Published in final edited form as:

J Subst Abuse Treat. 2009 December ; 37(4): 388–397. doi:10.1016/j.jsat.2009.05.004.

Choosing a Control Group in Effectiveness Trials of Behavioral Drug Abuse Treatments

Gregory S. Brigham, Ph.D.,

Maryhaven Research Institute Maryhaven Columbus, Ohio 43207

Daniel J. Feaster, Ph.D.,

Department of Epidemiology and Public Health Miller School of Medicine University of Miami Miami, Florida 33136

Paul G. Wakim, Ph.D., and

Center for Clinical Trials Network National Institute on Drug Abuse Washington D.C. 20892

Catherine L. Dempsey, Ph.D.

School of Medicine University of Arizona Tucson, Arizona 85721

Abstract

Effectiveness trials are an important step in the scientific process of developing and evaluating behavioral treatments. The focus on effectiveness research presents a different set of requirements on the research design when compared with efficacy studies. The choice of a control condition has many implications for a clinical trial's internal and external validity. The purpose of this manuscript is to provide a discussion of the issues involved in choosing a control group for effectiveness trials of behavioral interventions in substance abuse treatment. The authors provide a description of four trial designs and a discussion of the advantages and disadvantages of each.

Keywords

drug; abuse; clinical; trial; control; design

1.0 Introduction

Historically, the safety and efficacy of new treatments for drug abuse have been demonstrated primarily in specialized research settings, with somewhat restricted patient populations. The community-based treatment programs (CTPs) that provide the majority of drug abuse treatment services in the United States have faced many obstacles to implementing research in clinical settings resulting in limited bi-directional exchange of knowledge between researchers and clinical practitioners. This gap between research and practice has been an obstacle to improving drug abuse treatment, as many promising, evidence-based treatments have not been widely disseminated to, and adopted by, CTPs.

1.1 Stages of Science

The successful development of a new behavioral treatment can be conceptualized as progressing through several stages of science, to dissemination and, finally adoption. The

failure in this progression resulting in a gap between research and practice in the substance abuse field has been well documented (Lamb, Greenlick & McCarty, 1998). The stage model of behavioral therapies research, which has been influential in the design of behavioral therapy treatments for drug dependence, has also been well described and discussed in the literature (Onken, Blaine, & Battjes, 1997; Rounsaville, Carroll, & Onken, 2001). In this three-stage model, Stage I focuses on producing the minimal elements required to test the efficacy of a behavioral treatment. This includes generating new treatments, defining and documenting the procedures involved (often in the form of treatment manuals), developing measures of fidelity to evaluate the quality of treatment delivery, using preliminary findings to improve the therapies, and pilot testing. Promising treatments developed in Stage I may progress to Stage II and III where the focus is on testing efficacy and effectiveness respectively. The goals of Stage II are often achieved by comparing treatments in well-controlled single-site randomized clinical trials (RCTs). These trials adhere closely to procedures established in earlier research and provide a test of the treatment against other treatments of known efficacy or placebos. Considerable attention is given in the choice of comparison groups to maximize the internal validity of the design and therefore increase confidence that results are due to the therapy being tested and not other non-specific factors. If a treatment's efficacy is supported by well-controlled RCTs it may go on to be tested for effectiveness in Stage III. This stage of research seeks to evaluate the treatment in conditions more closely resembling those in which it will ultimately be adopted. Stage III research seeks to answer questions such as: will the treatment work in community settings with substance abuse treatment providers? What type of training and supervision is required to deliver a safe and effective treatment? What are the costs and benefits of the treatment and is it feasible and sustainable? Does the new treatment provide a needed alternative, a more cost effective alternative or, improvement over current practice? To address these questions investigative teams may simplify the treatment procedures, supervision and, training to make them more feasible in community treatment settings. It is often desirable to conduct effectiveness trials in multiple settings to provide an opportunity to test the intervention in a diverse range of settings which is more representative of the environment in which a successful intervention will ultimately be adopted. A number of research designs may be chosen to address the questions presented in Stage III. This paper provides guidance on choosing a control condition when investigative teams use a within-site group comparison design in a randomized clinical trial. The NIDA Clinical Trials Network [<http://www.nida.nih.gov/CTN/index.htm>], formed partly in response to a IOM report highlight the need for a clinical trials network to conduct multi-site effectiveness trials (Lamb, Greenlick & McCarty, 1998), provides an opportunity to look at the strengths and weaknesses of a number of Stage III clinical trial designs. While all of the examples used are studies conducted in a multi-site design the majority of the strengths and weaknesses of these designs apply to both multi-site and single site designs. As in earlier stages, in Stage III the choice of a control group has many implications for the internal and external validity of the study. From Stage II efficacy to Stage III effectiveness trials, the research question changes and the design must change to reflect new priorities.

