Developing Adaptive Treatment Strategies In Substance Abuse Research

Susan A. Murphy
University of Michigan
&
Kevin G. Lynch
University of Pennsylvania
&
James R. McKay
University of Pennsylvania
&
David W. Oslin
University of Pennsylvania
&
Tom R. Ten Have
University of Pennsylvania

Running Title: Developing Adaptive Treatment Strategies

Contact information:
Susan Murphy
Institute for Social Research
University of Michigan
Email: samurphy@umich.edu
Abstract:
For many individuals, substance abuse possesses characteristics of chronic disorders in
that individuals experience repeated cycles of cessation and relapse; hence viewing drug
dependence as a chronic, relapsing disorder is increasingly accepted by researchers. The
development of a treatment for a chronic disorder requires consideration of the ordering
of treatments, the timing of changes in treatment, and the use of measures of response,
burden and adherence collected during treatment to make further treatment decisions.
Adaptive treatment strategies provide a vehicle through which these issues can be
addressed and thus provide a means toward improving and informing the clinical
management of chronic substance abuse disorders. The sequential multiple assignment
randomized (SMAR) trial is particularly useful in developing adaptive treatment
strategies. Simple analyses that can be used with the SMAR trial are described.
Furthermore the SMAR trial is compared with standard experimental designs.

Keywords:
Stepped Care, Individualized Care, Treatment Algorithms, Dynamic Treatment Regimes,
Experimental Design
1. Introduction

For many individuals, substance abuse possesses characteristics of chronic disorders in that individuals experience repeated cycles of cessation and relapse (Hser et al., 1997); hence viewing drug dependence as an chronic, relapsing disorder is increasingly accepted by researchers (Donovan, 1986; Brown, 1998; O'Brian and McLellan, 1996; McLellan et al., 2000; McLellan, 2002). To effectively manage the chronic nature of this disorder, substance abuse researchers have proposed strategies composed of sequences of therapies. In some cases these strategies incorporate cost and burden considerations such as the stepped care models (Sobel and Sobel, 2000; Breslin et al, 1999), which advocate beginning with a minimally intensive but effective therapy, transitioning to more intensive or other types of therapy only if indicated. Other related strategies involve increasing both intensity of therapy and encouragement to adhere when indicated (Brooner and Kidorf, 2002; Kidorf, Neufeld and Brooner, 2004). Further strategies are designed to deal with acute problems when they arise, and to return to maintenance therapy once acute problems are resolved (McKay et al., 2004). We group all of these treatment/management strategies under the umbrella term “adaptive treatment strategies” (Lavori et al., 2000; Collins et al., 2004). Adaptive treatment strategies appear in a variety of health related areas; they are employed in the treatment of depression (Lavori et al., 2000; Rush et al. 2003; called treatment algorithms), in smoking cessation programs (Prokaska et al. 2001; called an expert system) and in acute HIV infection (Altfield & Walker, 2001, called structured treatment interruptions or Albert & Yun, 2001, called treatment strategies). In all cases, these strategies individualize therapy via decision rules that specify when and how the intensity or type of therapy should change.
depending on outcomes collected during treatment such as patient response, burden, adherence to treatment, and preference. Adaptive treatment strategies mimic the clinical practice of adapting and re-evaluating treatment options based on patient progress, but do so in a more systematic and rigorous way.

In developing an adaptive treatment strategy, questions that may need to be addressed include the best sequencing of therapies when individuals are not responding and whether other time varying information such as level of self-management skill or adherence should alter the sequencing of therapies. Other possible questions concern the timing of transition from more intensive therapies to less intensive therapies or maintenance therapy and vice versa. Because comorbidities such as homelessness, depression, and HIV infection are not uncommon, questions also arise concerning the sequencing or concurrent use of adjunctive therapies targeted at other disorders. We argue that these types of questions are best addressed in a holistic manner, taking into account the available subsequent therapies rather than considering each therapy in isolation. Furthermore randomized trials should be used as much as possible in the development of adaptive treatment strategies; however most experimental designs for randomized trials have been developed with the goal of evaluating whether a therapy is better than control. The work described in this paper is focused on developing new or refining existing adaptive treatment strategies rather than evaluating whether a particular adaptive treatment strategy is better than control. After an optimized adaptive treatment strategy is developed through the experimental procedures outlined in this paper, it can be compared against treatment as usual or other control conditions in more standard randomized trials. This paper proposes a first step toward the goal of providing an
experimental approach to developing adaptive treatment strategies; in particular we propose the use of sequential multiple assignment randomized (SMAR) trials.

