



Published in final edited form as:

Drug Alcohol Depend. 2007 March 16; 87(2-3): 107–118.

Training and Fidelity Monitoring of Behavioral Interventions in Multi-Site Addictions Research: A Review

John S. Baer¹, Samuel A. Ball², Barbara K. Campbell³, Gloria M. Miele⁴, Eugene P. Schoener⁵, and Kathlene Tracy⁶

1 University of Washington Department of Psychology and the Alcohol and Drug Abuse Institute, and the VA Puget Sound Health Care System

2 Yale University School of Medicine, Department of Psychiatry, 1 Long Wharf - Suite 321, New Haven, CT 06510, Samuel.Ball@yale.edu

3 Oregon Health & Science University, Department of Public Health & Preventive Medicine, 3181 SW Sam Jackson Park Road, CB-669, Portland, Oregon 97239, 503-227-3932, drbarbaracampbell@earthlink.net

4 New York State Psychiatric Institute, Columbia University College of Physicians and Surgeons, 6182 Palomar Circle, Camarillo, CA 93012, 805-987-0340, gmm23@columbia.edu.

5 Wayne State University, Departments of Psychiatry, Pharmacology and Community Medicine, 2761 East Jefferson Avenue, Detroit, MI 48207. 313-993-1364, eschoen@med.wayne.edu

6 New York University School of Medicine, Department of Psychiatry, 423 East 23rd Street, 12 North 122D, New York, NY 10010, 212/686-7500, x3167, kathlene.tracy@med.nyu.edu

Abstract

AIMS—Methods for the training and fidelity monitoring of behavioral interventions in multi-site addictions research are reviewed, including five published studies and seven ongoing studies sponsored by the National Institute on Drug Abuse-funded Clinical Trials Network.

METHODS—Methods are categorized and reviewed consistent with a technology model of treatment delivery. Topics include: therapist selection, training, certification, and supervision; selection, training, and certification of supervisors; scales and processes used for monitoring of the quality of treatment; and processes followed to provide new training for replacement staff once trials have begun.

RESULTS—The review reveals both a wide array of procedures and emerging standards for multi-site trials. Methodological weakness was observed with respect to limited empirical support for many adherence scales, little or no evaluation of supervisory processes, and no evaluation of retraining practices.

CONCLUSIONS—Methods used in multi-site trials are important not only to ensure validity of those trials, but also to inform the wider dissemination of empirically based treatment into community agencies. Studies examining noted weaknesses are needed. Training and fidelity models that delegate responsibility to participating sites appear most relevant for establishing best practices for dissemination of behavioral interventions. The effectiveness of these distributed training and

Corresponding Author: John S. Baer, Ph.D., Alcohol and Drug Abuse Institute, University of Washington, Box 354805, 1107 NE 45th St., Suite 120, Seattle, WA 98105-4631, 206-616-3397, jsbaer@u.washington.edu

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

supervision models should be subjected to empirical study at a level of rigor comparable to the evaluation of their corresponding treatments.

Keywords

Training; Multi-site Trials; Treatment Fidelity

1. Introduction

The development and evaluation of psychotherapeutic interventions for alcohol and drug problems have been a focus of research for over 40 years. In a recent review of treatments for alcohol problems, Miller and Wilbourne (2002) identified 361 randomized trials evaluating 87 different treatment modalities. However, small studies that are dependent on unique features of sample, treatment, and measurement procedures often lead to disparate findings and may not advance scientific knowledge efficiently (Elkin et al., 1985). In the last decade, a trend has emerged to evaluate treatments in multi-site trials. Multi-site or collaborative research trials that standardize research methods across sites can provide large samples that greatly increase experimental power to detect small treatment effects (Vickers, 2003). Increased experimental power also allows for complex designs that can be used to evaluate interactions among and between features of clients and treatments (c.f., COMBINE Study Research Group (2003), Project MATCH Research Group (1993). Recently the Clinical Trails Network (CTN) of the National Institute on Drug Abuse (NIDA) has designed multi-site trials to be conducted in community agencies, thereby providing a methodology to evaluate treatments in real world settings and support the dissemination of empirically supported treatments into everyday practice.

A key component of the internal validity of any clinical trial is the integrity and discriminability of the treatments delivered. Multi-site behavioral therapy clinical trials require therapist training and fidelity monitoring on a large scale across diverse treatment settings, thus naturally suggesting training standards and informing research on treatment adherence and competence. A technology model of treatment research (Bellg et al., 2004;Docherty, 1984;Rounsaville et al., 2001) has been advanced to encourage researchers to articulate key features of interventions. In research on psychotherapy, a technology model typically includes specification of “dosage” delivery through the use of detailed manuals of intervention content, standardized training procedures, assessment to ensure that providers have acquired skills, and regular monitoring to minimize drift in provider skills over time. For each of these processes, many specific strategies may be utilized (Bellg et al., 2004). Among these many options, there are few standards or best practices for the conduct of multi-site research. Furthermore, when applied and modified for trials conducted in community practitioners who have little or no previous experience in specific treatments, multi-site trial methodology also promises to inform dissemination efforts for empirically supported treatments for substance use disorders.

The need to understand how evidence based treatments can be transferred to diverse community practice settings has received increasing attention. The substantive differences between research and common community practice are now well publicized (Institute of Medicine, 1998). Most behavioral therapy trials for substance abuse have been conducted in academic-affiliated settings, and, until recently, there has been very little evaluation of training and supervision strategies used in controlled experimental settings. Lack of attention to training inevitably contributes to a gap between research knowledge of substance abuse treatment and the clinical strategies used in community-based agencies. Standardized and effective training methods, in contrast, offer one bridge for research and practice differences. The NIDA funded CTN represents a major national initiative aimed at bridging this gap between scientific research and clinical practice through the investigation of empirically tested substance abuse

treatments in large multi-site trials in community substance abuse treatment programs across the country.

As the CTN and other projects grapple with complex challenges regarding the transportability and effectiveness of empirically supported treatments, it is timely to review the field critically and identify best practices of behavioral therapy training and fidelity monitoring in large scale initiatives in substance abuse treatment research. There are many conceptual and logistical challenges inherent in conducting multi-site randomized clinical trials. Effectiveness studies emphasize external validity and often evaluate treatments in “real world” settings with less experimental control than existed in the single site used for initial therapy development (Carroll and Rounsaville, 2003). Typically, there is more variability in clients, therapists, training, and settings. With multiple sites, a select number of relevant design features (in addition to the site itself) can be allowed to vary and thus be evaluated. However, these and other more logistical issues present challenges for the feasibility of maintaining some aspects of internal validity. Practitioners involved in multi-site studies are often dispersed over considerable geographic distances. The education, training, and experience can vary widely, and selection criteria for therapists may or may not support such diversity. Expertise for supervision and monitoring may or may not exist prior to study inception. Essentially, the fidelity of the behavioral treatment becomes a common intersection for the demands of internal and external validity. We describe a range of methods for behavioral therapy training and fidelity monitoring in five major multi-site initiatives with published outcomes and seven ongoing or recently completed CTN protocols. We note strengths and limitations, highlight important themes, summarize emerging standards, and provide recommendations for future multi-site research efforts that ultimately might inform dissemination initiatives.

