57	0.0007
Baseline: PDNH	4.68	1.76	1.23	8.14	0.0079
Baseline: ethnic	−1.72	0.79	−3.27	−0.18	0.0289
Baseline: age	−0.04	0.02	−0.07	0.00	0.0302
Baseline: gender	0.68	0.46	−0.22	1.57	0.1402
Baseline: education years	0.10	0.06	−0.02	0.22	0.1194
A_1 (Nonresponse definition)	0.39	0.16	0.09	0.70	0.0123
A_{2R} (Second-stage intervention for responders)	−0.24	0.14	−0.52	0.05	0.1016
A_{2NR} (Second-stage intervention for nonresponders)	0.02	0.07	−0.11	0.15	0.7707
$A_1 * A_{2R}$	−0.05	0.15	−0.35	0.24	0.7114
$A_1 * A_{2NR}$	−0.14	0.07	−0.27	−0.01	0.0336

Abbreviations: PDHD, percent days of heavy drinking prior to entering the study; PDNH, percent days of drinking but not heavy drinking prior to entering the study.

coefficient of A_{2R} is marginally significant ($\hat{\beta}_2 = -0.24$, $p = 0.10$). The estimated coefficient for A_1 is positive, meaning that on average, assigning the lenient definition of nonresponse (five or more heavy drinking days: $A_1 = -1$) is better than assigning the stringent definition of nonresponse (two or more heavy drinking days: $A_1 = 1$) in achieving abstinence at the end of the second stage. The estimated coefficient of A_{2R} is negative, meaning that offering an intensive maintenance intervention (NTX+TDM: $A_{2R} = 1$) is better than offering the less-intensive maintenance intervention (NTX: $A_{2R} = -1$) in achieving abstinence at the end of the second stage. Notice that the regression coefficient for A_{2NR} is not significantly different from zero ($\hat{\beta}_3 = 0.02$), meaning that on average, there is no difference between the two treatment options for nonresponders in achieving abstinence.

Table 3 gives the estimated probability of drinking between the second and the third visit in stage two for each of the eight embedded adaptive interventions (holding the baseline measures constant at their values as specified in **Table 4**). Each adaptive intervention is abbreviated by a triplet. The first entry represents

the nonresponse definition at the first stage, the second entry represents the maintenance intervention assigned to responders, and the third entry represents the rescue intervention assigned to nonresponders. For example, (Stringent, NTX, CBI+MM+NTX) means: Begin with a combination of MM and NTX for up to eight weeks; if the patient has two or more heavy drinking days within eight weeks, offer NTX+CBI+MM; otherwise, after eight weeks offer NTX.

Table 6 compares four pairs of adaptive interventions in terms of the estimated $\text{logitP}(Y = 1)$. These four comparisons are of special interest: The four rows (in **Table 6**) contrast adaptive interventions that begin with different nonresponse definitions. We used the ESTIMATE statement in SAS GENMOD procedure to estimate the linear combination of the estimated coefficients corresponding to each pairwise comparison. We then used the MIANALYZE procedure with these outputs to obtain combined estimates based on 10 imputed data sets.

The results reported in **Table 6** indicate that the adaptive interventions in comparisons 1 and 2 are significantly different in terms of

Table 6 Four pairwise comparison of adaptive interventions based on the estimated regression coefficients in Table 5

Comp.	Pairs of adaptive intervention		Estimate	Robust standard error	95% Confidence interval		P-value
1	(Lenient, NTX, CBI+MM+placebo)	(Stringent, NTX, CBI+MM+placebo)	−1.17	0.48	−2.12	−0.23	0.015
2	(Lenient, NTX+TDM, CBI+MM+placebo)	(Stringent, NTX+TDM, CBI+MM+placebo)	−0.96	0.43	−1.79	−0.12	0.0249
3	(Lenient, NTX, CBI+MM+NTX)	(Stringent, NTX, CBI+MM+NTX)	−0.61	0.47	−1.54	0.31	0.1927
4	(Lenient, NTX+TDM, CBI+MM+NTX)	(Stringent, NTX+TDM, CBI+MM+NTX)	−0.39	0.42	−1.22	0.43	0.3502