In Section 2, we review adaptive treatment strategies and provide two concrete examples of these strategies. Section 3 introduces SMAR trials, provides examples, and gives an intuitive discussion of how questions such as those listed above can be addressed using data from a SMAR trial. Section 4 provides a discussion of the advantages and disadvantages of the SMAR trial relative to two alternative experimental designs. In Section 5, we discuss straightforward methods for data analyses. We close with an outline of outstanding issues. It should be noted that although the focus of this paper is on the development of adaptive interventions for the treatment of addictions, much of what will be discussed applies to treatments for other chronic disorders.

2. Adaptive Treatment Strategies

Adaptive treatment strategies change the type of therapy or dose of therapy depending on time varying outcomes such as patient response, addiction severity, burden, adherence to prior therapy, etc. Altering the intensity and type of therapy is crucial for many reasons, some of which are to obtain improvement if the patient is not responding to current treatment and to reduce costs and long term participant burden when intensive treatment is unnecessary. Adaptive treatment strategies individualize therapy via decision rules that specify how the intensity or type of therapy should change depending on tailoring variables. These variables could be versions of the outcome measure, assessed during treatment (e.g., substance use), or other variables thought to predict the ultimate outcome (e.g., self-efficacy, self-help attendance, craving, and so forth). Once developed, these decision rules can be used to improve substance abuse outcomes via
effective management of patients’ substance abuse disorders. Below we describe two adaptive treatment strategies in the substance abuse setting, the first taken from Brooner and Kidorf (2002) and the second currently under evaluation by one of us (McKay).

**EXAMPLE 1**
Brooner and Kidorf, (2002) treat opioid-dependent patients with a two-component adaptive treatment strategy; the therapeutic component provides methadone and steps up/down the intensity and type of counseling sessions based on the results of urine analyses. The second component of the adaptive treatment strategy can be viewed as an “encouragement to adhere” strategy; here the strength of behavioral contingencies is stepped up and down depending on adherence to counseling sessions. Here is an example of the decision rule applied to patients currently receiving the second highest level of care (one individual counseling session per week, three to four group counseling sessions/week, weekly urine testing). The decision rule is, “If the patient both attends all of his/her scheduled counseling sessions and provides drug-free urine samples for 2 consecutive weeks then the level of care is stepped down to one individual counseling session per week and weekly urine testing.”

**EXAMPLE 2**
“Effectiveness of Extended Telephone Monitoring” (McKay) is an ongoing study of continuing care, disease management strategies for alcohol dependent subjects. The continuing care strategy consists of telephone monitoring and counseling and employs operationalized decision rules to decide when to step up treatment for alcohol dependency and when to provide information about support services for co-occurring problems. At the beginning of each telephone session, the patient responds to an 11 –item risk indicator measure that is used to determine current risk for relapse. These questions address “negative” factors including any recent use, noncompliance with medications, depressed or anxious mood, craving, low confidence in being
able to not use, and spending time with users, and “protective factors” such as attending self-help meetings and engaging in positive social activities with non-users. Patients are categorized as low, moderate, or high risk on the basis of their answers to these questions. If risk rises above the low level, treatment is intensified through a stepped care protocol that includes more frequent telephone contacts, two in-person evaluation sessions, a course of in-person relapse prevention sessions, and finally a return to standard rehabilitation treatment. Treatment can also be stepped back down, after the patient has been re-stabilized and risk level has dropped. Responses to the risk indicator questions are also used to link the patient to services for co-occurring medical, psychiatric, parenting, housing, and employment problems, through a computerized data base of social services available in city neighborhoods.

Just as the decision rules may specify a change in intensity or type of therapy due to outcomes such as patient response, addiction severity, and burden, the decision rules may also specify a change in the level or type of “encouragement to adhere” as a result of adherence to past therapy (encouragement can occur by a variety of means such as motivational therapy, behavioral contingencies, or child care and travel assistance to attend therapy). This was the case in example 1 above. “Encouragement to adhere” strategies can be expected to grow in significance to combat the difficulties with adherence prevalent in substance abuse treatment (Hser et al., 1997; Pettinati et al., 2000).

To summarize, adaptive treatment strategies are composed of operationalized decision rules that link observed patient outcomes to alterations in therapy. Both the decision rules and the patient tailoring variables are specified prior to implementation of
an adaptive treatment strategy. More discussion about adaptive treatment strategies can be found in Collins et al. (2004).