2. Method

Studies were identified through literature review of known multi-site trials of behavioral treatments for drug and alcohol problems. Five published studies were identified: Matching Alcohol Treatment to Client Heterogeneity, (*MATCH*)(Project MATCH Research Group, 1997;Project MATCH Research Group, 1993); The Cannabis Youth Treatment Study (*Cannabis Youth*)(Dennis et al., 2004;Dennis et al., 2002); NIDA Collaborative Cocaine Treatment Study (*Cocaine Collaborative*)(Crits-Christoph et al., 1997;Crits-Christoph et al., 1999); Project COMBINE (*COMBINE*)(COMBINE Study Research Group, 2003); and the United Kingdom Alcohol Treatment Trial, (UKATT)(Team., 2005; UKATT Research Team, 2001). Additionally, we reviewed seven multi-site studies conducted within the NIDA CTN: Motivational Interviewing/Motivational Enhancement Therapy (*MI/MET*)(Carroll et al., 2006); HIV Risk Reduction During Detoxification (*HIV-Detox*)(Booth, 2005); Concurrent Treatment for Women with Trauma (*Women and Trauma*)(Hein, 2006); Smoking Cessation Treatment with Transdermal Nicotine Replacement Therapy in Substance Abuse Rehabilitation Programs (*Smoking Cessation*)(Reid, 2004); Job Seeker Training (*Job Seekers*)(Svikis, 2003); Safer Sex Skills for Women and Men (*Safe Sex Skills*)(Tross, 2005), (Calsyn, 2004); Brief Strategic Family Therapy (*BSFT*) (Szapocznik, 2004). We included all studies sponsored by the CTN to date that have included specific procedures for training and fidelity monitoring of behavioral interventions. Information for the review was obtained from published studies, protocol training manuals, and by contacting study investigators directly involved in the training and fidelity monitoring procedures, including authors of this review, all of whom have worked on training and fidelity in CTN trials.

Key design features and citations for the multi-site studies included in this review are listed in Table 1. The studies represent a wide array of behavioral interventions conducted at an average of nine collaborative sites. Sample sizes of client participants ranged from 360 to 1762; the studies trained an average of over 40 therapists each. With respect to experimental design, the

first five published studies are noteworthy for comparing different treatments. In contrast, CTN studies almost uniformly compared treatment as usual (TAU) to TAU augmented by an additional treatment method (i.e., additive designs).

We reviewed training and fidelity monitoring methods based on dimensions described by Elkin et al., (1985) for the NIMH Collaborative Study of Depression and consistent with a technology model of psychotherapy research (Bellg et al., 2004; Carroll et al., 1994; Docherty, 1984). These dimensions are reflected in the columns of Table 2. Therapists in trials first must be identified or *selected*. Once selected, therapists must *acquire skills* and *demonstrate competence* using the treatment. Typically, clinical trials therapists are supervised in their work, so we reviewed and included descriptions of *supervision processes*, as well as how *supervisors were selected, trained, and certified*. *Fidelity monitoring* procedures are described next. The types of *scales employed to both monitor and certify* therapists are reviewed as well. The next column in the table identifies procedures that were followed if therapists were performing below trial standards: *detecting low fidelity* (often referred to as “*redlining*”). Finally, the last column in Table 2 contains a brief description of procedures for the *training of new or replacement staff* while the trial is ongoing.

3. Results

Results of the review of training for multi-site trials in behavioral treatment for addictions are displayed in Table 2. Each of the elements in the training process, represented as columns in the table, are described below.

3.1. Therapist Selection

Across the 12 studies there was considerable variation in how therapists were identified, screened, and selected. Two different approaches to identifying study therapists were employed based on clinical setting and research design. Three studies (MATCH, Cocaine Collaborative, COMBINE) offered treatments primarily in research-affiliated settings and recruited clinical staff specifically for the multi-site trial. Other studies, including UKATT and all CTN trials, expected clinical staff to be employed (salaried or fee-for-service) within the participating community treatment agencies, most of which had not conducted treatment research previously. Cannabis Youth employed both of these methods differentially for its five therapeutic interventions.

Regarding screening and selection, seven studies required experience in treating addictions, typically 2 years, and five studies required specific professional educational achievement beyond bachelor’s degrees, most typically master’s degrees. Two studies required therapists to be allied professionally with the therapeutic approach under investigation. In contrast, some studies (i.e., CTN MI/MET; CTN Women and Trauma) excluded therapists who had prior training in the experimental treatment. Six studies used specific performance screens, including taped samples of clinical work to evaluate specific therapeutic skills (MATCH, Cocaine Collaborative, Women and Trauma, BSFT, UKATT), or taped role plays or clinical samples to observe more general skills such as empathy (COMBINE). Of the seven CTN studies included in this review, only two (Women and Trauma and BSFT) required that therapists demonstrate specific clinical skills for selection. Other CTN studies generally accepted therapists found in collaborating agencies who were interested in participating with relatively minor requirements (i.e., being a non-smoker to lead smoking cessation treatments, or having some experience in psycho-educational groups to lead Job Seeking group programs).

Variability in therapist selection appears to be a central dimension in external validity. Those studies seeking to demonstrate comparative efficacy of relatively new treatments or interactions between treatment and individuals tended to have the most restrictive procedures.

For example, in studies comparing the efficacy of different treatments, such as MATCH and Cannabis Youth, therapists with demonstrated experience and commitment to the therapeutic approach were sought. Both UKATT and COMBINE used clinical skill performance on tapes to determine eligibility prior to training. In contrast, most CTN studies (but not all, see BSFT) where treatments were compared to “treatment as usual” and conducted in community agencies, requirements for therapist selection were quite modest.

3.2. Centralized Training

There was little variability across the 12 studies with respect to the use of centralized trainings. All but one study brought therapists from multiple sites to a single location for group training prior to study initiation. The length of centralized training ranged from 1 to 4 days, with the exception of the Cocaine Collaborative, in which there were 4 separate 2-day trainings (8 full days) and CTN BSFT, which included one 4-day and three 3-day workshops (13 days total). All studies reported that training was based on detailed treatment manuals of the therapeutic method. All studies reported training to include didactic portions for background and rationale of treatment, reviews of specific treatment procedures, and demonstrations of treatment elements by video and/or role play. Variations in the amount of time devoted to centralized training presumably reflected differences in the nature and complexity of the behavioral treatments themselves, as well as the training goal. For example, the CTN Safe Sex Skills, which involved training of a psycho-educational group treatment for safer sex practices adjunctive to opioid treatment, required practice but no training beyond the initial centralized or core training. Other studies of more complex stand-alone treatments (e.g., Cocaine Collaborative and CTN BSFT Protocol) required considerable post-training practice and review after centralized training and prior to a certification process.

The one exception to the centralized training methodology was CTN MI/MET. Using an innovative design, each participating “node” (partnership of a research center and several community treatment programs) in the CTN was asked to identify a local MI Trainer who was trained centrally and subsequently conducted training (as well as supervision) of participating therapists and supervisors at the local performance sites.

3.3. Certification

All studies provide descriptions of certification procedures to ensure adequacy of therapist performance. The most common procedure involved audio- or videotaping a work sample (pilot care) that was reviewed by an expert in the treatment method. In one study, CTN Safe Sex Skills, certification was based on observed role-plays at a central training. In some studies where local supervision was available, there was an option for certification based on direct observation of actual intervention rather than tapes (CTN Smoking Cessation, CTN Job Seekers). Certification was based on session ratings using measures tailored to assess performance of specific clinical procedures (adherence measures are reviewed below). The specific decision rules for certification (cut scores on scales, specific behaviors observed on tape) were not described consistently.