1.2 Control Groups

In efficacy trials, the choice of control group is frequently dictated by concerns of maximizing internal validity. For example, a trial which randomly assigns participants to either a no-treatment control or an experimental treatment will be able to determine whether this treatment changes the trajectory of an outcome variable, with randomization ruling out the possibility that the change may have been a result of history, participant selection, testing or general time effects (Nathan, Stuart, & Dolan, 2003). Particularly in behavioral research, there is an additional concern that the level of contact alone, regardless of the content of the contact may have a measurable effect. To control for this possibility, an attention-placebo condition may be utilized as one of the experimental groups. If the experimental treatment group improves

relative to the attention-placebo group, there are stronger grounds for concluding that it was because of the particular content of the intervention and not the added attention. In some investigations a researcher might employ both a no-treatment and an attention-placebo group. Note that the use of these two different control groups allows the researcher to answer two different questions. In the case of the no-treatment control, the question is whether one can change the natural course of some phenomenon by applying a particular intervention. In the case of the attention-placebo, the question is whether the particular ingredients of the intervention and not the contact aspects of the intervention are changing the natural course.

In effectiveness trials, these two questions have normally been answered in prior efficacy research. Effectiveness research tends to focus more on external validity and attempts to establish the generalizability of an intervention's effects. In addition, because effectiveness trials are conducted in clinical settings in which patients are requesting services, a no-treatment control or an attention-placebo control maybe ethically untenable. Nevertheless, in effectiveness trials, internal validity is still a concern, and control groups are necessary.

The questions asked in an effectiveness trial will necessitate different control groups. With respect to the drug abuse treatment research, some interventions are comprehensive drug treatments whereas others are specific components to treatment or ancillary services. In the case of comprehensive treatments there can be differences in the amount of tailoring of the intervention necessary to maximize the feasibility of implementation in the clinical setting. For interventions where the feasibility of implementation in clinical settings is questioned and as a result the experimental treatment is significantly altered to enhance feasibility, the use of a standardized treatment control group may be appropriate. For example if the therapy is changed from a individual format to a group or the dose of treatment is substantially reduced to make the treatment more feasible in a community setting the investigators may question the degree to which the previous evaluation results for the longer individual treatment still apply and therefore may prefer a standard control condition to provide an element of efficacy evaluation within the effectiveness trial (see design 3 below for greater detail). Where the feasibility of implementation is not questioned and the intervention is not substantially altered, but the effectiveness of interventions across locations and clinics is of interest, a comparison to the treatments in use at each location (i.e. treatment-as-usual; TAU) is more appropriate. The issue here is a matter of degree, if the modifications are extensive, the applicability of the efficacy results lessens and the effectiveness trial shifts towards the efficacy side of the spectrum.

Sometimes new interventions that consist of components added to treatment are conceptualized as being a potentially good addition to TAU. In such cases, the question remains of whether the addition of the new component or ancillary service should be balanced by an attention control component or ancillary service in the comparison group.

2.0 Methods

Because of the many questions that may be posed in an effectiveness trial, there is no single best approach to choosing the control condition. This paper uses examples from the CTN to present four possible design options for randomized clinical trials conducted in community treatment programs involving treatment-seeking individuals:

Design # 1	New intervention vs. TAU
Design # 2	New intervention + TAU vs. alternative intervention + TAU
Design # 3	New intervention + TAU vs. alternative intervention + TAU
Design # 4	New intervention vs. standardized control treatment

Each of these design options is described below followed by an example from the CTN and a discussion of advantages and disadvantages from analytic (research design), practical

(treatment provider), and ethical (participant) perspectives. Because our current focus is to isolate and discuss the impact of the control group all design options in this paper are two-arm trials presented as Treatment A versus Treatment B, where Treatment A represents the new experimental intervention being tested and Treatment B the control treatment. The four design options are arranged in an order that proceeds from stronger generalizability to stronger internal validity. Table 1 clearly shows the shifting pattern of compromises and gains as the focus is shifted from external validity to internal validity (see Table 1).

Drug abuse is a complex condition affecting mind and body and, in response drug abuse treatments are complex and generally consist of multiple components including behavioral and pharmacological interventions. Whereas general principles of effective treatment derived from research have been published (NIDA, 1999) specific behavioral treatment guidelines are yet to be developed. Because no widely accepted standard of care has been established in drug abuse treatment in this paper we discuss treatment as usual as the standard practice of the community treatment providers. It is this community practice (TAU) to which the new treatments are compared to evaluate effectiveness and increase the likelihood of adoption in the substance abuse treatment field.