3. Sequential Multiple Assignment Randomization (SMAR)

Sequential multiple assignment randomized (SMAR) trials (Lavori & Dawson, 2000, 2003; Ten Have et al., 2003; Murphy, in press) are intended to be used in the building and refinement of adaptive treatment strategies. The SMAR trial capitalizes on the variation in patient response and adherence in order to inform the development of the decision rules and hence the adaptive treatment strategy. A number of SMAR trials have been, or are being, conducted. These include the CATIE trial for antipsychotic medications in patients with Alzheimer's (Schneider et al., 2001), STAR*D for treatment of depression (Rush et al., 2003; Lavori et al., 2001) and phase II trials at MD Anderson for treating cancer (Thall et al., 2000).

In order to make the following discussion of the SMAR trial concrete, consider the following simple example modeled after a SMAR trial currently underway by one of us (Oslo). This trial is motivated by the desire to address two questions related to the development of a strategy for treating alcohol dependent patients with the opiate antagonist Naltrexone (NTX). First, there is a potential variety of timing definitions concerning when a patient should be considered a NTX non-responder. Second, once a subject either responds to NTX (or does not respond), a variety of subsequent therapies are possible. The goal is to minimize the number of heavy drinking days over the 12 month study period.
EXAMPLE 3: Each subject is randomized twice in a sequential manner, first to a definition of nonresponse/response (first decision point) and second to a subsequent treatment (second decision point). The two possible definitions of nonresponse are: nonresponse if 2 or more heavy drinking days within a two month period or nonresponse if 5 or more heavy drinking days within a two month period. As soon as a subject meets the definition of nonresponse, the subject is randomized to either NTX + Combined Behavioral Intervention (CBI; NIH Pub. No. 04-5288, 2004) or to CBI alone. If the subject does not meet his/her assigned definition of nonresponse in the two month interval (a responder) then the subject is rerandomized to either a 6 month prescription of NTX or a 6 month prescription of NTX + Telephone Disease Management (TDM; Oslin, et al., 2003).

In a SMAR trial subjects are randomized at each decision point. In example 3, the decisions are first the timing of alterations in therapy and second the choice of the subsequent therapy for responders/nonresponders. Thus each subject is randomized twice, initially and then again once it is known if the subject is a responder or a nonresponder to initial therapy. It turns out that most of the time and indeed in all of the examples discussed in this paper, the multiple randomizations can be performed prior to initial treatment. For example in example 3, once a subject has provided consent, we can randomize the subject to the initial definition of nonresponse and then conduct a randomization to secondary treatment that will only be used if the subject responds to the initial treatment and conduct a randomization to secondary treatment that will only be used if the subject does not respond to the initial treatment. Thus if we want to conduct all randomizations pre-initial-treatment then for each subject we perform 3
randomizations of which only 2 will be used (depending on responder category). This issue is discussed further in Section 5.

To elucidate the form of the data resulting from a SMAR trial consider example 3. In figure 1, the adaptive treatment strategy,

“Strategy 1: Define nonresponse as 2 or more heavy drinking days in the two months following onset of therapy with NTX. An individual is a responder if during the entire two months, at most 1 heavy drinking day occurs. The treatment of nonresponders is augmented by CBI (NTX+CBI) and responders are provided a 6 month prescription of NTX and placed on the TDM program”

is indicated by dashed lines.

![Diagram of adaptive treatment strategy]

**Figure 1:** A NTX Study: Strategy 1 is indicated by the dashed lines.

To highlight the variety of adaptive treatment strategies consider a second strategy:
“Strategy 2: Define nonresponse as 2 or more heavy drinking days in the two months following onset of treatment with NTX. An individual is a responder if during the entire two months, at most 1 heavy drinking day occurs. Nonresponders are subsequently offered CBI without NTX and responders are provided a 6 month prescription of NTX and placed on a Telephone Disease Management (TDM) prevention program.”

This adaptive treatment strategy is highlighted by the dashed lines in Figure 2. This strategy differs from the previous adaptive treatment strategy only in regards to the treatment for nonresponders to NTX.

Figure 2: A NTX Study: Strategy 2 is indicated by the dashed lines.
Note that even though the SMAR trial randomizes subjects (for example, responders are randomized to one of two secondary therapies), the description of a particular adaptive treatment strategy does not involve randomization. In addition to the two adaptive treatment strategies highlighted above, there are 6 more adaptive treatment strategies, thus 8 in all.

We can conceptualize this SMAR trial as consisting of 8 groups of subjects, one per adaptive treatment strategy. However it is important to note that subjects belong to more than one group. For example a patient who is assigned to the nonresponse definition of 2 or more heavy drinking days and then is observed to respond belongs to both of the groups illustrated in figures 1 and 2. Thus observations on this subject can be used in assessing the effectiveness of both of these strategies.