3.4. Supervision and Therapist Support

In all multi-site trials reviewed, ongoing supervision and support of clinical staff was used to promote standardization and reliability of therapeutic activities. Supervision in this context was conceived as coaching, mentoring, and direction to support consistent and high fidelity treatment delivery. We have focused on the following dimensions as shown in Table 2: Supervision format, supervisor selection, supervisor training and certification, and monitoring format, setting and methods.

Supervisor qualifications include demonstrable knowledge and expertise in the content area as well as the process of supervision. A number of surrogate indicators were used in the studies we reviewed to determine qualifications and selection criteria. Key among these were professional degrees and credentials. Current position, years of experience, peer nomination (of an acknowledged expert), and unique skill sets were also used. None of the studies reviewed utilized an explicit process to establish the eligibility of supervisors, viz., core competencies for the role. An interesting distinction in supervisor selection seems to exist, however, between the NIDA CTN and other studies. The five published trials uniformly employed nationally acknowledged experts on specific treatments, whereas CTN trials, with some exceptions, engaged local clinical and research supervisors at study sites.

There was considerable variability in supervisor training across the studies reviewed. In one CTN study (BSFT) and three published studies (MATCH, Cocaine Collaborative, Cannabis Youth), *no* supervisor training (nor selection) was described, presumably obviated by the selection of national experts or previously trained individuals. Other studies, to varying degrees, described supervisor training that included both the content and process of supervision. For those instances in which supervisors received protocol-specific training, a session or mini-course of 4–16 hours typically was appended to a centralized training for clinical and research staff. In the UKATT study, supervisors were themselves supervised by a national expert. Some of these protocols required certification of supervisor proficiency by national experts using established criteria, both objective and subjective. Partial certification also was observed as in the CTN HIV-Detox study in which supervisors were certified as adherence raters using established criteria, but were not certified for other aspects of supervision. Three studies (CTN Women and Trauma, CTN Safe Sex Skills, and CTN HIV-Detox) evaluated the performance of study supervisors by requiring expert review of a subset of supervisor ratings. Approximately 25% of all sessions rated by the supervisor were co-rated by expert supervisors. Discrepancies in coding were reviewed on supervisor group conference calls run by intervention experts. In this way, the supervisors received feedback on their individual ratings and learned from the issues that arose in ratings by other supervisors.

The actual methods of supervision ranged considerably across studies, from on-site assistance with daily problem solving to formal weekly or bi-weekly communication between supervisors and their clinical or research staff. Supervision of clinical staff was described in all cases as based on performance criteria, but the mode of and tools for monitoring skill performance varied widely. The most rigorous approach called for supervisors to rate performance of the study clinicians with an objective assessment scale or skill check list using direct observation, video- or audiotape recording (i.e., MI/MET, HIV-Detox, Smoking Cessation, CTN Women and Trauma, CTN Safer Sex). In other cases, supervisors conducted regular case reviews with their study clinicians to discuss challenges they faced in implementing the protocol (i.e., CTN Job Seekers, CTN Smoking Cessation). Some clinicians were required to maintain detailed logs of their activity (e.g., Cannabis Youth).

Both centralized and local monitoring and supervision of behavioral treatment delivery were employed. While local supervision presumably employed face-to-face meetings, most studies relied heavily upon teleconferences for centralized supervision. Yet, across studies, conference calls varied with respect to participants (supervisors only, supervisors and counselors, or counselors only) and content (feedback from review of the adherence scales or general procedures and discussion surrounding the delivery of treatment). For example, the Cocaine Collaborative, COMBINE, Cannabis Youth, UKATT, MI/MET, Women/Trauma, Smoking, Job Seekers, Safe Sex Skills, HIV-Detox and BSFT trials made use of conference calling with central authorities. The Cocaine Collaborative, COMBINE, Cannabis Youth, UKATT, MI/MET, Women and Trauma and HIV-Detox studies utilized centralized expert raters for

adherence and treatment review, which guided some of the discussions on conference calls. In other trials, such as CTN Smoking, the conference calls were guided by feedback from the on-site supervisors' review of counselor adherence and competence with experts addressing the issues that were raised by the supervisors.

Many of the trials that included centralized monitoring also included some level of local monitoring. In Project MATCH, Cannabis Youth, and CTN MI/MET protocols, therapists completed self rating checklists in addition to the centralized expert ratings. Other CTN studies such as HIV-Detox, Women and Trauma, MI/MET, Safe Sex Skills, Smoking, and Job Seekers had local supervisors rate therapist delivery of treatment with adherence/competence scales. MI/MET utilized local experts as supervisors, who were coordinated by national experts who did not themselves interact with clinicians except when low fidelity was detected. Within all local monitoring, supervision ranged from weekly to monthly, or in some cases was based on counselor performance, with supervision decreasing as counselors demonstrated stable proficiency (e.g., CTN Women and Trauma and Safe Sex Skills).

We observed no obvious relationship between the content or complexity of the behavioral intervention and the supervisory processes used to support it. Given the range of potential supervisory activities, the presumed importance of supervision for therapeutic quality and consistency, and the many elements that could account for variability of supervision in multi-site trials, it is noteworthy that we were unable to find any systematic evaluations of supervisory processes. Indeed, we concluded that none of these multi-site studies subjected supervision to systematic inquiry.

3.5. Adherence Ratings Scales/Procedures for low fidelity

There was considerable similarity across studies in adherence scale procedures. All of the studies in this review used rating scales to monitor or evaluate therapist adherence to study interventions. With regard to construction, these scales ranged from session-specific content elements to broad domains of intervention methods. Differences along this dimension appeared to be related to the complexity of the intervention being assessed. Rating scales for simpler interventions, such as the CTN-HIV Detox study's single session therapeutic alliance intervention, itemized very specific intervention content. Ratings for more complex interventions, such as BSFT, itemized general intervention methods (e.g., use of "reframing"). Scales also differed in the measurement of competence or quality of intervention delivery as distinct from adherence, two factors that have been found predictive of treatment outcome (Barber et al., 1996a). Not all rating scales distinguished content adherence from quality of delivery. Eight of the 12 studies reviewed explicitly rated quality, labeled variously as quality, competence or skillfulness. Seven of those used scales that required separate ratings of content and skillfulness for the same intervention elements. Several studies (e.g., Cannabis Youth, CTN MI/MET, MATCH), employed an additional approach, using a general skillfulness scale to rate quality of treatment delivery across different interventions. This allowed for a direct comparison of therapist skill in treatment delivery across the interventions studied, a benefit not available with the methods used in other studies.

All studies reviewed used rating scales for initial therapist certification, establishing criterion scores for certification and ongoing quality standards. All studies also rated a selection of taped sessions across study implementation to assess and reduce intervention drift. Studies varied regarding who conducted ratings: independent experts, central supervisors or local supervisors. However, all of them used ratings' feedback as part of the supervision process to correct protocol deviations and improve quality. The majority of these studies used adherence scales scores to indicate low fidelity, although the specific criterion used was not always specified. In studies using this practice, therapists who fell below criterion performance received

enhanced supervision/training until performance exceeded threshold standards or they were suspended from accepting randomized participants. Most studies first used “redlining warnings” (Miller et al., 2005), a procedure in which monitoring, supervision and refresher training of therapists is increased while they continue to treat study participants. This was particularly true for studies employing group interventions, in which clinical continuity and cohesion were important. One study reviewed, the CTN-HIV Detox study, temporarily suspended therapists from accepting new study participants immediately upon falling below performance criteria.