3.0 Results

3.1 Design #1 New Intervention vs. Treatment-As-Usual

The CTN study evaluating Brief Strategic Family Therapy (BSFT) provides an example of the New Intervention vs. TAU design. This effectiveness study randomly assigns 480 drug-using adolescents to BSFT or TAU in 8 community treatment sites. This study was designed to compare the new treatment (BSFT) to TAU at each location (Feaster, Robbins, Horigian & Szapocznik, 2004, Robbins, Szapocznik, Horigian, Feaster, Puccinelli, et al. 2009). As such, every effort was made to minimize any effect of the trial on TAU. The trial was designed to account for variability across sites with the variability in treatment effects across sites considered to be random effects.

BSFT is a complete drug abuse treatment for adolescents, not a component of treatment, that targets interactions in the family system that have been shown to influence adolescent drug abuse (Szapocznik, Hervis, & Schwartz, 2003). This treatment consists of 12 to 16 sessions of 1 to 1.5 hours over 4 to 6-months. Therapists may also conduct up to 8 “booster sessions” after the 12-16 sessions with cases that relapse, have significant clinical problems, or to reinforce gains.

Whereas the BSFT protocol's choice of TAU as a control group and estimation of site variability in treatment effects were implemented to maximize external validity, the investigative team took steps to minimize the risk of compromising internal validity for external validity. Two aspects of the protocol design are instructive as to how this can be accomplished or considered in trial design. First, the protocol design team reviewed the existing TAUs for adolescent drug abuse treatment and concluded that nearly all community treatment providers' adolescent services involved at least as much, and in some case more, intensity and duration as does BSFT. This increases confidence that a positive finding for BSFT would not be caused by a planned imbalance in non-specific factors. This does not preclude the possibility that BSFT may have a greater propensity to engage and keep participants in treatment, which is a secondary hypothesis of the protocol. Second, all therapists for this protocol were drawn from the pool of available TAU therapists and randomly assigned to either TAU or BSFT. This randomization of therapist ensures that pre-existing skill level of therapist is not a factor in any treatment effect (other than the additional skill imparted by completing BSFT training).

3.1.1 Positive Aspects of Design

P1 Ecological Validity: The design has strong ecological validity. It does not require the creation of an artificial control group that may be unrelated to current practice. From the CTP's perspective, the results of the trial should give a robust answer as to whether the new treatment is an improvement over their current treatment, particularly if the number of participants per site is large enough to give stable estimates of effect size by site.

P2 Generalizability: This design allows for broad generalizability of trial results. This design should be able to answer the question of whether this new treatment would be expected to have a better outcome at any location within the scope of sites selected. Technically, this depends on the sample of CTPs and the method of handling site variability. First, if the sample of CTPs is representative of the population of CTPs, then inference is to any CTP. This speaks directly to the question: Will more people get better if we increase the implementation of this intervention across the country? If the sample of CTPs is not representative of the population, then the inference is to any CTP like the sample of CTPs in the study. Second, this type of inference (without reference to a specific location) is dependent on site variability in treatment effects being estimated as a random effect.

P3 Comparison with Different TAUs: This design yields more information than designs that compare a new intervention to a standardized control. It allows for the estimation of differences in outcomes across all the various TAUs within the sample. The pattern of these differences may yield important clinical understanding, particularly if examined with respect to the components and characteristics of the TAUs in the sample. It may even identify TAUs that could be more effective than the new intervention.

3.1.2 Negative Aspects of Design

N1 Drop out of New Intervention into TAU (Control): Because this effectiveness trial would be implemented with treatment-seeking individuals, if participants drop out of the New Intervention, they simply receive the TAU provided at the site where they are enrolled which in this design is the control condition. If the trial is conducted in multiple sites, as in our example, the study treatment has the potential to have this interaction in multiple different treatment environments. This could occur due to perceived negative aspects of the new treatment or higher desirability of the TAU. For example a family treatment that goes into the home to deliver services may be considered more invasive by some participants. A treatment as usual condition that offers some form of incentives or less burden for compliance might be perceived as more desirable.

N2 Increased Sample Due to Variability in TAU: The added variability in TAU will require a larger number of participants, and a larger number of sites in multi-site designs, than will designs that use a standardized control group.

N3 Potential for Mixed Results: When this design is implemented in a multi-site design there is the possibility of getting a "mixed" answer. The new intervention may on average be an improvement over TAU; it may also be less effective than TAU at some sites (see Figure 1 below). However, this would be important clinical information, which may require examination and possibly even additional study to determine whether this was a "random" occurrence, or predictable from some aspect of management or clinical program at the site. This possibility highlights the value of collecting information about the sites and the programs that they administer.