A variety of questions that inform the development of an adaptive treatment strategy can be addressed using data from a SMAR trial. For example, “Is it useful to continue providing NTX to a nonresponding patient in addition to CBI? Does the comparison between CBI and NTX+CBI for nonresponders change depending on the definition of nonresponse to initial NTX? Does the level of adherence to NTX interact with the effect of NTX+CBI relative to CBI for nonresponders? A sufficiently strong interaction would indicate that adherence should be used in deciding which secondary treatment is best for which nonresponder. All of these questions can be addressed using comparisons between randomized groups.

Other interesting hypotheses concern two-way comparisons between adaptive treatment strategies beginning with different initial decisions. In our example, this would be a comparison of a strategy that defines nonresponse on the basis of 2 or more heavy
drinking days with a strategy that defines nonresponse on the basis of 5 or more heavy
drinking days. Alternately we might be most interested in comparisons between groups of
nonresponders assigned different subsequent treatments. For example, in the SMAR trial
represented by example 3, the primary analysis was a comparison of augmentation
(NTX+CBI) versus a switch in treatment (CBI) for nonresponders. Indeed the SMAR
trial allows us to test a variety of hypotheses, all of which use randomization to justify the
attrition that differences in outcome are due to differences between the strategies
(Green and Byar, 1984).

We view the SMAR trial as one trial in the multi-stage optimization experimental
strategy (MOST) proposed by Collins et al. (in press) (see also Onken et al., 1997).
MOST has been developed specifically for developing and refining multi-component
treatments and proposes a sequence of treatment building/refining trials so as to optimize
the treatment strategy prior to conducting a full evaluation in a confirmatory randomized
trial. According to Collins et al., (in press), the viewpoint that SMAR trails are adaptive
treatment strategy building/refining trials has a number of consequences in the
experimental design and analysis. First as in example 3, a SMAR trial need not assign
subjects to a “treatment as usual” condition or a control. As stated above once a
treatment strategy is formulated and/or refined then the strategy would be evaluated in a
randomized multi-group trial, e.g. one group assigned to the refined adaptive treatment
strategy and the other group to the appropriate standard. Second there is greater concern
with missing important differences than would be the case in a confirmatory study, thus
the Type 1 error rate might be set at a higher than .05 rate (e.g. .10) and in calculating the
sample size we might use a higher power (e.g. .85 or .9).
To minimize sample size requirements it is best to power the SMAR trial to address only a few primary strategy-building hypotheses. Even though the testing of other hypotheses will be based on randomized comparisons between groups, these hypotheses must be viewed as secondary. This is because the SMAR trial is not powered to perform these secondary analyses and thus only large effects will be detectable.

Specifying the secondary hypotheses prior to data collection allows us to attach greater confidence to significant results of these analyses (Yusuf et al., 1991; ICH, 1999). Of course as we increase the number of analyses the chance of detecting a spurious result increases (Yusuf et al., 1991; Green, 2002) but recall the purpose of the SMAR trial is to build or refine adaptive treatment strategies; once developed an adaptive treatment strategy should be evaluated in a confirmatory two group randomized trial (the other group would be assigned an appropriate alternative). In general we advocate stating all strategy-building hypotheses prior to data collection. Post hoc hypotheses and the results of associated analyses are best used to suggest hypotheses for future trials.

After a SMAR trial we may be able to proceed directly to a two group confirmatory trial in which our developed strategy is compared to an appropriate alternative or we may decide that further randomized trials are needed. Consider example 3 once more. Suppose that in preparing to conduct the trial we plan several tests. First we plan to test if TDM alone versus TDM+NTX produces differing outcomes among responders to the initial NTX. Second we plan to test if among nonresponders to initial NTX, CBI alone produces different outcomes than CBI+NTX. Third we will compare the strategies beginning with different definitions of nonresponse. Lastly we wish to test for the following interaction: “Does adherence to NTX differentiate between
nonresponders who do better on CBI alone and nonresponders who do better on CBI+NTX?” We conduct the study and find the following.

EXAMPLE 3 CONTINUED: There is no difference between providing TDM or TDM + NTX for responders and on average better outcomes result if we provide CBI+ NTX rather than CBI alone to nonresponders. Also we find that strategies that define nonresponse on the basis of 2 or more heavy drinking days result in better outcomes than strategies that use a more lenient definition of nonresponse.

We also test for the anticipated interaction and find that the interaction is sufficiently strong so that nonresponders who are adherent do better when provided only CBI. Then we can choose to eliminate further refining trials and proceed directly to a confirmatory randomized trial of an appropriate alternative versus the following adaptive treatment strategy, “Subjects begin on NTX; if the subject experiences 2 or more heavy drinking days within two months and the subject was adherent to NTX then the subject is provided CBI. If the subject was non-adherent to NTX then the subject is provided CBI+NTX. On the other hand if the subject has at most 1 heavy drinking day within two months, then the subject is provided TDM.”