Abbreviations: CBI, combined behavioral intervention; MM, medical management; NTX, naltrexone; TDM, telephone disease management.

achieving drinking abstinence at the end of stage two [estimated $\text{logit } P(Y = 1) = -1.17; -0.96, p < 0.05$, for comparison 1 and 2, respectively]. Notice that comparisons 1 and 2 include adaptive interventions in which nonresponders are offered the less-intensive second-stage intervention (CBI+MM+placebo). These results indicate that if we are to provide nonresponders with the less-intensive rescue intervention (CBI+MM+placebo), then beginning with the lenient definition of nonresponse (five or more heavy drinking days) will result in better achievement of abstinence at the end of the second stage relative to beginning with the stringent definition of nonresponse (two or more heavy drinking days).

The results reported in **Table 6** also indicate that the adaptive interventions in comparisons 3 and 4 are not significantly different in terms of achieving drinking abstinence at the end of stage two [estimated $\text{logit } P(Y = 1) = -0.61; -0.39$, ns, for comparison 3 and 4, respectively]. Notice that comparisons 3 and 4 include adaptive interventions in which nonresponders are offered the more-intensive maintenance intervention (CBI+MM+NTX). These results indicate that we are to provide nonresponders with the more-intensive second-stage

intervention (CBI+MM+NTX); it is less clear whether we should begin with the lenient or the stringent definition of nonresponse.

SUMMARY

Adaptive interventions are both conceptually and scientifically appealing. Here, we discussed the Sequential Multiple Assignment Randomized Trial (SMART), an experimental design useful for developing adaptive interventions in the area of behavioral health. We also described four selected examples of studies in which SMART designs were utilized to address a variety of critical research questions that inform the development of high-quality adaptive interventions. Similarities and differences between these SMART studies along four key dimensions were discussed. Finally, we used data from the ExTEND SMART study to illustrate data analysis methods for addressing questions concerning adaptive intervention. To conclude, SMART designs are increasingly being used to improve the efficacy and effectiveness of interventions in behavioral health. These designs deserve continued research attention and a prominent role in treatment and prevention research.

DISCLOSURE STATEMENT

The authors receive NIH funds focused on the development of SMART studies; this might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We thank Xi Lu for data imputation and Daniel Almirall for valuable comments. We also thank William Pelham, Connie Kasari, and Hendrée Jones for valuable comments and revision in the SMART examples section. We acknowledge support from NIMH grant R01-MH-080015, NIAAA grant R01 AA-014851, NIDA grant P50-DA-010075, and NIAAA grant RC1 AA-019092.