N4 Differences in Factors Not Specific to Treatment: TAU may differ from the new intervention in nonspecific factors, such as the amount of attention given to participants and the quality of implementation, which may compromise the internal validity and draw into question any conclusions based on the group comparisons. In this design the availability of prior efficacy can increase confidence that the observed effects are due to the specific treatment.

N5 Observed TAU is Not True TAU: Treatment as usual is rarely true TAU in effectiveness behavioral trials, simply because it is being observed, described and assessed. Whereas the direction of bias is not clear, to the extent that these procedures cause TAU therapists to “be at their best,” the bias will be conservative with respect to finding an effect for the new treatment.

N6 Potential Overlap between the Two Groups: If TAU includes significant components that are related to the new intervention, the observed effect size will be lessened. This is a problem common to comparing new interventions to any type of active treatment control.

Figure 1 below presents one potential set of findings from this type of control group, in a multi-site design where the effect of TAU may vary greatly. In this example, the overall effect of the new intervention is positive when compared to TAU. This overall finding is consistent with the results in sites A and C. However, the TAU condition was superior in site B.

3.2 Design #2 New Intervention + TAU vs. TAU

There are situations where it may be desirable to add a new therapeutic regimen to TAU. For example the new therapeutic regimen may deal with a different dimension than TAU; however, without TAU, the new therapeutic regimen may not be effective. One such example involves opiate-abusing women with young children. These women are often poor caregivers for their children, who are at elevated risk for developing substance abuse and other behavioral problems themselves (Johnson & Leff, 1999). Thus, when these mothers are stably enrolled in a methadone maintenance (MM) program, it is desirable to augment MM with a program, such as the Strengthening Families Program (SFP), which provides parenting and family communication skills training. The MM provides an opportunity to access these women during a time when they may be open to receiving services and are in regular contact with a treatment agency. However, without the stable context of MM treatment, there may be little reason to implement SFP.

Another reason for adding a new intervention to TAU is that the new intervention may enhance TAU to provide a more beneficial treatment. For example, research suggests that there are high rates of unemployment in drug dependent individuals and that drug dependent patients express strong interest in vocational services (French, Dennis, McDougal, et al., 1993). A CTN study designed to evaluate the addition of a Job Seekers Workshop developed specifically for drug dependent patients (Hall, Loeb, & Norton, 1977) provides an example of the New Intervention + TAU versus TAU design. This multi-site effectiveness trial was designed to randomize 624 treatment-seeking individuals, 312 in methadone maintenance programs and 312 in drug free outpatient treatment programs, to receive either standard treatment (TAU) or standard treatment plus Job Seekers' Workshop (Svikis, 2004).. The primary outcomes included a) time to either a new taxed job or enrollment in a job-training program during a 3-month follow-up period and b) total hours worked in a taxed income job and/or hours accumulated in job skills training during the 3-month follow-up period.

3.2.1 Positive Aspects of Design—Several previously described positives also apply to this design including: P1 Ecological Validity, P2 Generalizability and, P3 Comparison with Different TAUs (see Table 1). This design also offers several additional advantages:

P4 Synergistic Interaction of Treatments: The new intervention may interact with the TAU to provide a more beneficial treatment. The new intervention may catalyze the effectiveness of TAU in addition to whatever additive effects each may have (Borkovec, 1993). This design may demonstrate whether the combination of a supplemental treatment and TAU is superior to the sum of its components.

P5 Reduces Impact of Differences in TAUs: In multi-site designs, when randomization is done within CTPs, this design considerably reduces the impact of variability in the effectiveness of TAUs across participating CTPs if there is no synergistic interaction of treatments (P4). Because the same TAU is included in both conditions at each site, the added effectiveness of the new intervention is being analyzed regardless of the effectiveness of TAU. Thus, the effectiveness of TAU cancels out (see Figure 2 below that shows very different TAUs across sites).

P6 Evaluates Add-on to TAU: This design answers a simple and practical question of whether an add-on to what clinics currently do is worthwhile. Although it does not protect against the possibility that any non-specific increase in therapeutic contact and quality of supervision would improve outcome. Assuming prior efficacy trials have established specific effects for the intervention, this is less of a concern in effectiveness trials.

P7 All Participants Receive Established TAU: Since both interventions are designed to provide benefit to participants and include TAU, there is no chance that an established treatment will be given up for one of unknown utility.

3.2.2 Negative Aspects of Design:

Numerous negatives described above also apply to this design including: N4 Differences in Factors Not Specific to Treatment, N5 Observed TAU is not True TAU, and N6 Potential Overlap between the Two Groups (see Table 1). This design also introduces additional negatives:

N7 Limited to Testing Add-On Interventions: The intervention being studied must be a type of treatment that can be added to TAU.