Alternately suppose we carried out additional post-hoc analyses and detected a strong, unanticipated interaction (that is, this interaction was not hypothesized prior to conducting the SMAR trial). Suppose we find that among responders to the initial NTX, subjects with poor social support do better on TDM + NTX prescription than these
subjects do with TDM alone. Thus a refining trial might seek to replicate this result and/or may add a component to improve social support so that more subjects could be managed effectively with the lower cost intervention (i.e., TDM without medication). This second trial might only study responders (those who experience one or fewer heavy drinking days in the first two months on NTX). On the other hand if the post-hoc analyses detect a second strong unanticipated interaction, this time between the definition of nonresponse and whether nonresponders have better outcomes under CBI versus CBI+NTX then we might want to repeat the SMAR trial so as to replicate both of these unexpected interactions.

4. Disadvantages and Advantages of the SMAR Trial Relative to Other Experimental Designs

In this section, we contrast the SMAR trial with two other plausible experimental designs; in each case the goal is to develop and/or refine adaptive treatment strategies. To make the discussions concrete consider the following example in which we wish to develop an adaptive treatment strategy for opiate dependent subjects.

**EXAMPLE 4**: This strategy will have three steps or stages. In the initial stage we will provide methadone + group counseling. We have the option of providing standard or enhanced group counseling. The enhanced counseling includes additional work on the development of self-management skills and it is more expensive than standard counseling due to required staff training and oversight. This stage lasts for two months; during this time patients are closely monitored for nonresponse. If a patient responds (as judged by the clinical staff) then the patient moves to
a maintenance stage; in this stage we provide methadone + telephone monitoring. Our options at this stage are to whether to include brief counseling in the telephone appointments. If a patient does not respond (as judged by clinical staff) in the initial stage then the patient is provided methadone + two standard group counseling sessions per week + behavioral contingencies; in this secondary stage we have the option of providing more or less intense behavioral contingencies. The goal of the behavioral contingencies is to improve attendance at counseling and self-help groups.

**SMAR Trial:** In this trial subjects enter in the acute stage; we randomize each subject between enhanced or standard group counseling (both with methadone). If the subject responds we randomize the subject between no and low level telephone counseling. If the subject does not respond in the acute stage then we randomize between the two levels of behavioral contingencies (both include methadone and two standard group counseling sessions per week).

**Alternative Experimental Approach 1:** An alternative to a SMAR trial is to use results from the available experimental literature to decide which of the initial (acute stage) therapies to use and then conduct a randomized trial for the responders and a randomized trial for the nonresponders. Indeed, following the scenario in example 4, suppose that a review of the literature indicates that the two initial therapies appear equivalent with regards to substance use outcomes, even out to several months post-treatment. Thus we might dispense with the randomization to initial treatment (methadone + standard counseling versus methadone + enhanced counseling) and assign all subjects to methadone + standard counseling (it costs less). Next we conduct a two-group randomized trial for the responders (methadone + telephone counseling versus
methadone + telephone monitoring). And similarly we conduct a two-group randomized trial for the nonresponders (more intense behavioral contingencies versus less intense).

There are multiple advantages and disadvantages of the SMAR trial relative to experimental approach 1). First the disadvantages of the SMAR trial are the greater complexity and unfamiliarity of the SMAR trial. However consider the advantages of the SMAR trial relative to experimental approach 1. First consider advantages related to the fact that subjects who are enrolled in and who remain in the historical trial may be quite different from the subjects in the SMAR trial. Consider the issue of adherence; in many historical trials subjects were assigned a fixed treatment, that is, there were no options besides non-adherence for subjects who were not improving. This often leads to higher than expected drop-out or non-adherence. This is particularly the case in longer studies where continuing treatments that are ineffective is likely associated with high non-adherence especially if the subject doesn’t know if they are receiving treatment such as in a double bind study. As a result the subjects who remained in the historical trial may be quite different from the subjects that remain in a SMAR trial, which by design provides alternates for non-improving subjects. This issue is part of the greater issue involving cohort effects and occurs also with the use of historical controls (Byar, 1980; Green and Byar, 1984; Byar, 1988; Byar 1991). The problem is that subjects in historical trials may differ from current subjects in unknown ways. Thus an advantage of the SMAR trial over experimental approach 1) is that we avoid the general issue of cohort effects.