LITERATURE CITED

- Adamson LB, Bakeman R, Deckner DF. 2004. The development of symbol-infused joint engagement. *Child Dev.* 75:1171–87
- Am. Acad. Child Adolesc. Psychiatry. 2007. Practice parameters for the assessment and treatment of children and adolescents with attention-deficit/hyperactivity disorder. *J. Am. Acad. Child Psychiatry* 46:894–921
- Anderson DK, Lord C, Risi S, Shulman C, Welch K, et al. 2007. Patterns of growth in verbal abilities among children with autism spectrum disorder. *J. Consult. Clin. Psychol.* 75:594–604
- Angst J, Gamma A, Baldwin DS, Ajdacic-Gross V, Rossler W. 2009. The generalized anxiety spectrum: prevalence, onset, course and outcome. *Eur. Arch. Psychiatry Clin. Neurosci.* 259:37–45
- Bandyopadhyay N, Dragalin V. 2007. Implementation of an adaptive group sequential design in a bioequivalence study. *Pharmaceut. Statist.* 6:115–22
- Bierman KL, Nix RL, Maples JJ, Murphy SA. 2006. Examining clinical judgment in an adaptive intervention design: the Fast Track program. *J. Consult. Clin. Psychol.* 74:468–81
- Brown RT, Antonuccio DO, DuPaul GJ, Fristad MA, King CA, et al. 2007. *Childhood Mental Health Disorders: Evidence Base and Contextual Factors for Psychosocial, Psychopharmacological, and Combined Interventions*. Washington, DC: Am. Psychol. Assoc.
- Cafiero J. 2005. *Meaningful Exchanges for People with Autism: An Introduction to Augmentative and Alternative Communication*. Bethesda, MD: Woodbine House
- Chow SC, Chang M. 2008. Adaptive design methods in clinical trials—a review. *Orphanet J. Rare Dis.* 3:11
- Ciraulo DA, Piechniczek-Buczek J, Iscan EN. 2003. Outcome predictors in substance use disorders. *Psychiatr. Clin. North Am.* 26:381–409
- Collins LM, Murphy SA, Bierman KL. 2004. A conceptual framework for adaptive preventive interventions. *Prev. Sci.* 5:185–96
- Crits-Christoph P, Siqueland L, Blaine J, Frank A, Luborsky L, et al. 1999. Psychosocial treatments for cocaine dependence—National Institute on Drug Abuse Collaborative Cocaine Treatment Study. *Arch. Gen. Psychiatry* 56:493–502
- Dawson R, Lavori PW. 2011. Efficient design and inference for multi-stage randomized trials of individualized treatment policies. *Biostatistics*. In press
- Dragalin V. 2006. Adaptive designs: terminology and classification. *Drug. Inf. J.* 40:425–35
- Fabiano GA, Chacko A, Pelham WE, Robb J, Walker KS, et al. 2009. A comparison of behavioral parent training programs for fathers of children with attention-deficit/hyperactivity disorder. *Behav. Ther.* 40:190–204
- Fabiano GA, Pelham WE, Waschbusch DA, Gnagy EM, Lahey BB, et al. 2006. A practical measure of impairment: psychometric properties of the impairment rating scale in samples of children with attention deficit hyperactivity disorder and two school-based samples. *J. Clin. Child Adolesc.* 35:369–85
- Hammerness P, McCarthy K, Mancuso E, Gendron C, Geller D. 2009. Atomoxetine for the treatment of attention-deficit/hyperactivity disorder in children and adolescents: a review. *Neuropsychiatr. Dis. Treat.* 5:215–26