N8 Site Variability Due to Synergistic Treatment Interactions: In multi-site designs the new intervention may catalyze the effectiveness of TAU in addition to whatever additive effect each may have (Borkovec, 1993). If this catalysis is constant across sites/TAUs, then this is a positive aspect of the design as described above in Positive 4. If, however, the new intervention catalyzes the effectiveness of the TAUs in different ways depending on the TAU, then it is no longer positive and may complicate interpretation of the results.

3.3 Design #3 New Intervention + TAU vs. Alternative Intervention + TAU

In this design, both the new and alternative interventions are added to Treatment-As-Usual (TAU). All participants receive the usual care provided by the participating clinic in addition to either the new or alternative intervention.

A CTN study designed to test the effectiveness of a treatment for women with substance abuse and symptoms of post traumatic stress provides an example of this design. The primary objective of this trial was to implement and evaluate the effectiveness of Seeking Safety (SS), in comparison to a control intervention, Women's Health Education (WHE). The Seeking Safety intervention is a cognitive behavioral therapy designed for women. It addresses both substance abuse and symptoms related to trauma. In this trial, it is conducted in an open-ended group format, for 12 sessions. The WHE intervention is matched to SS for group format,

frequency and duration of sessions and, provides equivalent facilitator attention, expectancy of benefit and, issue oriented focus, but does not provide the theory-driven techniques of Seeking Safety or psycho-education specific to substance abuse and PTSD.

Participants in the study were substance-abusing women seeking treatment who also met DSM-IV (APA, 1994) criteria for Post-Traumatic Stress Disorder (PTSD) or Sub-threshold Post-Traumatic Stress Disorder (SPTSD). This randomized, parallel group effectiveness trial planned for implementation in eight separate clinics with 480 women randomized to one of the two study treatment groups in addition to TAU. The primary outcomes include changes in substance abuse and severity of PTSD symptoms. Participants at each of the CTPs were enrolled in outpatient treatment for drug abuse. These programs consist of a variety of individual and group treatment components. The precise nature of the services provided by each CTP for either drug abuse or PTSD is unknown and is not a major factor in determining the site's participation in the study (see Figure 3 below that illustrates very different TAUs across sites).

3.3.1 Positive Aspects of Design—Several of the positives described above also apply to this design including: P5 Reduces Impact of Differences in TAUs, P6 Evaluates Add-on to TAU and, P7 All Participants Receive Established TAU (see Table 1). In addition this design offers the following additional positives.

P8 Control of factors not specific to intervention: When the control condition (Alternative Treatment) is chosen or developed to match the New Intervention on factors not specific to the content of the treatment such as the amount of attention participants receive, the quality and quantity of supervision and training for therapist, the internal validity of the trial remains intact. It yields differences between arms at the end of the study that can be attributed to differences in content between the new and alternative interventions.

P9 Reduces Across Site Variability in Treatment Effects: By comparing two standardized add-on treatments across sites, the variability in treatment effects is reduced, resulting in greater statistical power with fewer participants. Further benefit is realized when the standardized control treatment has been previously tested, as this allows for more informed estimates of the expected differences between the groups.

3.3.2 Negative Aspects of Design

Several of the negatives described above also apply to this design including: N6 Potential Overlap between the Two Groups, N7 Limited to Testing Add-On Interventions and, N 8 Site Variability Due to Synergistic Treatment Interactions (see Table 1). This design introduces several additional negatives.

N9 Failure to Detect a Difference between Two Effective Treatments: With this design, it is more difficult to show a treatment effect. Both interventions may be much more effective than TAU alone but not different from each other (null trial result) If results are not statistically significant, does that mean that both interventions were beneficial or that both were not? Both interventions may equally provide substantial added benefit to TAU or no added benefit. The results would not allow the investigative team to distinguish between these two possibilities.

N10 No Comparison to TAU: There is no direct comparison to what is currently practiced. If results are statistically significant, they do not indicate how different the effectiveness of each intervention is from that of TAU alone.

3.4 Design #4 New Intervention vs. Standardized Control Intervention

In this design, both the New Treatment and the control condition are standardized for uniform implementation across sites. The standard treatment may be one that has been designed and validated in previous studies or it may have been developed for the current study. This design allows for control of many threats to internal validity by matching the two groups on nonspecific factors such as amount of attention and quality of implementation. This design can be viewed as a variant on “New Intervention + TAU vs. Alternative Treatment + TAU”, except that in this design, the experimental and control treatment completely replace any TAU.

In situations where no usual care is available, the standardized control treatment may take the form of a treatment that is equal in nonspecific ways to the new intervention and believed to be neutral in its effect on the primary outcome variable. In this case, the investigative teams are not concerned with providing a direct comparison to current clinical practice since there are no interventions currently in use. In a treatment seeking population this situation is unlikely to exist except in ancillary intervention areas and is only desirable in that situation when there are no reasonable active interventions available. When some form of usual care does exist, an alternative to the standard inactive or neutral treatment approach is to use a standardized control treatment, which matches the new intervention on nonspecific factors and incorporates the most common treatment practices or philosophies. This approach has advantages for both generalizability and internal validity. It provides a comparison of the new intervention to a treatment similar in content to usual care.

The CTN protocol developed to evaluate Community Reinforcement and Family Training (CRAFT) provides an example of this design. The objective of this study was to investigate the effectiveness of CRAFT for motivating drug abusers to enter substance abuse treatment. This protocol was designed to include concerned significant others (CSOs) of treatment-resistant drug-abusing persons as the participants. The CRAFT approach utilizes a cognitive behavioral approach to train CSOs to use behavioral techniques for motivating drug-abusing persons to volunteer to enter treatment (Meyers & Smith, 1997).

A survey of CTP practices conducted during the protocol development process, revealed that for the participating CTPs the usual care for CSOs of treatment refusing persons with drug abuse disorders was an informal referral to self-help programs, primarily Al-Anon or Nar-Anon. The CRAFT intervention, consists of 12 manual-guided, 60-minute, individual counseling sessions. This study presented some interesting challenges from a design perspective since the usual care practice of an informal referral is similar to no treatment. However, in this case the type of referral, Al-Anon or Nar-Anon, indicates a specific treatment philosophy. The protocol design team decided to use a standardized control treatment, which reflects the most common usual care treatment philosophy and matches the new intervention in nonspecific factors, including amount, frequency and duration of attention to participants, training and supervision models, and therapy adherence and quality measures. The protocol defined the standardized control treatment as 12 manual-guided, 60-minute, individual counseling sessions of Al-Anon/Nar-Anon Facilitation Therapy (ANFT). The ANFT control was adapted from the Twelve Step Facilitation intervention used in Project MATCH and was also used in 2 Stage II randomized clinical trials investigating CRAFT (Miller, Meyers & Tonigan, 1999; Meyers, Miller, Smith & Tonigan, 2002).

3.4.1 Positive Aspects of Design—This design most closely resembles models used in efficacy trials as is reflected in the positives and negatives. Two positives described above also apply to this design including: P8 Strong Internal validity and P9 Reduces Across-Site Variability in Treatments (see table 1). Additional positives also apply to this design:

P10 Provides a Control Condition in Absence of Established TAU: This design works particularly well when TAU is not well specified. This is more likely the case in treatments that are a component of therapy, ancillary service or otherwise focused on those additionally affected by the disorder being studied (situations in which CTPs are more likely to have uneven levels of service).

P11 Control Can be Matched to TAU Philosophy: If the standardized control treatment is chosen or constructed to reflect the TAU philosophy, it allow CTPs to assess the effectiveness of the new intervention in comparison to a treatment similar to their usual practice.

P12 Defines What is Being Compared in The Two Treatments: By providing the full content of both treatments in a standard format with fidelity measures, the investigative team is able to define precisely what the treatments consist of, and therefore what is being compared in the two treatments.

3.4.2 Negative Aspects of Design—Several of the negatives described above also apply to this design including: N 9 Failure to Detect a Difference between Two Effective Treatments, and N10 No Comparison to TAU (see Table 1). This design introduces one additional substantial negative:

N11 No Inclusion of TAU: This design is the furthest removed from inclusion of TAU. This is a significant flaw in effectiveness research unless current practices are considered in designing the standardized control treatment to provide a meaningful point of comparison. If the control treatment does not have practical relevance to providers, even a well-designed trial with significant findings may fail to influence practice.

4.0 Discussion and Recommendation

We have noted that the successful development and implementation of a new behavioral treatment can be conceptualized as progressing through several stages of science, to dissemination and, finally adoption. Also noted is the failure of this progression and the resulting gap between research and practice in the substance abuse treatment (Lamb, Greenlick & McCarty, 1998). Numerous behavioral therapies have demonstrated efficacy in treating substance abuse however considerably fewer effectiveness trial outcomes are available. The NIDA CTN is one of the largest and most recent national efforts to conduct effectiveness trials of drug abuse treatment in community treatment settings. We have used four examples of trials developed in the CTN, to examine important considerations in choosing a control group. The complexity of clinical trial design increases considerably when moving from a single-site study at a carefully controlled research clinic to a community-based treatment program with a primary mission to deliver clinical care. These conditions may be further complicated in the case of multi-site effectiveness trial, where each site has unique characteristics that may interact with the new treatment. Community-based effectiveness trials are ultimately useful when the treatments being tested are applied and shown to work in practice. The more representative the sites in the study are of the range of actual settings and practices in the community the greater the generalizability of the results.