Most rating instruments used did not have established psychometric properties. The Cocaine Collaborative reported initial reliability and validity data for adherence and competence scales developed for individual drug counseling, one of its three interventions (Barber et al., 1996b). For a second intervention, supportive expressive therapy, it used a scale with previously established reliability/validity data (Barber and Crits-Christoph, 1996). Other studies used unpublished, but previously developed scales (e.g., Seeking Safety Adherence Scale by Najavits, (2003) for CTN’s Women’s Treatment for Trauma, the Family Therapy Skills Rating System for BSFT by Santisteban), adapted previously published instruments (CTN MI/MET), or developed new scales to assess adherence (CTN Smoking Cessation study). Six of the 12 studies used some form of reliability check for ratings. For example, the CTN studies Women and Trauma, HIV- Safer Sex, and HIV Detox used a co-rating procedure in which supervisors’ and experts’ ratings were compared for selected sessions throughout the study. In COMBINE and MI/MET, central supervisory co-ratings were used as checks on ratings that fell below threshold performance standards. BSFT conducted two independent ratings followed by consensus agreement for initial therapist certification, and then a Head Training Supervisor held weekly meetings with supervisors and randomly listened to supervision sessions to minimize supervisor differences. Two studies, MATCH and MI/MET, did not check supervisory inter-rater reliability, but did use independent raters (blind to condition and results) to determine that the interventions were indeed reliably discriminable.

3.6. Re-Training

As described above, there is typically a high level of rigor and commitment of resources to the initial training of interventionists in multi-site clinical trials. Unfortunately, therapist attrition is a common problem for studies lasting extended periods of time. All trials reviewed described procedures for training new therapists during the study and conducted two or more additional trainings for replacement therapists. In three CTN trials, Smoking Cessation, HIV-Detox, and BSFT, sites were required to have back-up therapists who were certified (in HIV-Detox they were supervisors) and immediately available so as not to disrupt the treatment due to therapist attrition.

Subsequent trainings differed widely in their resemblance to the initial training. For example, MATCH trained as many additional, replacement therapists over the course of the trial as were trained in the initial cohort. These 10–12 subsequent trainings consisted of small numbers of clinicians, and the 3-day training was compressed into one-day. Although the original trainers were used for the face-to-face training, therapists were expected to assume greater personal responsibility for learning than was the case with the initial cohort: these additional trainings relied upon reading the manual, reviewing tapes of the original training event, watching session tapes, sitting in on supervision groups at the local site, and even beginning a pilot case prior to the face-to-face training. Although it is unclear whether the same pre-qualification criteria were always used as in the original cohort, the same post-training certification criteria were used before treating randomized cases.

Similarly, Cannabis Youth lost a significant number of therapists (1/4 of total) after the initial training. New therapists received a compressed training by the therapist coordinator, who had been an original trainer along with each treatment model developer, after reviewing training tapes, treatment manuals, and taped cases from their local site. In contrast, the time and resources involved in the initial training phase (which was a study in itself; Crits-Christoph et al., 1998) for the Cocaine Collaborative made the repetition of this process prohibitive. New therapists were instructed to read the treatment manuals, watch the videotapes made of the initial four weekend training sessions, were assigned one or more training cases, and began seeing randomized participants after certification.

The CTN MI/MET protocols did not rely on a centralized training model for therapists, but rather centralized the training/calibration of local MI experts who provided both the initial training and subsequent re-trainings at their sites. Thus, re-trainings maintained the same pre-training expectations of therapists and number of training days, although trainings typically involved fewer therapists. Other CTN trials blended aspects of all of these approaches. For example, the initial centralized training for the HIV-Detox and Smoking Cessation protocols trained interventionists and supervisors together. Subsequent trainings for new interventionists were managed by the on-site intervention specialist who presented didactic material and reviewed the intervention manual with the trainee who was also required to watch videotapes of the centralized training and practice the intervention through role plays with the supervisor. Thereafter, the new therapist demonstrated the treatment protocol on standardized certification cases with teleconference input provided by the national expert. A similar remote training process was offered when replacement site supervisors were needed. Other CTN protocols (e.g., Safe Sex Skills; Job Seekers) similarly relied upon review of initial training videotapes and reading training manuals and materials under the direction of a local certified supervisor with minimal involvement from the national expert trainers or participation in an intensive, centralized face-to-face training. In contrast, the Women and Trauma protocol maintained central control over the training procedures by either conducting centralized trainings for new therapists or, when that was not logistically feasible, presenting all of the original training material via slideshow teleconference facilitated by one of the original trainers and videotape review with local role plays managed by the on-site, certified supervisor.

4. Discussion

Over 20 years ago, Elkin et al. (1985) noted perceptively that all clinical research requires careful deliberation about methods used to train and deliver treatments, but that such details are seldom included in research reports (c.f., Borrelli et al., 2005). Our effort has been to review and explicate the many options for training in multi-site trials of behavioral treatments for addictions. As shown in Table 2, the magnitude of choices is large.

Both consistency and variability were observed across these 12 studies. For the most part, these studies were impressive in their attention to therapist training and treatment fidelity. All employed a core set of procedures. These procedures, arguably emerging standards, follow those first used in the landmark treatment collaborative study of depression and include: employment of treatment manuals, provision of standardized competency based training, use of rating scales for adherence measurement and quality improvement, employment of specific performance and certification procedures, monitoring of intervention delivery via review of sessions (in vivo or taped), supervision and support processes for those delivering treatments, and regular oversight of supervision and support procedures. With careful application of all of these procedures, readers of most studies should have confidence in the fidelity of the treatments provided in these multi-site trials.

Variation in training and fidelity practices were associated with differences in study objectives and design to some degree. For example, CTN studies, as well as UKATT, recruited therapists working in community agencies, whereas MATCH, COMBINE, and Cocaine Collaborative selected individuals with specific experience or training. The former procedure appears to be a key feature in studies that evaluate treatments conducted in community settings.

Another distinction between studies related to the selection of clinical supervisors. Studies comparing different established treatments and those examining client-treatment or treatment interactions usually employed national experts for initial training and ongoing supervision. The qualities and abilities of these experts were not measured. In contrast, most CTN studies employed additional procedures for identifying and training local supervisors. While certification procedures were employed for supervisors in some CTN studies, we are not aware of any empirically-based standards for supervisor selection and training.

The studies reviewed used a mixture of centralized and distributed models for training, supervision, and monitoring. In the more centralized model, the responsibility and quality control often remained with national experts at a coordinating center. This model presumably retains more consistency across sites, but must transcend geographic challenges. A more distributed training and supervision model, where responsibility rests with performance sites, minimizes geographic impediments and permits more immediate face-to-face interactions. On the other hand, distributed models create another source of variability for treatment fidelity, that of selection, training, and functioning of local supervisors. Despite this additional source of variability, distributed models provide more flexibility for the initiation of a trial and training of replacement staff once a trial has begun. Such models have the additional benefit of promoting more widespread use of study interventions. Staff at participating sites not only become skilled in the treatment itself, but in a range of activities that support and maintain treatment fidelity. It is noteworthy that a recent review of dissemination of empirically supported treatments for addictions emphasized that, in addition to initial workshop training, feedback from practice and ongoing “coaching” or supervision of actual performance is necessary to bring clinicians to proficiency levels and maintain changes in their treatment provision (Miller et al., 2006b). While some combination of central and distributed training and supervision will be used in future multi-site trials, further evaluation of the broader impact of distributed supervisory and monitoring processes is warranted to inform best practices in the dissemination of empirically based treatments.