A second advantage of the SMAR trial vis-a-vi the alternative experimental approach 1) is that data from a SMAR trial permits us to examine if there are synergistic effects in a sequence of therapies resulting in improved longer term outcomes. Again
consider example 4 and again suppose a review of historical trials involving the acute stage therapies appear equivalent with regards to substance use outcomes. The two therapies (methadone + weekly standard group counseling versus methadone + weekly enhanced group counseling) are likely to have been evaluated in a setting in which the subjects did not receive a maintenance therapy or if they did, the maintenance therapy differed from those considered here. So the evaluation of the two initial therapies is actually a comparison of methadone + standard counseling followed by a mixture of maintenance therapies as available in the community for responders versus a comparison of methadone + enhanced counseling followed by a mixture of maintenance therapies as available in the community for responders. This comparison may say very little about the usefulness of methadone + enhanced counseling when followed (if a responder) by telephone counseling. Thus it is plausible that the sequence of methadone + enhanced counseling followed (if a responder) by telephone counseling may be more effective than any other sequence. For example, the self management skills learned early may enable the subject to gain from the minimal counseling offered with the telephone monitoring. A SMAR trial has the potential to detect the advantage of the sequence whereas the other experimental approach does not.

**Alternative Experimental Approach 2**: A second alternative experimental design to the SMAR trial is to survey the available literature, use clinical experience and employ a variety of established principles (e.g. start patients on the least intensive treatment as in the stepped care model (Sobel & Sobell, 2000) or use the principles espoused by Collins, et a., (2004)) to formulate the decision rules underlying the adaptive treatment strategy. Further guidance might be obtained from existing clinical practice guidelines, such as
those of ASAM (ASAM, 1996). Alternately we might use behavioral, psychosocial and/or biological theories (for an example, see Velicer et al., 1993) to formulate the decision rules. Once the decision rules and associated therapies are selected, we conduct a two group confirmatory randomized trial with one group assigned the formulated adaptive treatment strategy and the other group assigned the best alternative treatment available (or other appropriate alternative).

An advantage of the experimental approach 2) over the SMAR trial is related to the sample size. If $n$ is the total number of available subjects, then in this approach $\frac{1}{2}n$ subjects are in each group; in contrast the SMAR trial for example 4 has only $\frac{1}{4}n$ subjects in a group corresponding to a particular strategy. This is the price the SMAR trial pays for allowing us to investigate a variety of strategies.

An apparent advantage of formulating an adaptive treatment strategy and then contrasting the formulated strategy versus an appropriate alternative is speed. However this advantage is only apparent as it hinges on a significant effect size resulting from an analysis of the two-group confirmatory randomized trial. If this occurs we have an effective, already formulated, adaptive treatment strategy. Hopefully our clinical experience, theories, etcetera are sufficiently accurate so that there are no unexpected negative interactions between the decisions comprising the adaptive treatment strategy. For example we assume that the burden imposed by the selected initial treatment will not interfere with adherence to the selected subsequent treatment or, in the case of example 3 (involving consideration of when to alter treatment), the chosen definition of responder/nonresponder is not poorly matched with the selected subsequent treatments.
for responders and nonresponders (e.g. a responder according to our definition is ready for the selected maintenance treatment).

A disadvantage of formulating an adaptive treatment strategy and then conducting a two group randomized trial as compared to a SMAR trial is that in the former once we have compared the two groups we do not know which decision rules comprising the adaptive treatment strategy are necessary. We have not looked into the “black box.” We do not know if any of the selections we have made in formulating the strategy are inert as compared to other possible treatments or decisions. For example we can not test if it is best to use adherence level in deciding which subsequent treatment to provide nonresponders (we provided only one subsequent treatment).

**Additional Comparisons:** Recently scientists have been studying the combination of treatments in two – four group randomized trials, particularly for subjects with co-morbidities. Consider the trial described in Oslin (in press) for subjects with alcohol dependence and depression. In this randomized trial one group was assigned a combination of three treatments (psychotherapy + sertraline + naltrexone) and the other group was assigned a combination of two treatments (psychotherapy + sertraline + placebo). The implicit assumption motivating this trial was that two drugs is always better than one; this ignores the possibility that a sequence may actually result in better outcomes either because for some individuals one medication may be better alone rather than two and because treating one of the disorders may lead to improvement in the remaining disorder without need for further medication. An alternative SMAR trial would start with psychotherapy + sertraline or another medication for depression and then randomize to an augmentation with naltrexone only if drinking continued or craving was
high. This provides a different, more focused assessment of the utility of naltrexone in that only subjects whose alcohol dependence is resistant may be randomized to naltrexone.

All of the trials discussed above can be subject to further post-hoc analyses. These include dose-response analyses in which the dose is the amount of treatment received (for example the number of counseling sessions attended). These analyses may lend credibility to one or more explanations for the observed effects sizes. Note however these analyses do not depend on randomization and hence are subject to bias or confounding (Green and Byar, 1984).