As with many aspects of clinical trial design, choice of a control group is largely driven by the key research question the trial is designed to answer. For example: What is the central question about treatment that the proposed clinical trial is being designed to answer? What is the right question to ask given the stage of scientific knowledge about the treatment being investigated? What are the needs of the treatment community and patients?

Our examination of the four designs presented here leads us to view the merits and compromises of each control group along a continuum from internal to external validity. As the choice of control increased in ecological validity it decreased in internal control.

We recommend the following considerations when choosing a control group or when interpreting the results of effectiveness trials. First, what is the strength of prior evidence for the efficacy of the intervention. If strong efficacy data are available investigators may be much less concerned with evaluating the treatment specific effects of the new intervention. In this situation the importance is placed on providing the most compelling evidence for adoption of the new intervention if the results show it to be superior and designs 1 & 2 may be the most desirable. However, if strong efficacy data are not available, or if the new treatment is modified from the version tested in previous efficacy studies, there may remain a need to establish the efficacy of the intervention. As a result the design may benefit from greater internal control to ensure the observed difference are specific to the new treatment and not due to non specific factors such as the amount of attention given to patients. In this circumstance designs 3 & 4 may provide more appropriate comparisons. A second recommendation is to consider the degree to which the control group provides a comparison that is useful to providers who are ultimately the adopters of the new treatment. Here the more representative the control condition is of the most common clinical practices the more easily the results can be interpreted by providers. In the case of designs 1 & 2 a direct comparison to TAU is provided and the results are easily understood. The provider can easily see the new treatment was inferior, equal, or superior to their practice depending on the results of the trial. Design 3 also incorporates TAU and provides a useful comparison however since neither condition is TAU only the interpretation is not as direct. Design 4 provides the least amount of information regarding a direct comparison with current practice. A final recommendation is to give careful consideration to what constitutes current treatment settings. Providing evaluation of new treatments in settings representative of those in which interventions will ultimately be adopted, provides the strongest support for adoption when the intervention is demonstrated to be effective. This approach anticipates the next set of questions in the process of moving from science to practice. These questions are posed by the providers; is this new treatment an improvement over my current practice, can I adopt it and, should I adopt it?

Acknowledgments

This paper is based in large part on a guidance document developed by the NIDA CTN Design and Analysis Workgroup to support lead investigative teams in the design of multi-site randomized clinical trials. Additional authors on the original guidance document include: Morton Brown, PhD, Cynthia Kleppinger, MD, Janet Levy, PhD, and Edward Nuñez, MD.

We would like to thank David Liu, MD, and Theresa Montini, PhD for their thoughtful comments and suggestions on the original guidance document. Finally, we would like to thank the CTN Publications Subcommittee for their review and comments on this paper.

This work was supported by National Institute on Drug Abuse (U10 DA13732-07)

References

- American Psychiatric Assn, Washington, DC, US. Diagnostic and statistical manual of mental disorders. Vol. 4th ed.. American Psychiatric Publishing, Inc.; Washington, DC, US: 1994.
- Brigham, GS. NIDA Clinical Trials Network; 2003. A Randomized, Multi-Site Clinical Trial Evaluating Community Reinforcement and Family Training (CRAFT) and AI-Anon Nar-Anon Facilitation Therapy (ANFT) for Assisting the Concerned Significant Others of Treatment-Refusing Drug Abusers. CTN 0025 Unpublished research protocol

- Borkovec, TD. Between-group therapy outcome research: Design and methodology. In: Onken, LS.; Blaine, JD.; Boren, JJ., editors. Behavioral Treatments for Drug Abuse and Dependence. National Institute on Drug Abuse Research; 1993. p. 249-289. Monograph 137
- Feaster DJ, Robbins MS, Horigian V, Szapocznik J. Statistical Issues in Multi-Site Effectiveness Trials: The Case of Brief Strategic Family Therapy for Adolescent Drug Abuse Treatment. *Clinical Trials* 2004;1:428-439. [PubMed: 16279281]
- French MT, Dennis ML, McDougal GL, et al. Training and employment programs in methadone treatment. *Journal of Substance Abuse Treatment* 1992;9:293-303. [PubMed: 1336067]
- Hall SM, Loeb P, Norton J. Improving vocational placement in drug treatment clients: A pilot study. *Addictive Behaviors* 1977;2:227-234. [PubMed: 343521]
- Johnson JL, Leff M. Children of substance abusers: Overview of research findings. *Pediatrics* 1999;103:1085-1099. [PubMed: 10224196]
- Lamb, S.; Greenlick, MR.; McCarty, D., editors. Bridging the gap between practice and research: forging partnerships with community-based drug and alcohol treatment. National Academy Press; Washington, DC, US: 1998.
- Meyers RJ, Smith JE. Getting off the fence: Procedures to engage treatment resistant drinkers. *Journal of Substance Abuse Treatment* 1997;14:467-472. [PubMed: 9437616]
- Meyers RJ, Miller WR, Smith JE, Tonigan JS. A randomized trial of two methods for engaging treatment-refusing drug users through concerned significant others. *Journal of Consulting and Clinical Psychology* 2002;70(5):1182-1185. [PubMed: 12362968]
- Miller WR, Meyers RJ, Tonigan JS. Engaging the unmotivated in treatment for alcohol problems: A comparison of three intervention strategies. *Journal of Consulting and Clinical Psychology* 1999;67(5):688-697. [PubMed: 10535235]
- Nathan, PE.; Stuart, SP.; Dolan, SL. Research on psychotherapy efficacy and effectiveness: between Scylla and Charybdis?. In: Kazdin, AE., editor. Methodological issues and strategies in clinical research. Vol. 3rd Edition. American Psychological Association; 2003.
- National Institute on Drug Abuse. Principles of Drug Addiction Treatment; A research-based guide. National Institute on Drug Abuse; Rockville, MD: 1999. NIH Publication No. 00-4180
- Onken, LS.; Blaine, JD.; Battjes, RJ. Behavioral therapy research: A conceptualization of a process. In: Henggeler, SW.; Santos, AB., editors. Innovative approaches for difficult-to-treat populations. American Psychiatric Press; Washington, DC: 1997. p. 477-485.
- Robbins MS, Szapocznik JS, Horigian VE, Feaster DJ, Puccinelli M, Jacobs P, Burlew K, Werstlein R, Bachrach K, Brigham G. Brief Strategic Family Therapy™ for Adolescent Drug Abusers: A Multi-Site Effectiveness Study. *Contemporary Clinical Trials*. in press
- Rounsaville BJ, Carroll KM, Onken LS. Methodological diversity and theory in the stage model: Reply to Kazdin. *Clinical Psychology: Science & Practice* 2001;8:152-154.
- Rounsaville BJ, Carroll KM, Onken LS. A stage model of behavioral therapies research: Getting started and moving on from stage 1. *Clinical Psychology: Science & Practice* 2001;8:133-142.
- Svikis, D. NIDA Clinical Trials Network; 2004. Job Seekers' Training for Patients with Drug Dependence. CTN 0020 Unpublished research protocol
- Szapocznik, J.; Hervis, OE.; Schwartz, S. National Institute on Drug Abuse Therapy Manuals for Drug Addiction. National Institute on Drug Abuse; Rockville, MD: 2003. Brief strategic family therapy for adolescent drug abuse. NIH Publication No. 03-4751

Figure 1.
New Intervention vs. Treatment-As-Usual (Positive Overall Result - Different Treatment Effects)

Figure 2.
New Intervention + TAU vs. TAU (Positive Overall Result - Similar Treatment Effects)

Figure 3.
New Intervention + TAU vs. Alternative Intervention + TAU (Positive Overall Result)

Table 1

Positives & Negatives	Designs			
Positive	1 New intervention vs. TAU	2 New intervention + TAU vs. TAU	3 New intervention + TAU vs. alternative intervention + TAU	4 New intervention vs. standardized control treatment
P1 Ecological Validity	X	X		
P2 Generalizability	X	X		
P3 Comparison with Different TAUs	X	X		
P4 Synergistic Interaction of Treatments		X		
P5 Reduces Impact of Differences in TAUs		X	X	
P6 Evaluates Add-on to TAU		X	X	
P7 All Participants Receive Established TAU		X	X	
P8 Control of Factors Not Specific to the Intervention			X	X
P9 Reduces Across Site Variability in Treatment Effects			X	X
P10 Provides a Control Condition in Absence of Established TAU				X
P11 Control Can be Matched to TAU Philosophy				X
P12 Defines What is Being Compared in The Two Treatments				X
Negatives				
N1 Drop Out of New Intervention Into TAU (Control):	X			
N2 Increased Sample Due to Variability in TAU	X			
N3 Potential for Mixed Results	X			
N4 Differences in Factors Not Specific to Treatment	X	X		
N5 Observed TAU is Not True TAU	X	X		
N6 Potential Overlap Between the Two Groups	X	X	X	
N7 Limited to Testing Add-On Interventions		X	X	
N8 Site Variability Due to Synergistic Treatment Interactions		X	X	
N9 Failure to Detect a Difference Between Two Effective Treatments			X	X
N10 No Comparison to TAU			X	X
N14 No Inclusion of TAU				X