Critical inspection of the training methods in these multi-site trials reveals three general areas of methodological weakness. First, although studies were consistent in the use of scales to rate adherence to clinical process, empirical validation for many of these scales is lacking. This is a difficult criticism to advance, as reliable and valid adherence scales have not been required as a standard for empirically based treatments. Yet our experience in conducting multi-site trials suggests that a reliable and valid rating scale becomes essential when attempting to standardize treatments across collaborating sites and supervisory processes. Rating scales certify competence in quantitative terms and provide an objective feedback tool to coach therapists in protocol adherence. Multi-site trials conducted without validated, reliable rating scales should consider publishing secondary papers on the performance of adherence scales, as well as the monitoring process, in order to advance the empirical base for these procedures. Barber et al. (1996b) provide an example in their report on the development and validation of adherence/competence scales for individual drug counseling in the Cocaine Collaborative.

Second, procedures used to supervise therapists in multi-site trials are both important and understudied. The studies reviewed reported an impressive array of approaches that varied as a

function of geographic distance, available technology, and demands of the specific clinical intervention. Yet, we were unable to discern patterns in supervisory processes based on research design or treatment content. Variations in supervisory competence and processes could decrease consistency of treatment delivery and imperil study integrity. Investigators for future trials must weigh the costs of time and resources in determining the nature and frequency of conference calls, tape reviews, methods of feedback, and absolute performance standards. As recruitment and training of supervisors becomes more common with current interest in empirically based treatments in community treatment settings, methods for evaluating supervisory processes must be validated, and controlled trials are needed to determine the efficacy of different models of supervision.

Third, training of new therapists due to staff turnover during study protocols occurred frequently and the methods used were highly variable across studies. This important area of variability receives comparatively limited attention in designing or reporting of clinical trials and is potentially problematic. Therapist effects have been identified as important as differences between patients and treatment conditions (Crits-Christoph et al., 1999). An important, unrecognized source of therapist effects within multi-site clinical trials could be the different training experiences of therapists depending on when they began participating in the clinical trial. Intangible characteristics of an initial training (e.g., excitement of launching a major national effort; traveling to a respected research university for training; meeting professional colleagues from across the country; being trained by the developers of the model) do not apply to subsequent training (e.g., local training provided by phone with expert or coordinated on-site by a recently trained “expert;” diminished sense of participating in a therapist team providing a common approach; sense of hurried expectations to take on new participants). The urgency to bring new therapists on board that local sites experience can influence training and certification processes to approve the therapist. Differences in training experience, compounded by the tendency for therapist and supervisory adherence monitoring procedures to drift later in the life of a protocol, may contribute to significant therapist effects and fidelity differences for patients treated at different times of a clinical trial. Fortunately, there are data available on therapist fidelity in most of the studies reviewed here. Thus, the potential impact of training variability is an empirical question awaiting study through a careful evaluation of therapist adherence and competence and patient outcome for therapist cohorts trained in the initial versus subsequent waves of training.

If it is shown that therapists given less intensive training mid-way through clinical trials function comparably to those trained initially with more intensive training, then investigators would be justified to revisit traditional methods used for initial training. Centralized trainings present many logistical difficulties, including distance, costs, and simultaneous availability of staff from different programs. In our experience, sites vary in their ability to hire staff and prepare for a clinical research protocol. Thus, different participating sites are rarely ready for training at the same time. As clinical trials are implemented in community treatment programs, this also requires that clinical operations (reimbursements, productivity) be suspended or greatly curtailed to accomplish training. If initial training can be handled remotely and more flexibly, the resources that are conserved could be redirected to observe and rate initial skill acquisition.

Finally, in order to implement new treatments, agencies administrators and practitioners are challenged with learning new treatment methods as well as adopting protocols for the acquisition of therapeutic skills and ongoing monitoring and enhancement of performance (Miller et al., 2006b). Difficulties in making such changes are numerous, including increased cost, time, lack of knowledge of treatments, and a lack of established systems to support the acquisition and maintenance of treatments (c.f., Miller et al., 2006a). To date we have few clear

and concrete recommendations for practitioners facing these dilemmas. We conclude by noting that multi-site trial methodology offers one important potential source for guidance for optimal methods to deliver effective behavioral therapy. Studies evaluating training outcomes are rare, and we are unaware of any studies of comparative costs or cost effectiveness of different training models. Such data are critically important to inform and support broader dissemination efforts. Multi-site studies of behavioral addictions treatments, particularly those conducted in community treatment settings, can facilitate dissemination of empirically based treatments in several ways: 1) providing data on generalizability of treatments to multiple and varied settings; 2) exposing and engaging treatment providers to research-based clinical protocols; 3) developing technologies for training, supervision and fidelity monitoring that can be used clinically; and 4) evaluating cost effectiveness for training and monitoring protocols. The procedures summarized in this review provide a beginning model for the training and fidelity methods that can be used beyond research to disseminate empirically supported therapies into community practice.

Acknowledgements

This paper was supported by grants from the National Institute on Drug Abuse as part of the Cooperative Agreement on National Drug Abuse Clinical Trials Network (CTN) U10 numbers DA13714 (Dennis Donovan, PI), DA13038 (Kathleen Carroll, PI), DA13036 (Dennis McCarty, PI), DA13035-05 (Edward Nunes, PI), DA13710 (Charles R. Schuster, PI), and DA 13046 (John Rotrosen, PI). Contents of the review are solely the responsibility of the authors and do not necessarily represent the official views of NIDA.

References

- Barber JP, Crits-Christoph P. Development of an adherence competence scale for dynamic therapy: Preliminary findings. *Psychother Res* 1996;6:79–92.
- Barber JP, Crits-Christoph P, Luborsky L. Effects of therapist adherence and competence on patient outcome in brief dynamic therapy. *J Consult Clin Psychol* 1996a;64:619–622. [PubMed: 8698958]
- Barber JP, Mercer D, Krakauer I, Calvo N. Development of an adherence/competence rating scale for individual drug counseling. *Drug Alcohol Depend* 1996b;43:125–132. [PubMed: 9023068]
- Bell AJ, Borrelli B, Resnick B, Hecht J, Minicucci DS, Ory M, Ogedegbe G, Orwig D, Ernst D, Czajkowski S. Enhancing treatment fidelity in health behavior change studies: Best practices and recommendations from the NIH Behavior Change Consortium. *Health Psychol* 2004;23:443–451. [PubMed: 15367063]
- Booth, R. Addiction Research and Treatment Services. Department of Psychiatry, University of Colorado Health Sciences Center; Denver, Colorado 80206: 2005. HIV and HCV risk reduction interventions in drug detoxification and treatment settings (Protocol for NIDA-CTN-0017).
- Borrelli B, Sepinwall D, Ernst D, Bell AJ, Czajkowski S, Breger R, DeFrancesco C, Levesque C, Sharp DL, Ogedegbe G, Resnick B, Orwig D. A new tool to assess treatment fidelity and evaluation of treatment fidelity across 10 years of health behavior research. *J Consult Clin Psychol* 2005;73:852–860. [PubMed: 16287385]
- Calsyn, DA. HIV/STD safer sex skills groups for men in methadone maintenance or drug-free outpatient treatment programs (Protocol for NIDA-CTN-0019). Alcohol and Drug Abuse Institute, University of Washington; Seattle, WA 98105: 2004.
- Carroll KM, Ball SA, Nich C, Martino S, Frankforter TL, Farentinos C, Kunkel L, Mikulich-Gilbertson S, Morgenstern J, Obert JL, Polcin D, Snead N, Woody GE. Motivational interviewing to improve treatment engagement and outcome in individuals seeking treatment for substance abuse: A multisite effectiveness study. *Drug Alcohol Depend* 2006;81:301–312. [PubMed: 16169159]
- Carroll KM, Farentinos C, Ball SA, Crits-Christoph P, Libby B, Morgenstern J, Obert J, Polcin D, Woody GE. MET meets the real world: Design issues and clinical strategies in the clinical trials network. *J Subst Abuse Treat* 2002;23:73–80. [PubMed: 12220604]