Also all of the experimental approaches (including the SMAR trial) discussed above can be following by further randomized trials. Indeed, since the SMAR trial is strategy developing/refining trial, it must eventually be followed by the standard two group randomized trial, with one group assigned the refined adaptive treatment strategy and the other group assigned an appropriate alternative (best available alternative or standard care, etc.). Also all of these experimental designs are subject to each of the difficulties inherent in the conduct of a randomized trial (subjects attrit from the trial, subjects withdraw consent, incomplete assessments, implementation issues etc.).

5. Further Statistical Considerations

Below we discuss both the design and analysis of the SMAR trial in greater detail.

**Design Considerations:** As discussed previously often we are able to conduct all of the multiple randomizations prior to the onset of treatment. Consider example 4. In this case at the beginning of the trial we can conduct three randomizations. The first
randomization is to one of the two initial treatments (methadone + weekly standard group counseling versus methadone + weekly enhanced group counseling). At the same time we conduct a randomization that will be used if the subject responds to initial treatment (randomize between methadone + telephone counseling versus methadone + telephone monitoring) and conduct a randomization that will be used if the subject does not respond to initial treatment (randomize to more intense behavioral contingencies versus less intense). Since a subject can either respond or not respond to initial treatment only one of the two latter randomizations will be used. This approach is equivalent to randomizing each subject to one of the eight possible treatment strategies.

Recall that in classical randomized control trials, one stratifies the randomization by variables that are likely to be strongly predictive of the response (for example, presence of co-morbidities) (Piantadosi, 1997, p. 208-211). Similarly it is best to stratify randomization of the responders and nonresponders if we believe that there is a variable observed during initial treatment that will be strongly predictive of response to secondary treatment; adherence to initial treatment might be one such variable. In this case, we randomize the nonadhering nonresponders separately from the adhering nonresponders. In order to stratify the randomization of the secondary treatments by an outcome variable measured during initial treatment, we must delay the randomization. Why go to the trouble to stratify randomization by adherence? We know that on average, the groups associated with different secondary treatments will not differ in terms of adherence to prior treatment. However this is only on average; in our study a random imbalance between groups assigned different secondary treatments may occur. If imbalance does occur then the observed difference in outcome to the secondary treatments maybe due in
part to the imbalance rather than due to a difference between secondary treatments. Consider example 4. Suppose that approximately 30% of the nonresponders to the initial treatment were nonadherent. Then we expect that due to the randomization, 30% of the individuals assigned high intensity behavioral contingencies will have been nonadherent to initial treatment. However, even with a sample size as high as 50 subjects, the chance that 50% or more nonadherent subjects will be randomized to the high intensity behavioral contingencies group is .84. Suppose that individuals who are nonadherent to initial treatment are more likely to be nonadherent to secondary treatment (and thus suffer poorer outcomes). Then any difference in outcomes between the high and low intensity behavioral contingencies groups may be due an imbalance in the proportion of subjects who have been nonadherent to past treatment. Stratifying the randomizations eliminates this problem.

**Analysis Considerations:** In this exposition we assume that the primary outcome, Y, is continuous. In example 3, Y might be (the logarithm of) the percent heavy drinking days. Many analyses are familiar. Consider the comparison of any two adaptive treatment strategies each beginning with a different initial decision. Consider example 3 once again. For example we might be interested in comparing the two different definitions of nonresponse. One way to do this is to compare the mean outcome to Strategy 1 stated previously with the mean outcome to:

“Strategy 3: Define nonresponse as 5 or more heavy drinking days in the two months following onset of treatment with NTX. An individual is a responder if during the entire two months, at most 4 heavy drinking days occur. The treatment of nonresponders is augmented by CBI (NTX+CBI) and responders are provided a 6 month prescription of NTX and placed on a Telephone Disease Management (TDM) prevention program.”
These two strategies differ only in their definition of nonresponse. To compare the two strategies we conduct a two group test (a t-test). The first group consists of all subjects with treatment patterns consistent with strategy 1 and the second group consists of all subjects with treatment patterns consistent with strategy 3. Recall that ¼ of the subjects belong to each group thus each two group comparison uses ½ of the subjects. See Murphy (in press) for further discussion of these two group comparisons.