- Hancock TB, Kaiser AP. 2006. Enhanced milieu teaching. In *Treatment of Language Disorders in Children*, ed. R McCauley, pp. 203–36. Baltimore, MD: Brookes
- Hernán MA, Brumback B, Robins JM. 2000. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology* 11:561–70
- Hernán MA, Brumback B, Robins JM. 2002. Estimating the causal effect of zidovudine on CD4 count with a marginal structural model for repeated measures. *Stat. Med.* 21:1689–709
- Honigsmann P, Fischer H, Kurmann A, Audigé L, Schüpfer G, Metzger J. 2007. Investigating the effect of intra-operative infiltration with local anaesthesia on the development of chronic postoperative pain after inguinal hernia repair: a randomized placebo controlled triple blinded and group sequential study design [NCT00484731]. *BMC Surg.* 7:22
- Hser YI, Longshore D, Anglin MD. 2007. The life course perspective on drug use: a conceptual framework for understanding drug use trajectories. *Eval. Rev.* 31:515–47
- Jones H. 2010. *Reinforcement-Based Treatment for Pregnant Drug Abusers (HOME II)*. Bethesda, MD: Natl. Inst. Health. <http://clinicaltrials.gov/ct2/show/NCT01177982?term=jones+pregnant&rank=9>
- Jones HE, O'Grady KE, Tuten M. 2011. Reinforcement-based treatment improves the maternal treatment and neonatal outcomes of pregnant patients enrolled in comprehensive care treatment. *Am. J. Addict.* 20:196–204
- Jones HE, Wong CJ, Tuten M, Stitzer ML. 2005. Reinforcement-based therapy: 12-month evaluation of an outpatient drug-free treatment for heroin abusers. *Drug Alcohol. Depend.* 79:119–28
- Joss RA, Alberto P, Bleher EA, Ludwig C, Siegenthaler P, et al. 1994. Combined-modality treatment of small-cell lung cancer: randomized comparison of three induction chemotherapies followed by maintenance chemotherapy with or without radiotherapy to the chest. *Ann. Oncol.* 5:921–28
- Kasari C. 2009a. *Developmental and Augmented Intervention for Facilitating Expressive Language (CC-NLA)*. Bethesda, MD: Natl. Inst. Health. <http://clinicaltrials.gov/ct2/show/NCT01013545?term=kasari&rank=5>
- Kasari C. 2009b. *CCNLA Developmental and Augmented Intervention for Facilitating Expressive Language*. Cent. Autism Res. Treat., Univ. Calif., Los Angeles. <http://www.semel.ucla.edu/autism/research/project/ccnia-developmental-augmented-intervention-facilitating-expressive-language>
- Kasari C, Freeman S, Paparella T. 2006. Joint attention and symbolic play in young children with autism: a randomized controlled intervention study. *J. Child Psychol. Psychiatry* 47:611–20
- Kasari C, Paparella T, Freeman S, Jahromi LB. 2008. Language outcome in autism: randomized comparison of joint attention and play interventions. *J. Consult. Clin. Psychol.* 76:125–37
- Lavori PW, Dawson R. 2000. A design for testing clinical strategies: biased adaptive within-subject randomization. *J. R. Statist. Soc. A* 163:29–38
- Lavori PW, Dawson R. 2004. Dynamic treatment regimes: practical design considerations. *Clin. Trials* 1:9–20
- Lavori PW, Dawson R, Rush AJ. 2000. Flexible treatment strategies in chronic disease: clinical and research implications. *Biol. Psychiatry* 48:605–14
- Liu-Seifert H, Osuntokun OO, Feldman PD. 2011. Factors associated with adherence to treatment with olanzapine and other atypical antipsychotic medications in patients with schizophrenia. *Compr. Psychiatry*. In press
- Longabaugh R, Zweben A, Locastro JS, Miller WR. 2005. Origins, issues and options in the development of the combined behavioral intervention. *J. Stud. Alcohol.* 15(Suppl.):179–87
- Lord C. 2000. Commentary: achievements and future directions for intervention research in communication and autism spectrum disorders. *J. Autism Dev. Disord.* 30:393–98
- Lunceford JK, Davidian M, Tsiatis AA. 2002. Estimation of survival distributions of treatment policies in two-stage randomization designs in clinical trials. *Biometrics* 58:48–57
- Marlowe DB, Festinger DS, Arabian PL, Dugosh KL, Benasutti KM, et al. 2008. Adaptive interventions in drug court: a pilot experiment. *Crim. Justice Rev.* 33:343–60
- McKay JR. 2005. Is there a case for extended interventions for alcohol and drug use disorders? *Addiction* 100:1594–610
- McKay JR. 2009. Continuing care research: what we have learned and where we are going. *J. Subst. Abuse Treat.* 36:131–45

- McKay JR, Van Horn DHA, Oslin DW, Lynch KG, Ivey M, et al. 2010. A randomized trial of extended telephone-based continuing care for alcohol dependence: within-treatment substance use outcomes. *J. Consult. Clin. Psychol.* 78:912–23
- McLellan AT. 2002. Have we evaluated addiction treatment correctly? Implications from a chronic care perspective. *Addiction* 97:249–52
- McLellan AT, Lewis DC, O'Brien CP, Kleber HD. 2000. Drug dependence, a chronic medical illness—implications for treatment, insurance, and outcomes evaluation. *J. Am. Med. Assoc.* 284:1689–95
- Miller WR, Moyers TB, Arciniega LT, DiClemente C, LoCastro J, et al. 2003. A combined behavioral intervention for treating alcohol dependence. *Alcohol Clin. Exp. Res.* 27(Suppl. 5):113A
- Murphy, SA. 2005. An experimental design for the development of adaptive treatment strategies. *Stat. Med.* 24:1455–81
- Murphy SA, Collins LM, Rush AJ. 2007a. Customizing treatment to the patient: adaptive treatment strategies. *Drug Alcohol Depend.* 88(Suppl. 2):1–3
- Murphy SA, Lynch KG, Oslin D, McKay JR, TenHave T. 2007b. Developing adaptive treatment strategies in substance abuse. *Drug Alcohol Depend.* 88(Suppl. 2):24–30
- Murphy SA, van der Laan MJ, Robins JM, CPPRG. 2001. Marginal mean models for dynamic regimes. *J. Am. Stat. Assoc.* 96:1410–23
- Nahum-Shani I, Qian M, Almirall D, Pelham WE, Gnagy B, et al. 2011a. Experimental design and primary data analysis methods for comparing adaptive interventions. Manuscript submitted
- Nahum-Shani I, Qian M, Almirall D, Pelham WE, Gnagy B, et al. 2011b. Q-Learning: a data analysis method for constructing adaptive interventions. Manuscript submitted
- Orellana L, Rotnitzky A, Robins JM. 2010. Dynamic regime marginal structural mean models for estimation of optimal dynamic treatment regimes, Part II: proofs of results. *Int. J. Biostat.* 6(2):article 9
- Oslin D. 2005. *Managing Alcoholism in People Who Do Not Respond to Naltrexone (EXTEND)*. Bethesda, MD: Natl. Inst. Health. <http://clinicaltrials.gov/ct2/show/NCT00115037?term=oslin&rank=8>
- Pelham WE, Fabiano GA. 2008. Evidence-based psychosocial treatments for attention deficit/hyperactivity disorder. *J. Clin. Child Adolesc. Psychol.* 37:184–214
- Pelham WE, Gnagy EM, Greenslade KE, Milich R. 1992. Teacher ratings of DSM-III-R symptoms for the disruptive behavior disorders. *School Psychol. Rev.* 21:285–99
- Pettinati HM, Weiss RD, Dundon W, Miller WR, Donovan D, et al. 2005. A structured approach to medical management: a psychosocial intervention to support pharmacotherapy in the treatment of alcohol dependence. *J. Stud. Alcohol Suppl.* 15:170–78
- Pettinati HM, Weiss RD, Miller WR, Donovan D, Ernst DB, Rounsaville BJ. 2004. A clinical research guide for medically trained clinicians providing pharmacotherapy as part of the treatment for alcohol dependence. In *Medical Management Treatment Manual COMBINE Monograph Series, Volume 2*, ed. ME Mattson. Bethesda, MD: Natl. Inst. Alcohol Abuse Alcohol.
- Pocock SJ. 1977. Group sequential methods in the design and analysis of clinical trials. *Biometrika* 64:191–99
- Robins JM. 1986. A new approach to causal inference in mortality studies with sustained exposure periods—application to control of the healthy worker survivor effect. *Comput. Math. Appl.* 14:1393–512
- Robins JM, Hernán MA, Brumback B. 2000. Marginal structural models and causal inference in epidemiology. *Epidemiology* 11:550–60
- Robins JM, Orellana L, Rotnitzky A. 2008. Estimation and extrapolation of optimal treatment and testing strategies. *Stat. Med.* 27(23):4678–721
- Rubin DB. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. *J. Educ. Psychol.* 66:688–701
- Rubin DB. 2007. The design versus the analysis of observational studies for causal effects: parallels with the design of randomized trials. *Stat. Med.* 26:20–36
- Rush AJ, Fava M, Wisniewski SR, Lavori PW, Trivedi MH, et al. 2004. Sequenced treatment alternatives to relieve depression (STAR*D): rationale and design. *Control. Clin. Trials* 25:119–42
- Rush AJ, Trivedi MH, Wisniewski SR, Nierenberg AA, Stewart JW, et al. 2006. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. *Am. J. Psychiatry* 163:1905–17

- SAS Inst. 2008. *SAS/STAT 9.2 User's Guide*. Cary, NC: SAS Inst.
- Schneider LS, Ismail MS, Dagerman K, Davis S, Olin J, et al. 2003. Clinical antipsychotic trials of intervention effectiveness (CATIE): Alzheimer's disease trial. *Schizophr. Bull.* 29:57–72
- Sobell MB, Sobell LC. 2000. Stepped care as a heuristic approach to the treatment of alcohol problems. *J. Consult. Clin. Psychol.* 68:573–79
- Stone RM, Berg DT, George SL, Dodge RK, Paciucci PA, et al. 1995. Granulocyte-macrophage colony-stimulating factor after initial chemotherapy for elderly patients with primary acute myelogenous leukemia. *N. Engl. J. Med.* 332:1671–77
- Stroup TS, McEvoy JP, Swartz MS, Byerly MJ, Glick ID, et al. 2003. The National Institute of Mental Health Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) project: schizophrenia trial design and protocol development. *Schizophr. Bull.* 29:15–31
- Stulz N, Gallop R, Lutz W, Wrenn GL, Crits-Christoph P. 2010. Examining differential effects of psychosocial treatments for cocaine dependence: an application of latent trajectory analyses. *Drug Alcohol. Depend.* 106:164–72
- Tannir NM. 2010. *Sequential Two-Agent Assessment in Renal Cell Carcinoma Therapy*. Bethesda, MD: Natl. Inst. Health. <http://clinicaltrials.gov/show/NCT01217931>
- Thall PF, Logothetis C, Pagliaro LC, Wen SJ, Brown MA, et al. 2007a. Adaptive therapy for androgen-independent prostate cancer: a randomized selection trial of four regimens. *J. Natl. Cancer Inst.* 99:1613–22
- Thall PF, Sung HG, Estey EH. 2002. Selecting therapeutic strategies based on efficacy and death in multicourse clinical trials. *J. Am. Stat. Assoc.* 97:29–39
- Thall PF, Wathen JK. 2005. Covariate-adjusted adaptive randomization in a sarcoma trial with multi-stage treatments. *Stat. Med.* 24:1947–64
- Thall PF, Wooten LH, Logothetis CJ, Millikan RE, Tannir NM. 2007b. Bayesian and frequentist two-stage treatment strategies based on sequential failure times subject to interval censoring. *Stat. Med.* 26:4687–702
- Tummarello D, Mari D, Graziano F, Isidori P, Cetto G, et al. 1997. A randomized, controlled phase III study of cyclophosphamide, doxorubicin, and vincristine with etoposide (CAV-E) or teniposide (CAV-T), followed by recombinant interferon-alpha maintenance therapy or observation, in small cell lung carcinoma patients with complete responses. *Cancer* 80:2222–29
- van Buuren S, Groothuis-Oudshoorn K. 2010. MICE: multivariate imputation by chained equations in R. *J. Stat. Softw.* In press
- Wahed AS, Tsiatis AA. 2004. Optimal estimator for the survival distribution and related quantities for treatment policies in two-stage randomization designs in clinical trials. *Biometrics* 60:124–33
- Wahed AS, Tsiatis AA. 2006. Semiparametric efficient estimation of survival distributions in two-stage randomisation designs in clinical trials with censored data. *Biometrika* 93:163–77
- Weisz JR, Chu BC, Polo AJ. 2004. Treatment dissemination and evidence-based practice: strengthening intervention through clinician-researcher collaboration. *Clin. Psychol. Sci. Pract.* 11:300–7
- Wigal SB, Wigal T, Schuck S, Brams M, Williamson D, et al. 2011. Academic, behavioral, and cognitive effects of OROS (R) methylphenidate on older children with attention deficit/hyperactivity disorder. *J. Child Adolesc. Psychopharmacol.* 21:121–31
- Zhou X, Liu SY, Kim ES, Herbst RS, Lee JL. 2008. Bayesian adaptive design for targeted therapy development in lung cancer—a step toward personalized medicine. *Clin. Trials* 5:181–93