- Carroll KM, Kadden RM, Donovan DM, Zweben A, Rounsaville BJ. Implementing treatment and protecting the validity of the independent variable in treatment matching studies. *J Stud Alcohol* 1994;Supplement No.12:149–155.
- Carroll KM, Rounsaville BJ. Bridging the gap: a hybrid model to link efficacy and effectiveness research in substance abuse treatment. *Psychiatr Servi* 2003;54:333–339.
- COMBINE Study Research Group. Testing combined pharmacotherapies and behavioral interventions in alcohol dependence: Rationale and methods. *Alcohol Clin Exp Res* 2003;27:1107–1122. [PubMed: 12878917]
- Crits-Christoph P, Siqueland L, Blaine J, Frank A, Luborsky L, Onken LS, Muenz L, Thase ME, Weiss RD, Gastfriend DR, Woody G, Barber JP, Butler SF, Daley D, Bishop S, Najavits LM, Lis J, Mercer D, Griffin ML, Moras K, Beck AT. The National Institute on Drug Abuse Collaborative Cocaine Treatment Study. Rationale and methods. *Arch Gen Psychiatry* 1997;54:721–726. [PubMed: 9283507]
- Crits-Christoph P, Siqueland L, Blaine J, Frank A, Luborsky L, Onken LS, Muenz LR, Thase ME, Weiss RD, Gastfriend DR, Woody GE, Barber JP, Butler SF, Daley D, Salloum I, Bishop S, Najavits LM, Lis J, Mercer D, Griffin ML, Moras K, Beck AT. Psychosocial treatments for cocaine dependence: National Institute on Drug Abuse Collaborative Cocaine Treatment Study. *Arch Gen Psychiatry* 1999;56:493–502. [PubMed: 10359461]
- Crits-Christoph P, Siqueland L, Chittams J, Barber JP, Beck AT, Frank A, Liese B, Luborsky L, Mark D, Mercer D, Onken LS, Najavits LM, Thase ME, Woody G. Training in cognitive, supportive-expressive, and drug counseling therapies for cocaine dependence. *J Consult Clin Psychol* 1998;66:484–492. [PubMed: 9642886]
- Dennis M, Godley S, Diamond G, Tims F, Babor T, Donaldson J, Liddle H, Titus J, Kaminer Y, Webb C, Hamilton N, Funk R. The Cannabis Youth Treatment (CYT) study: Main findings from two randomized trials. *J Subst Abuse Treat* 2004;27:97–213.
- Dennis M, Titus J, Diamond G, Donaldson J, Godley S, Tims F, Webb C, Kaminer Y, Babor T, Roebuck MC, Godley M, Hamilton N, Liddle H, Scott C. The cannabis youth treatment experiment: Rationale, study design and analysis plans. *Addiction* 2002;97:16–34. [PubMed: 12460126]
- Docherty, JP. Implications of the technology model of psychotherapy. In: Williams, JB.; Spitzer, RL., editors. *Psychotherapy Research: Where are We and Where Should We Go?*. Guilford Press; New York: 1984. p. 139.-149.
- Elkin I, Parloff MB, Hadley SW, Autry JH. NIMH Treatment of Depression Collaborative Research Program. Background and research plan. *Arch Gen Psychiatry* 1985;42:305–316. [PubMed: 2983631]
- Project MATCH Research Group. Matching alcoholism treatments to client heterogeneity: Project MATCH posttreatment drinking outcomes. *J Stud Alcohol* 1997;58:7–29. [PubMed: 8979210]
- Hein, D. Social Intervention Group. Columbia University School of Social Work; New York, NY 10025: 2006. Women's treatment for trauma and substance use disorders: A randomized clinical trial (NIDA-CTN-0015).
- Miller WR, Baca C, Compton WM, Ernst D, Manuel JK, Pringle B, Schermer CR, Weiss RD, Willenbring ML, Zweben A. Addressing substance abuse in health care settings. *Alcohol Clin Exp Res* 2006a; 30:292–302. [PubMed: 16441278]
- Miller WR, Moyers TB, Arciniega L, Enst DAF. Training, supervision and quality monitoring of the COMBINE Study behavioral interventions. *J Stud Alcohol* 2005:188–195.
- Miller WR, Sorenson JL, Selzer JA, Brigham GS. Disseminating evidence-based practices in substance abuse treatment: A review with suggestions. *J Subst Abuse Treat* 2006b;31:25–39. [PubMed: 16814008]
- Miller WR, Wilbourne PL. Mesa Grande: A methodological analysis of clinical trials of treatments for alcohol use disorders. *Addiction* 2002;97:265–277. [PubMed: 11964100]
- Najavits, LM. Seeking Safety Adherence Scale Score Sheet. McLean Hospital; Belmont, MA: 2003.
- Project MATCH Research Group. Project MATCH: Rationale and methods for a multisite clinical trial matching patients to alcoholism treatment. *Alcohol Clin Exp Res* 1993;17:1130–1145. [PubMed: 8116822]

- Reid, M. Smoking cessation treatment with transdermal nicotine replacement therapy in substance abuse rehabilitation programs (Protocol for NIDA-CTN-0009). Department of Psychiatry, New York University School of Medicine; New York, NY 10010: 2004.
- Rounsaville BJ, Carroll KMLSO. A stage model of behavioral therapies research: Getting started and moving on from stage I. *Clin Psychol Science Practice* 2001;8:133–142.
- Svikis, D. Job seekers training for patients with drug dependence (Protocol for NIDA-CTN-0020). Department of Psychology, Virginia Commonwealth University; Richmond, VA 23298: 2003.
- Szapocznik, J. Brief strategic family therapy for adolescent drug abusers (Protocol for NIDA-CTN-0014). Center for Family Studies, University of Miami School of Medicine; Miami, FL 33136: 2004.
- UKATT Research Team. Effectiveness of treatment for alcohol problems: findings of the randomised UK alcohol treatment trial (UKATT). *BMJ* 2005;331:541. [PubMed: 16150764]
- Tober G, Godfrey C, Parrott S, Copello A, Farrin A, Hodgson R, Kenyon R, Morton V, Orford J, Russell I, Slegg G. Setting standards for training and competence: the UK alcohol treatment trial. *Alcohol Alcohol* 2005;40:413–418. [PubMed: 16027128]
- Tross, S. HIV/STD safer sex skills groups for women in methadone maintenance or drug-free outpatient treatment programs (Protocol for NIDA-CTN-0019). HIV Center, New York State Psychiatric Institute; New York, NY 10032: 2005.
- UKATT Research Team. United Kingdom Alcohol Treatment Trial (UKATT): hypotheses, design and methods. *Alcohol Alcohol* 2001;36:11–21. [PubMed: 11139410]
- Vickers AJ. Underpowering in randomized trials reporting a sample size calculation. *J Clin Epidemiol* 2003;56:717–20. [PubMed: 12954462]

Table 1

Multi-site Studies Included in Review

Study	Design	No. of Conditions	No. of Sites	No. of Participants	No. of Therapists Trained/Certified	Primary citations
MATCH	Compared CBT, 12-step and MET for alcohol problems as initial outpatient treatment or aftercare treatment in interaction with client individual differences	3	10	1,762	91 trained, 80 completed	(Group., 1997; Project MATCH Research Group, 1993)
Cocaine Collaborative	Compared Group Drug Counseling (GDC) alone vs. GDC + supportive-expressive dynamic therapy vs. GDC+CBT vs. GDC +individual in cocaine users	4	5	487	65 trained	(Crits-Christoph et al., 1997; Crits-Christoph et al., 1999; Crits-Christoph et al., 1998)
COMBINE	Compared 16-session Combined Behavioral Intervention (CBI) with Medication Management (MedM) in the context either or both of two medications (or placebo) for alcohol problems	9	12	1383	MedM: 50 trained/44 completed CBI: 44 completed	(COMBINE Study Research Group, 2003; Miller et al., 2005)
Cannabis Youth	Compared MET/CBT 5 session; MET/CBT 12 session; family support network; adolescent community reinforcement; multidimensional family therapy in adolescent cannabis users	5	4	600	20 trained, 17 certified	(Dennis et al., 2004; Dennis et al., 2002)
UKATT	Compared MET (3-sessions) and Social Behaviour and Network Therapy (8-sessions) for alcohol problems	2	7	742	76 recruited, 72 trained, 52 completed	(Tober et al., 2005; UKATT Research Team, 2001)
CTN:						
MI/MET	Compared 1-session MI Assessment vs. TAU Assessment; compared 3 session MET + TAU vs. Counseling-as-Usual + TAU in outpatient substance abusers	3	9	884	100 trained	(Carroll et al., 2006; Carroll et al., 2002)
HIV-Detox	Compared 2-session HIV counseling and education +TAU vs. 1-session therapeutic alliance with outpatient counselor +TAU vs. TAU only in outpatient substance abusers	3	8	632	45 trained, 45 certified	(Booth, 2005)
Women and Trauma	Compared 12-session Seeking Safety group + TAU vs. Health Education group + TAU in outpatient	2	5	353	35 trained, 33 certified	(Hein, 2005)

Study	Design	No. of Conditions	No. of Sites	No. of Participants	No. of Therapists Trained/Certified	Primary citations
	women with substance use disorders and PTSD					
Smoking Cessation	Compared smoking cessation groups+nicotine patch+TAU vs. TAU only in outpatient substance abusers	2	7	223	22 trained, 20 certified	(Reid, 2004)
Job Seekers	Compared 3-session Job Seekers Workshop + TAU vs. TAU only in outpatient substance abusers	2	11*	624 (312 DF, 312 MM)	22 total (1 primary per site plus 1 backup per site)	(Svikis, 2003)
Safe Sex for Women (W) and Men (M)	Compared gender-specific Safer Sex Groups (5-sessions) + TAU vs. HIV education (1 session) + TAU for men and women in outpatient and MM programs	2	12 (W), 14 (M)	516 (W), 594 (M)	52 trained and certified (W); 38 trained, 35 certified (M)	(Tross, 2005; Calsyn, 2004)
BSFT	Compared BSFT (12-16 sessions)+ TAU vs. TAU only in substance abusing adolescents	2	8	480 (projected)	23 trained, 17 certified	(Szapocznik, 2004)

CBT-Cognitive Behavior Therapy; MET – Motivational Enhancement Therapy; MI – Motivational Interviewing; TAU – Treatment as Usual; BSFT – Brief Strategic Family Therapy; DF – Drug

Free; MM – Methadone Maintenance; PTSD-Post-Traumatic Stress Disorder

* includes both drug free and methadone maintenance sites

Table 2
Training and Fidelity Monitoring Design Features of Multi-site Trials of Behavioral Interventions for Addictions

Study	Therapist Selection	Centralized Training (length)	Therapist Certification Process	Adherence Scale/ Certification Criteria	Supervision Format	Supervisor Selection	Supervisor Training /Certification	Monitoring Format/ Setting/ Method	Detecting Low Fidelity	Retraining/new Training: same vs. different/hours
MATCH	<ul style="list-style-type: none"> MA Degree + 2 yrs exp Exp w/ alcohol Tx Taped work Commitment to model 	3-day	<ul style="list-style-type: none"> 2 videotaped trng cases Independent ratings of adherence and competence 	<ul style="list-style-type: none"> Items rated as occurring for the approach had to be at least "good" 	<ul style="list-style-type: none"> Central telephone supervision Local site group supervision 	Local expert clinicians	Trng: Not described Cert: Not described	Central and local	2 sessions with ratings 2 or below out of 5	Compressed central 1-day Trng after review of tapes of original trng and materials
Cocaine Collaborative	<ul style="list-style-type: none"> 2 audio tape samples Agree to RCT and one-year trng phase 	Four 2-day trngs	<ul style="list-style-type: none"> Tape review by central supervisors 	<ul style="list-style-type: none"> Score at least average on rating scale 	<ul style="list-style-type: none"> Central-Weekly supervision 	National experts	Trng: Not described Cert: Not described	Central-Expert judges provide additional review	None were instituted Supervision of behavior appeared sufficient	2-day booster trng for all. New staff read and watch tapes of initial trngs
COMBINE 1) = MM 2) = CBI	<ul style="list-style-type: none"> 1) Med professional with 2 yrs exp 2) MA degree + 2 yrs exp + 2 tapes for empathy 	<ul style="list-style-type: none"> 1) 1-day 2) 3- or 4-day 	<ul style="list-style-type: none"> 1) Simulated cases & tape review 2) 2 clients tape review 	<ul style="list-style-type: none"> 7 pt rating scales for style and adherence to manual Average score of 5 required for cert. 	<ul style="list-style-type: none"> 1) Weekly local review 2) National calls and feedback from central tape review 	<ul style="list-style-type: none"> 1) Attend central trng 2) No information 	Trng1) Attend central trng Trng2) Not described Cert1/Cert2: Not described	<ul style="list-style-type: none"> Weekly local review Central tape review with feedback 	Below 5 on 7 pt scale	New trng at local sites. Submit tapes for central review/ Cert
Cannabis Youth	<ul style="list-style-type: none"> Existing staff + advertising Most MA degree "+" Clinical orientation congruent with study intervention. 	15-25 hours	<ul style="list-style-type: none"> Tape review of minimum 1 full training case 	<ul style="list-style-type: none"> Therapist Skillfulness Scale (TSS) most items must be rated "4" on 5-pt scale Other intervention specific scales 	<ul style="list-style-type: none"> Central twice weekly: 1 individual; 1 group tele-conference Local ongoing clinical issues 	National intervention specific clinical supervisor	Trng: Not described Cert: Not described	<ul style="list-style-type: none"> Central-2 tapes per month 	No formal criteria -	Review of tapes, manual, and practice cases
Study	Therapist Selection	Centralized Training (length)	Therapist Certification Process	Adherence Scale/ Criteria	Supervision Format	Supervisor Selection	Supervisor Training/ Certification	Monitoring Format/ Setting/ Method	Detecting Low Fidelity	Retraining/new Training: same vs. different/hours
UKATT	<ul style="list-style-type: none"> 2- yrs exp (relaxed during trial) Willing to be randomized Videotape sample of MI skills Able to work with 2 clients at one time 	Separate 3-day trng for each Tx	<ul style="list-style-type: none"> Judged competent by central trainer following review of 2 pilot cases for MET and 1 pilot case for SBNT 	<ul style="list-style-type: none"> Checklists developed for UKATT trial 	Video Tape review and telephone (or in person) consultation	Central expert	Trng: Supervised by PI for each Tx Cert: Not described	Central review of 1/3 of session tapes	Drift determined by central monitor Additional supervisory procedures	Same procedure
CTN: MI/MET	<ul style="list-style-type: none"> CTP counselor Willing to be randomized No MI trng past 90 days 	2-day trng of local experts who then trained counselors	<ul style="list-style-type: none"> Local Nodes provide initial trng Counselors conduct 3 or more pilot sessions 	<ul style="list-style-type: none"> Average MET/MI adherence rated adequate All competence items rated above average 	<ul style="list-style-type: none"> Local Monthly individual and monthly small group supervision 	Clinical leader	Trng: Local trng Cert: First certified as therapists; then as supervisors based on reliability of tape rating;	<ul style="list-style-type: none"> Local supervisor Central expert trainer 	Falling below certification criteria for three consecutive sessions	Review tapes centrally, increase supervision Local trainers trained new staff. Followed same model
HIV-Detox	<ul style="list-style-type: none"> CTP counselor Clinical exp with population 	3-day	<ul style="list-style-type: none"> Conduct minimum one trng case meeting adherence criteria 	<ul style="list-style-type: none"> 80% "3" on 5-pt scale of adherence and quality. 	<ul style="list-style-type: none"> Local bi-weekly for therapists Central monthly for supervisors 	<ul style="list-style-type: none"> Selected by CTPs supervisory exp recommend-ed 	Trng: 4 hours central Cert: First certified as therapists; then as supervisors with 80% reliability with expert rating of standard session tape	<ul style="list-style-type: none"> Local- supervisors rate 1/ 2 sessions Central- experts co-rate 1 of 3 sessions rated by supervisors 	Same as certification standard.	Teleconference + local trng - Manual review, standard tape review; role play practice
Women - Trauma	<ul style="list-style-type: none"> Female CTP counselor Consent for randomization Competent on problem solving demonstration tape 	3-day	<ul style="list-style-type: none"> Conduct pilot trng group Reviewed by central authority. 	<ul style="list-style-type: none"> Intervention specific avg. scores 2 on 0-3 pt scale 3 on 1-5 pt scale 	<ul style="list-style-type: none"> Local weekly supervision with tape review 	<ul style="list-style-type: none"> Same as therapist, plus Selected by CTP Prior supervisory exp 	Trng: 4 hours at central trng Cert: First certified as therapist; then as supervisors with 70% agreement with expert on adherence ratings of pilot group sessions	<ul style="list-style-type: none"> Central, expert co-rating of supervisor ratings Weekly conference calls 	6 consecutive sessions below adequate	Small central or teleconference and videotape

Study	Therapist Selection	Centralized Training (length)	Therapist Certification Process	Adherence Scale/ Criteria	Supervision Format	Supervisor Selection	Supervisor Training and Certification	Monitoring Format/ Setting/ Method	Detecting Low Fidelity	Retraining/new Training: same vs. different/hours
Smoking Cessation	<ul style="list-style-type: none"> •CTP counselor •6-mo SA Tx exp •BA or state certified SA counselor •Non-smoker 	3-day	<ul style="list-style-type: none"> •Practice w/ Observation 	<ul style="list-style-type: none"> •Frequency and skill of techniques rated on 1-8 likert scale across all 9 sessions •75% of the maximum score 	<ul style="list-style-type: none"> •Local monthly supervision 	<ul style="list-style-type: none"> •Ph.D. or MA & SA Certified •Exp in cognitive behavioral therapy 	Trng: Trained on coursework for intervention and supervision issues incorporated into the 3-day training.	<ul style="list-style-type: none"> •Local tape review and supervision of 2 sessions per month •Central bi-weekly national supervisory calls 	75% of maximum score.	Onsite and conference call trng. Practice and ongoing sessions rated with the adherence monitoring form.
Job Seekers	<ul style="list-style-type: none"> •CTP counselor •Psycho-education and clinical exp 	2-days	<ul style="list-style-type: none"> •Practice workshop presentation with supervisor observation & rating scale •Telephone supervision with national experts. 	<ul style="list-style-type: none"> •Single adherence monitoring form 	<ul style="list-style-type: none"> •Local weekly supervision •Monthly supervisor observation of workshop 	<ul style="list-style-type: none"> •CTP supervisors •Ph.D. or MA plus SA cert 	Trng: Same as therapists plus 2 days on supervision issues. Cert: completion of centralized trng and observation by national expert	<ul style="list-style-type: none"> •Local supervisor observation •Central calls with national expert 	No formal score. Judged below criterion by national experts	None reported -- ongoing monitoring and correction as needed.
HIV Safe Sex women and men	<ul style="list-style-type: none"> •CTP counselor •HIV-Group Exp. 	2 ½ days	<ul style="list-style-type: none"> •At trng •Role Plays rated by national experts 	<ul style="list-style-type: none"> •Intervention specific scale •all scores "adequate" (3) or better on 1-5 pt scale 	<ul style="list-style-type: none"> •Local monthly with tape review •National monthly for first 6 months of study 	<ul style="list-style-type: none"> •Selected by CTP directors •Group exp preferable 	Trng: 3 hours at central trng Cert: At central trng, 70% agreement with ratings made by national experts	<ul style="list-style-type: none"> •Local tape review and supervision •Central review of supervisor ratings •National monthly calls 	Ratings below "adequate"	Video of original trng Review manual Tape of role play reviewed by national expert
BSFT	<ul style="list-style-type: none"> •CTP counselor •BA+2 years exp •Video of family Tx session 	1 4-day and 3 3-day workshops (13 days total)	<ul style="list-style-type: none"> •Videotaped pilot cases •Weekly phone supervision 	<ul style="list-style-type: none"> •Two independent ratings on each of 4 intervention domains •avg 3.75/5 on rating scale 	<ul style="list-style-type: none"> •Central 	<ul style="list-style-type: none"> •BSFT national experts 	Trng: Not described Cert: Not described	<ul style="list-style-type: none"> •Central weekly phone supervision •Random, weekly tape review 	Unable to meet criterion after 3 pilot cases	Not described

Cert – Certification; CTP - Community Treatment Program; Exp - experience; MA - Master's Degree; SA - Substance Abuse; Trng - Training; Tx - Treatment