If we wish to compare two adaptive treatment strategies both with the same initial decision then essentially we are comparing outcomes to subsequent decisions. Here an analysis of covariance or regression including both initial and subsequent decisions as covariates can be employed. Again we discuss this in the context of example 3. Denote the initial randomization by $A_1$ (equal to 1 if randomized to the 2-or-more-heavy-drinking-days definition of nonresponse and equal to 0 otherwise), denote the second randomization for nonresponders by $A_{2, NR}$ (equal to 1 if NTX +CBI and equal to 0 if CBI only), denote the second randomization for responders by $A_{2, R}$ (equal to 1 if prescription +TDM and equal to 0 if prescription only). Similarly denote responder by $R=1$ and nonresponder by $R=0$. The following regression formula for the mean outcome, $Y$, can be used to test hypotheses concerning the secondary treatments for both responders and nonresponders to initial NTX:

$$
\beta_0 + \beta_1 A_1 + \beta_2 R + \beta_3 A_{2,R} + \beta_4 (1-R)A_{2, NR}
$$

For example, a test of $\beta_4 = 0$ is a test of the hypothesis that providing CBI+NTX to nonresponders results in a different mean $Y$ than providing CBI only to the nonresponders. A natural interesting question in the context of example 3 is whether the
level of adherence to the initial provision of NTX differentiates between nonresponder who do better on CBI alone as compared to nonresponders who do better on CBI+ NTX. To address this question we can include terms and interactions involving adherence level to initial NTX in the above regression.

This regression formulation permits us to address a variety of secondary hypotheses. These regressions are standard in terms of implementation. There is however one difference between interpretation of these regressions and the standard regression with which most researchers are familiar. This difference lies in the interpretation of $\beta_1$, the regression coefficient of the first decision; this regression coefficient should *not* be used to assess the significance of the first decision (timing decision in example 3; initial treatment in example 4). Because we have included a variable that has been affected by A$_1$, (that is the response indicator, R) in our regression formula, the regression coefficient of A$_1$ measures a complex combination of both a) the effect of A$_1$ that effects Y through other causal pathways than via R and b) a correlation between A$_1$ and Y induced by the fact that there are many potential unobserved causes of both response to NTX and drinking outcome Y. This is one of the reasons why clinical trialists avoid including post treatment outcomes as independent variables in their analyses of clinical trial data (Rochon, 1999; ICH, 1999). In the case that a contrast between different initial decisions is desired, we recommend comparing strategies with different initial decisions as discussed above.

Exciting new methods with an emphasis on improving power are in development (Murphy, 2003; Robins, 2004; Blatt et al., 2004) and which permit a greater variety of outcome variables (Thall, 2000) or time to an event outcomes (Lunceford et al., 2002).
Many of these methods generalize well to the analysis of trials with more than two randomizations per subject (e.g. as in STAR*D in which up to 4 randomizations occur).

6. Discussion

An ongoing challenge to the development of adaptive treatment strategies is the fact that these strategies are most naturally multi-component treatments, with potential components including medications, psychosocial therapies, and interventions to improve adherence (e.g. motivational interviewing). Moreover, each component may be delivered through several possible formats (e.g., group therapy versus individual therapy, telephone vs. in-person, and so forth). Additionally in adaptive treatment strategies there are also timing components (how soon should one alter treatment?). As a result the refining and building of adaptive treatment strategies based on experimental evidence appears difficult. The multi-stage optimization (MOST) design as proposed by Collins et al. (in press) may be useful in generalizing the SMAR design; the MOST design has been developed specifically for developing and refining multi-component treatments.

Many clinicians who help patients manage their chronic illness find results from standard trials to be insufficiently informative. For example, clinical trials that evaluate the usefulness of a fixed treatment (e.g. NTX for 6 months) do not deal with issues that clinicians must face such as what to do if the patient does not appear to be responding. Also researchers who run clinical trials are often frustrated with high rates of nonadherence. Nonadherence can result if a patient has improved and is thus ready for a reduction in intensity of treatment; similarly nonadherence can result if a patient is not improving and there are no alternatives within the trial. To the extent that these are true,
SMAR trials hold the promise of improving adherence and thus provide an attractive alternative to clinical trials of fixed treatments.

Given that for many individuals drug dependence is chronic, we believe that the SMAR trial provides an alternative experimental design that is well suited toward addressing the sequencing and timing questions that arise in the development of adaptive treatment strategies. The process described in this paper is designed to develop an optimal adaptive treatment strategy, through a systematic, experimental test of different treatment alternatives. We believe that the use of these trials will improve our ability to develop the strongest possible adaptive strategies.

Acknowledgements:

We acknowledge support from NIH grants: R21 DA019800, K02 DA15674, P50 DA10075 (Murphy); DA-005186 (Lynch); K08 MH01599, 5P30MH52129; NIAAA 1R01AA014851, and the Department of Veterans Affairs MERP Award (Oslin); K02-DA00361, R01 AA14850 (McKay); R01 MH61892, P30 MH066270 (Ten Have).
References:


