Published in final edited form as:

J Subst Abuse Treat. 2012 September ; 43(2): 211–220. doi:10.1016/j.jsat.2011.11.003.

The effectiveness of community-based delivery of an evidence-based treatment for adolescent substance use

Sarah B. Hunter, Ph.D.^a, Rajeev Ramchand, Ph.D.^b, Beth Ann Griffin, Ph.D.^b, Marika J Suttorp, M.S.^a, Daniel McCaffrey, Ph.D.^c, and Andrew Morral, Ph.D.^b

^aDrug Policy Research Center, RAND Santa Monica, CA 90407

^bDrug Policy Research Center, RAND Arlington VA, 22202

^cDrug Policy Research Center, RAND Pittsburgh, PA, 15213

Abstract

This study evaluates the effectiveness of Motivational Enhancement Therapy/Cognitive Behavioral Therapy-5 (MET/CBT-5) when delivered in community practice settings relative to standard community-based adolescent treatment. A quasi-experimental strategy was used to adjust for pre-treatment differences between the MET/CBT-5 sample ($n = 2293$) and those who received standard care ($n = 458$). Results suggest that youth who received MET/CBT-5 fared better than comparable youth in the control group on five out of six 12-month outcomes. A low follow-up rate (54%) in the MET/CBT-5 sample raised concerns about nonresponse bias in the treatment effect estimates. Sensitivity analyses suggest that while modest differences in outcomes between the MET/CBT-5 nonrespondents and respondents would yield no significant differences between the two groups on two of the six outcomes, very large differences in outcomes between responders and nonresponders would be required for youth receiving MET/CBT-5 to have fared better had they received standard outpatient care.

Keywords

adolescent outpatient treatment; marijuana; outcomes; evidence supported treatments; quality assurance

1. Introduction

In 2007 there were over 200,000 adolescent admissions to publicly funded substance abuse treatment programs in the United States, and over half of these admissions involved cannabis use (Substance Abuse and Mental Health Services Administration. Office of Applied Studies, 2009). In fact, cannabis has been the predominant drug used by adolescent substance abuse treatment clients seen in the US for the past half-century (M. Dennis, Babor, Roebuck, & Donaldson, 2002). Trends in cannabis use among adolescents have fluctuated during this time, though rates of use were higher in 2008 than they were twenty

© 2011 Published by Elsevier Inc.

Corresponding author: Sarah Hunter, PhD, RAND, 1776 Main Street, Santa Monica, CA 90407-2138, Phone: (310) 393-0411, ext. 7244, shunter@rand.org.

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.

years earlier (Child Trends, 2010). Together, these data suggest that developing, evaluating, and disseminating treatments effective for cannabis-related problems and disorders among adolescents are critically important. Recognizing this, in 1997 the Center for Substance Abuse Treatment (CSAT) created the Cannabis Youth Treatment (CYT) cooperative agreement. The goals of the CYT agreement were two-fold: first, to develop effective short-term (i.e., less than 3 month) outpatient adolescent treatment with cannabis-use problems that could be easily disseminated; and second, to conduct a field experiment to evaluate the cost, effectiveness, and cost-effectiveness of those treatments (M. Dennis et al., 2004).

One of the primary findings from the CYT field experiment was that the briefest option, a five-session manual-guided treatment designed to be delivered over a 5–7 week period, termed Motivational Enhancement Treatment/Cognitive Behavioral Therapy-5 or MET/CBT-5 (Diamond et al., 2002; Sampl, Kadden, & Center for Substance Abuse Treatment (U.S.), 2001) achieved clinical outcomes similar to those of the other four treatments to which it was compared, all of which typically consisted of double the number of sessions delivered over twice as long a time frame. In light of these findings, starting in 2003, CSAT launched the Effective Adolescent Treatment (EAT) initiative that provided support for 37 community based programs across the U.S. to adopt MET/CBT-5 to treat adolescent cannabis use. The selected community-based programs were provided three years of support, including training and supervision in the MET/CBT-5 protocol and were required to collect client outcome data using a standardized instrument.

Historically there has been little overlap between the substance abuse treatments used in community practice and those with scientific evidence of efficacy (Miller et al., 2006). This may be in part due to the challenges associated with transferring such treatments to community based programs (Institute of Medicine (U.S.) Committee on Crossing the Quality Chasm, 2006; Liddle et al., 2002). A number of factors have been cited as potential barriers, including differences between the mix of clients seen in research and usual care settings as well as in the resources available (e.g., to train and supervise providers) in research versus typical community practice settings (Weisz, Donenberg, Han, & Weiss, 1995; Weisz, Weiss, & Donenberg, 1992). Supporting community-based providers to adopt treatments with scientific evidence of efficacy could be a potentially valuable policy lever used to improve behavioral health care quality in the United States (Institute of Medicine (U.S.) Committee on Crossing the Quality Chasm, 2006), but first a better understanding of whether the adoption of an evidence-supported treatment, like MET/CBT-5, will improve on upon existing outcomes in community based settings is needed. Since the efficacy of usual care community treatment for adolescents has not been well established (Morrall, McCaffrey, & Ridgeway, 2004), it is not clear whether MET/CBT-5 as delivered in usual care settings will improve outcomes.

To date, only one published study provides some insight on the value of disseminating MET/CBT-5 to improve outcomes in standard outpatient care. A variant of this treatment that offered two additional family therapy sessions (termed MET/CBT-7) was tested and deemed comparably effective, though at significantly less cost, to a community-based care control group (Godley et al., 2010). This study was conducted under controlled research conditions, using a randomized intent-to-treat study design where the MET/CBT-7 treatment was compared to standard care delivered in the same treatment program. In addition, the study was conducted by the same organization that participated in the CYT field experiment, therefore the resources may not reflect those of a typical treatment program attempting to adopt and implement MET/CBT-5 for the first time. In contrast, examining outcomes from the 37 community-based programs funded to deliver MET/CBT-5 under the CSAT initiative may help to address the generalizability of these findings, that is, how well MET/CBT-5 performs when delivered in real world treatment settings.

In order to evaluate MET/CBT-5 effectiveness in usual care settings, we compared the outcomes achieved by the EAT funded programs to standard care delivered in other CSAT-funded outpatient community-based programs. A similar approach was used in a quasi-experimental study comparing MET/CBT-5 in experimental conditions (CYT) to community-based treatment (Ramchand, Griffin, Suttrop, Harris, & Morral, 2011). In both that and the current study, “standard” care was approximated by three outpatient treatment programs deemed “exemplary” by the Substance Abuse and Mental Health Services Administration/Center for Substance Abuse Treatment (SAMHSA/CSAT). In the absence of a representative sample of contemporary outpatient programs, these three outpatient programs offer a viable comparison sample, though to the extent that these programs are “exemplary” they represent a conservative comparison group for assessing the relative effectiveness of MET/CBT-5 in community-based treatment settings. In the current study, strategies were used to adjust for any pre-treatment heterogeneity in client characteristics between the two samples. More specifically, propensity score weights were used to account for differences between youth who received MET/CBT-5 in the EAT initiative and youth who received treatment in the standard outpatient treatment settings. Propensity score weights allow us to compare outcomes across the two settings by controlling for any observed pre-treatment differences across the two samples. Consistent with the results from studies of MET/CBT-5 conducted under more controlled research conditions, we hypothesized that the outcomes achieved under MET/CBT-5 in the EAT initiative would be better than those demonstrated in standard outpatient treatment settings.

2. Methods

2.1 MET/CBT-5 Intervention: The EAT Sample

Youth in the MET/CBT-5 arm were drawn from the EAT program initiative. EAT included a sample of adolescents who received MET/CBT-5 from one of 37 different treatment programs located across the country including such locations as Burlington, VT, Pinellas Park, FL, Pontiac, MI, Laredo, TX, and Santa Barbara, CA. Two cohorts of programs were funded, the first in 2003 ($n = 22$) and the second cohort in 2004 ($n = 15$). The initial grant funding was for a three year period and included training and technical assistance in delivering the therapy (MET/CBT-5) and in implementing the Global Appraisal of Individual Needs (GAIN) assessment tool (see *Measures*, below). The support for the therapy and assessment tool included didactic training and active role plays followed by a certification process whereby program staff had audiotaped practice sessions (for the therapy or assessment) reviewed by an expert who provided feedback (Godley, Garner, Smith, Meyers, Meyers and Associates, Godley, 2011).

There were 4,225 youth in the EAT sample who were eligible to have a 12-month follow-up interview during the study funding period (i.e., the client was not enrolled during the last 12 months of the grant period) and 2,293 (54% of the eligible EAT sample) completed the 12-month follow-up interview.

2.2. Comparison Outpatient Treatment: The Adolescent Treatment Model (ATM) Sample

Youth in the comparison condition were drawn from three outpatient treatment programs funded in 1998–99 under the auspices of the Adolescent Treatment Model (ATM) program also supported by SAMHSA/CSAT (M.L. Dennis, Dawud-Noursi, Muck, & McDermeit, 2003). Criteria for funding included: (1) two years of continuous operation, (2) evidence or testimony of the merits of the treatment program, (3) willingness to participate in an evaluation of program effectiveness and collaboration with other selected program sites, and (4) development of a treatment manual that would be publicly available (M.L. Dennis et al., 2003). The three outpatient programs used in the current analysis were: (1) Chestnut Health

Systems' outpatient adolescent services located in Bloomington, IL (Godley, Risberg, Adams, & Sodetz, 2003), (2) the Epoch Counseling Center in Baltimore, MD (Battjes et al., 2003), and (3) the Teen Substance Abuse Treatment in Maricopa County, AZ (Stevens et al., 2003). Since the time it was funded as an ATM site, each program has provided evidence that its clients exhibited some reductions in adverse behaviors (Battjes et al., 2004; Godley, Jones, Funk, Ives, & Passetti, 2004; Stevens, Schwebel, & Ruiz, 2007) and all programs performed comparably in analyses that controlled for pretreatment differences in client characteristics (Morral, McCaffrey, Ridgeway, Mukherji, & Beighley, 2006).

Our analysis used data from the 431 youth who were re-contacted at 12-months post-intake from the 463 youth that were assessed at these three community-based treatment programs at intake (for a 93% response rate).

2.3. Measures

Data documenting pre-treatment client characteristics and 12-month treatment outcomes in both EAT and ATM program sites was obtained from the same biopsychosocial instrument, the Global Appraisal of Individual Needs (GAIN) (M. L. Dennis, 1999). The GAIN has eight sections assessing background and demographic characteristics (asked at baseline only), substance use, physical health, risk behaviors, mental health, environment, legal, and educational/vocational problem areas. It includes over 100 symptom counts, change scores, and service utilization indices, most with Cronbach alphas over .85 (M. L. Dennis, 1999). A test-retest analyses of key GAIN substance use and substance use problem indices in a sample of 210 adolescents indicated good reliability over a 90-day interval ($r > .72$ for each index) (M. Dennis et al., 2000).

2.3.1 Pre-treatment characteristics—To adjust for observed pre-treatment heterogeneity between the two samples, we began our propensity score analysis with 60 pre-treatment variables related to substance abuse patient placement criteria established by the American Society for Addiction Medicine (Mee-Lee & American Society of Addiction, 2001). These variables represent multiple domains that affect treatment outcomes such as pre-treatment substance use, mental health symptoms, substance use and mental health treatment utilization, criminal activity and justice system involvement, scholastic attendance and academic performance, environmental characteristics (e.g., peer and familial substance use), and sociodemographics. Since controlling for variables unrelated to the outcomes may not improve estimates of treatment effects (Wooldridge, 2002), we included in our propensity score model only those pretreatment covariates that explained at least one percent of the variance in our outcomes. This resulted in a total of 22 pretreatment variables which are listed in Table 1. Missing values on the 22 pre-treatment variables were low (mean = 0.73% and max = 4.85%) and controlled for in our propensity score model by balancing the two groups on missing value indicators for each variable as well as the pre-treatment variables (details in following section).

2.3.2. Outcomes—Given the range of adverse outcomes associated with harmful adolescent substance use, we examined 12-month outcomes in six domains consistent with those reported in the CYT trial: recovery, institutionalization, emotional problems, illegal activities, substance use problems, and substance use frequency. *Recovery* was operationalized as a dichotomous (0/1) variable that combined the traditional outcome of abstinence with that of being in the community, and therefore at risk for using. More specifically, *Recovery* was coded as a "1" if in the past month the client was living in the community, abstinent from all substances and reported no substance problems. *Institutionalization* was also dichotomous and indicated whether a youth spent one or more of the past 90 days in any of several types of controlled environments in which illicit drug

use and liberty were substantially constrained for the whole day (e.g., jail, inpatient treatment, group homes, or probation camps). The *Emotional Problem Scale* (EPS) is a 7-item scale assessing psychological distress recency, frequency and severity during the 90 days preceding assessment with higher scores indicating greater emotional problems ($\alpha = 0.80$ in the ATM dataset). The *Illegal Activities Scale* (IAS) is a three-item scale assessing the recency and days of illegal activity and illegal financial support in the past 90 days, with larger values indicating more recent activity, more days of illegal activity and more dependence on illegal activities for financial support ($\alpha = 0.74$ in the ATM dataset). To assess problems from use, we utilized the *Past Month Substance Problem Scale* (SPS) that is based on recency ratings on 16 symptoms: seven consistent with the DSM-IV criteria for drug dependence, four consistent with the criteria for abuse, two for substance-induced health and psychological problems and three related to lower severity symptoms (e.g., hiding use, people complaining about use, and weekly use). Interitem reliability on this scale was excellent ($\alpha = 0.92$ in the ATM dataset). The *Substance Frequency Scale* (SFS) is an 8-item scale that assesses the average proportion of alcohol and other drug using days in the past 90 taking into account heavy use and problem days. Higher scores indicate increased frequency of substance use in terms of days used, days staying high most of the day, and days causing problems. Interitem reliability was good ($\alpha = 0.80$ in the ATM dataset).

2.4. Data Analysis

2.4.1 Propensity scores weights—Our analytic approach was designed to answer the following question: *When disseminated to a set of community-based outpatient programs, is MET/CBT-5 effective relative to other outpatient treatment offered in the community?* To answer this question, we used propensity score weighting to compare the outcomes of youth receiving MET/CBT-5 in the EAT initiative with those of similar clients receiving outpatient care at the ATM sites. It is important to note that this strategy does not allow us to interpret how those receiving outpatient care in the ATM study would have fared had they received MET/CBT-5 in the EAT study.

Propensity score weighting is a quasi-experimental strategy that requires that we estimate “counterfactual” outcomes for youth receiving MET/CBT-5 in EAT that measure the outcomes that would have been observed had they received treatment in ATM (Rosenbaum & Rubin, 1983; Rosenbaum, 1995; Hirano, Imbens & Ridder, 2003; Imbens, 2004). To do this, we created propensity score weights to weight the ATM group who completed the 12 month follow-up interview to match to the EAT (MET/CBT-5) group who completed the 12-month follow-up interview with respect to the distributions on the 22 pre-treatment variables (and corresponding missing value indicators). Thus, we “upweight” ATM clients who are similar to youth in the EAT (MET/CBT-5) study with respect to the 22 variables and “downweight” those who are not.

The resulting weighted ATM group is thus balanced to look similar to the EAT group on the 22 variables, but as demonstrated by the “effective sample size” this is achieved with a possible loss in precision of the weighted group mean relative to that of the unweighted group mean. However, these propensity score weights help to ensure that any observed differences in our outcomes between the EAT (MET/CBT) and ATM groups are not due to systematic differences between the groups on these 22 pre-treatment variables.

The propensity score weights in our analysis were estimated using a Generalized Boosted Model (GBM) which is a flexible, non-parametric estimation technique that can account for a large number of pre-treatment covariates in the model of the binary treatment group assignment variable (e.g. ATM vs. MET/CBT-5 in EAT). GBM adaptively captures the functional form of the relationship between the covariates and treatment group selection

with less bias than traditional approaches (Lee, Lessler, & Stuart, 2009; McCaffrey, Ridgeway, & Morral, 2004). The weight estimation was implemented using the “twang” package in R (R Development Core Team) and balance between the treatment groups was assessed across a number of criteria (Ridgeway, McCaffrey, & Morral, 2010).

To show how comparable the ATM and EAT (MET/CBT-5) groups were after applying the propensity score weights, we calculated the pre- and post-weighting standardized mean difference (SMD) for each variable used to create the weights. The SMD is estimated by (Ridgeway et al., 2010):

$$\frac{\hat{\mu}_t - \hat{\mu}_c}{\hat{\sigma}}.$$

Here, $\hat{\mu}_j$ denotes the estimated mean value for the treatment and control conditions ($j = t$ and c , respectively); $\hat{\sigma}$ denotes the estimated standard deviation for the variable in question and is estimated using the observed variance in the (unweighted) treatment (i.e., MET/CBT-5) condition. Values of 0 for a SMD represent no difference in means while values of ± 1 represent one standard deviation difference between the two groups. Since standardized differences of 0.25 are generally considered to be small to medium effect size differences, we denote group differences to be notable when the absolute standardized mean difference is greater than 0.25 (Cohen, 1992). These variables were included as covariates in the outcome analyses to control for any lingering confounding of the variable after weighting (Neugebauer and van der Laan, 2005).

2.4.2. Outcome Analyses—We estimated the treatment effect for each of our outcomes separately at 12-months post-intake by computing the mean difference between the youth receiving MET/CBT-5 in the EAT and the propensity score weighted ATM youth, along with corresponding 95% confidence intervals (CI).

2.4.3. Sensitivity analyses—The EAT initiative obtained relatively low response rates for the 12-month follow-up interview (54% among those eligible for a 12-month follow-up interview) in comparison to the ATM study (12-month follow-up rate of 93%). This could bias our results to favor the youth receiving MET/CBT-5 in the EAT program if youth with worse outcomes were less likely to complete the 12-month interview, or to favor ATM if youth with better outcomes were less likely to complete the 12-month interview. We therefore developed a series of sensitivity analyses to explore the magnitude of the association that would have to exist between loss to follow-up and our outcomes to explain our results (Griffin, McCaffrey, Ramchand, Hunter & Suttrop, under review).

Our sensitivity analysis involved a series of simulation studies which randomly “censored” youth in the ATM study in order to make the 12-month follow-up response rates in the ATM and EAT the same and then re-ran our analysis for each outcome. Although we do not know which 46% of ATM clients to censor from the sample (i.e., those with good or bad outcomes), by censoring clients based on their 12-month outcomes we can determine the degree to which differential response rates in the two samples might bias our results. We varied the magnitude of the relationship between the outcome and the probability, or odds, of censoring to determine how differential the outcome would need to be between those who responded and those who did not before our observed results were likely different from an unbiased estimate. Our first simulation assumed no relationship between the censoring indicator and the 12-month outcome while our subsequent simulations gradually increased the strength of the relationship between the censoring indicator and the 12-month outcome

such that youth in the ATM trial with worse outcomes were more likely to be censored (excluded) than youth with better outcomes.

The magnitude of the relationship between censoring and the 12-month outcome in the simulations is described by traditional effect sizes which compare the standardized mean outcome of the simulated non-censored cases (i.e., “responders”) in the ATM study to the standardized mean outcome of the simulated censored (i.e., “non-responders”) in the ATM study. Thus, effect sizes differences greater than 0 indicate that youth who were censored had worse mean outcomes than youth who were not censored and we use effect size guidelines specified in Cohen (1992) to determine whether the magnitude of the relationship between follow-up and outcomes is small, medium, or large (0.20, 0.50, and 0.80, respectively). One hundred simulations were performed for each assumed magnitude.

3. Results

3.1. Participant characteristics

Table 1 shows a comparison of our analytic sample before and after propensity score weighting across the outcome measures and the 22 pre-treatment variables used in the weighting strategy. The table shows values for the youth receiving MET/CBT-5 in EAT who had 12-month follow-up data and the unweighted and weighted ATM sample with 12-month follow-up data. Differences between samples were identified for variables that had absolute standardized mean differences greater than 0.25. As shown, EAT youth looked different from ATM youth on a number of variables before weighting. The ATM youth had higher baseline means on the following continuous outcomes: EPS, IAS, and SPS. Also, before weighting, the ATM sample consisted of a greater percentage of White Caucasian youth whereas the MET/CBT-5 sample from EAT consisted of a greater percentage of Hispanic youth. The ATM youth had higher mean scores on many of the substance and other related risk factor scales: *Crime and Violence*, *Substance Problem* (past year), *Substance Dependence* (past year), *Environmental Risk* (time spent with people using alcohol/drugs, involved in illegal activity, who argue, who are not in school or work and who have never been in treatment), *Training Activity* (days in school or training accounting for days in trouble, missed or suspended), *Social Environmental Risk* (acquaintances involved in illegal activities or substance use and are not employed or in school, attending self-help groups), and the mean maximum number of days in controlled environment.

After weighting, the EAT and ATM youth look very similar on almost all measures except the percentage female and the substance dependence scale in the past year whose absolute SMDs were only just above the 0.25 cut-off ($=0.253$ in both cases). The effective sample size of ATM was reduced to 47 which signals that the two groups were initially very different. However, weighting accounts for these differences and this effective sample size still provides sufficient power to detect significant differences between the EAT and weighted ATM samples (see Figures 1 and 2).

3.2. Effectiveness of MET/CBT-5 vs standard outpatient treatment

Our treatment effect estimates are presented graphically in Figures 1 (binary outcomes) and 2 (continuous outcomes). These estimates demonstrate that youth who received MET/CBT-5 under the EAT initiative fared better with respect to 12-month outcomes than similar youth who received treatment at one of the three ATM outpatient programs. At 12-months, youth receiving MET/CBT-5 in the EAT initiative had significantly higher rates of recovery when compared to similar youth in the ATM programs (mean difference 23.2%; 95% CI = 11.7%, 34.7%). Youth receiving MET/CBT-5 also had significantly lower mean values on the EPS (mean difference = -0.10 ; 95% CI = -0.15 , -0.04), the IAS (mean difference = -0.08 ; 95%

CI = -0.12, -0.04), the SPS (mean difference = -0.08; 95% CI = -0.13, -0.02), and the SFS (mean difference = -0.03; 95% CI = -0.06, 0.00) than similar clients receiving ATM care. Each of the differences corresponds to the following effect size differences: 0.64 for EPS, 0.80 for the IAS, 0.44 for the SPS, and 0.27 for the SFS. These effect sizes suggest the magnitude of the difference between youth receiving MET/CBT-5 in the EAT initiative and similar clients receiving treatment in the ATM initiative was quite large for all outcomes except substance frequency where the difference would be considered to be a medium effect size.

3.3. Sensitivity Analysis

The bold lines in Figure 3 show the range of effect size differences on outcomes that would be required between simulated ATM responders and non-responders for the 95% simulation bands to contain zero as a possible mean difference between the ATM and EAT (MET/CBT-5) groups (note that for Recovery and IAS, the simulation band extends beyond 2, which represents a very large effect size). The circles denote the exact effect size difference on an outcome for which the mean treatment effect estimate comparing the ATM and EAT (MET/CBT-5) groups across the 100 simulations is zero (i.e., null or no treatment effect). For example, using the SPS, we find that effect size differences ranging from 0.09 to 1.30 would yield 95% simulation bands that contain zero and that an effect size difference of 0.73 results in an estimated mean treatment effect of zero across all 100 simulations. In other words, while we only need a small effect size of 0.09 to get a null effect as possible, we would need a very strong relationship between follow-up and worse 12-month SPS scores (i.e., effect sizes greater than 1.3) to find evidence that youth receiving ATM care would have fared better than youth receiving MET/CBT-5. Moreover, for both recovery and the IAS, we were unable to find an effect size that would result in ATM youth having significantly better outcomes than youth receiving MET/CBT-5. For example, even in the extreme case where we censor the 46% of ATM youth who had the *best* values on the IAS scale, we do not find statistically significant evidence that comparable ATM youth would have fared the same as or better than those who received MET/CBT-5; in this case, our estimated mean difference between ATM and MET/CBT-5 youth remains above 0 and equals 0.01 (95% CI = -0.01, 0.03; results not shown).

4. Discussion

4.1 Summary of Findings

Prior to the current study, there was limited information about the effectiveness of MET/CBT-5 as delivered in typical delivery settings relative to standard approaches for addressing adolescent substance use. Moreover, the available information focused on clients who participated in or looked similar to those in the CYT field experiment (M. Dennis et al., 2004; Godley et al., 2010; Ramchand et al., 2011). Since this group of clients may not necessarily generalize to adolescents who typically present to outpatient programs (Tims et al., 2002) and because the resources available to community-based treatment programs may differ from those available in experimental settings (Weisz et al., 1995; Weisz et al., 1992), we conducted the current study to assess how youth who received community-based MET/CBT-5 fared relative to what CSAT considered to be “exemplary” outpatient treatment. Consistent with our hypotheses, we found evidence to suggest that those receiving MET/CBT-5 generally fared as well as or better to comparable youth receiving “exemplary” outpatient care.

The strengths of these findings were manifested in several ways. First, after weighting the ATM sample to be comparable to the EAT sample on a number of pretreatment covariates, the effective sample size of the ATM group was reduced to 47. Even with this small

effective sample size, a comparison across five of the six outcomes showed statistically significant differences favoring youth receiving MET/CBT-5, demonstrating large effect sizes between the EAT and ATM groups. In exception was a measure of institutionalization where both groups reported relatively low rates at the 12-month follow-up (i.e., 15% in the EAT group and 18% in the ATM group). Second, we were able to confirm these findings even after conducting a sensitivity analyses that accounted for differential response rates between the two samples. Specifically, we examined how strong the relationship between poor outcomes and follow-up would need to be to make outcomes equivalent between the two groups and, more critically, to reverse our findings in favor of the standard outpatient treatment arm (ATM). Results from sensitivity analyses addressing concerns about the low follow-up rate suggest that while modest differences in outcomes between nonrespondents and respondents in the MET/CBT-5 condition would yield no significant differences between the two groups on two of the six outcomes (substance problems and substance frequency scales), very large differences in outcomes would be required for comparable youth receiving exemplary community based care to have fared *better* on all of the outcomes as compared to MET/CBT-5.

4.2 Clinical and policy implications and future research

There are several important clinical and policy lessons to be drawn from this study as well as recommendations for future research. First, the 12-month outcomes observed among the EAT sample showed that approximately 50% met the recovery criteria relative to similar youth in standard community-based care at 27%. This 50% recovery figure surpasses the proportion that was “recovered” at 12-months in the CYT field experiment of MET/CBT-5, which was 35% (M. Dennis et al., 2004). However, it is possible that the high rate of nonresponse among the EAT sample may contribute to these relatively positive findings. Also, it is important to consider the characteristics of the youth served in the EAT initiative when interpreting these results. Before weighting the two datasets, we found that EAT youth exhibited less severe substance use symptoms and related risk factors than the youth that participated in the ATM initiative. While MET/CBT-5 was shown to be effective, our findings do not suggest that MET/CBT-5 is a superior approach for treating all youth referred to outpatient treatment, but may be targeted for youth that are deemed appropriate for a brief cannabis use treatment program. The evidence presented here suggests that MET/CBT-5 is likely to lead to outcomes similar or slightly better than standard outpatient care for less severe cases of cannabis use that present to outpatient treatment programs in the United States. MET/CBT-5’s low cost relative to other treatments previously studied make it an appealing option, but note that future research should address that there is still room for improvement in adolescent treatment quality, as 50% of the EAT youth reported continued use at 12 months.

Second, our findings regarding the relatively low response rate observed in the EAT sample suggest the need for continued support to optimize response rates in dissemination trials such as the EAT initiative. As part of receiving support from SAMHSA/CSAT to deliver evidence-supported treatments like MET/CBT-5, funded sites are expected to monitor and track 80% of clients that participate in the intake interview for up to 12 months. Under the EAT initiative, the highest response rate achieved across the 37 sites for cases eligible for follow-up during the grant period was 72% and the median response rate was much lower (41%), demonstrating the challenges of tracking substance using populations which have been well documented (Desmond, 1995; Scott, 2004; Walton, Ramanathan, & Reischl, 1998). These findings suggest that the resources in this initiative devoted to follow-up were insufficient. Allocating adequate resources, for instance, providing resources for the interim tracking of clients and access to professionally trained field interview staff is needed. Also,

allowing for follow-up assessments to be conducted by staff nonaffiliated with the treatment center may help to boost response rates by reducing client self-reporting concerns.

Third, low response rates create uncertainty about the likely size of true causal effects. Recognizing the need to address the differential follow-up rates between the two samples, we used a sensitivity analyses that allows us to understand the strength of our study findings (Griffin et al., under review). In our opinion, this strategy represents one of the best analytic options currently available. Other strategies were considered such as use of attrition weights and principal stratification. In our sample, attrition weights are not appropriate because it is highly unlikely that the outcomes among those who responded could be extrapolated to draw inferences about the outcomes among those who did not. Principal stratification, on the other hand, is a model-based approach that works best when one has covariates that are highly predictive of loss to follow-up and that improve both precision and identifiability in the model fit (Griffin, McCaffrey, & Morral, 2008; McCaffrey, Morral, Ridgeway, & Griffin, 2007); in our case, such covariates were not found. We also examined outcome data collected at the 6-month interview to potentially address the differential follow-up concern. However we discovered differential follow-up rates were also an issue at that time point (i.e., the EAT response rate was 62% as compared to a 94% response rate in the ATM sample for a 32% difference). Moreover, a similar pattern of results emerged between the weighted EAT and ATM groups to those demonstrated at 12-months. Therefore, we recommend that similar studies that compare outcomes across samples with differential follow-up rates might consider using this approach to help discern the strength of the findings.

4.3 Limitations

It is worthwhile to point out certain limitations to our study. First, we considered that a possible source of the relatively positive outcomes for youth who participated in the EAT initiative may be that they received more treatment during the 12-month period than the comparison participants in the exemplary community-based programs. To address this alternative explanation, we examined self-reported receipt of treatment following the delivery of MET/CBT-5 and found that in fact, the ATM youth were more likely to report more outpatient treatment at 6-months than the EAT youth, and there were no differences between the two groups at 12 months. A second concern about the study is whether the study samples were truly balanced. With propensity weighting, there is always the possibility that unmeasured (i.e., unobserved) confounders may influence the study results. Although our analytic strategy succeeded in identifying a comparison group of youth in standard outpatient community-based care using a comprehensive set of demographic and known risk factors, the possibility remains that there may have been some important unobserved differences between the groups not accounted for in our analyses. Moreover, there is a possibility of measurement error among those variables we purport to adjust for in our analyses. Third, it is also important to note that the sensitivity analyses only provide simulated outcomes and uncertainty remains in the study findings even when this approach is utilized, particular because it is was not possible to ascertain how follow-up and outcomes are related. Fourth, our findings do not directly address the reasons for the superior effectiveness of MET/CBT-5 relative to standard community care. Because we did not have access to the quality of treatment delivered in the two study conditions, we are limited in attributing the source of the observed effects. This is an important question for future research. Finally, the outcomes from the two studies were measured more than five years apart when the context for using drugs and standard treatment might have differed. Although there were notable declines in cannabis use over this period, the trends in daily use were flatter which may be more indicative of the population under study (i.e., youth in treatment) (Johnston, O'Malley, Bachman, Schulenburg, 2011). It is possible that “standard”

community based care may have evolved over time and current approaches may outperform those observed in the ATM study.

4.4. Summary

This study helps to address the research to practice gap by providing information about the relative effectiveness of an evidence-supported treatment as delivered in usual care practice settings as compared to standard community based approaches to address adolescent substance use. Understanding how well evidence-supported treatments perform when delivered in usual care practice settings is an important question to address with the increasing movement towards their use in these settings. The study results generally indicate that MET/CBT-5, as delivered to youth across 37 community based outpatient treatment programs, resulted in outcomes that are at least as good as, if not slightly better than, similar adolescents receiving longer community-based outpatient treatment. Although these results are favorable for MET/CBT-5, the observed 12-month outcomes were far from ideal and two important caveats to these findings remain. One is that it is possible that the differential follow-up in the two samples biased our results in favor of MET/CBT-5, although sensitivity analyses indicate that a strong relationship would be required between follow-up and outcomes for comparable youth receiving exemplary outpatient care to have achieved better outcomes. Second is that although we carefully weighted the samples so that they were balanced on a number of important pre-treatment characteristics, there is the possibility that unobserved differences between the two samples or measurement error in the observed variables biased the study results. Finally, the study findings indicate the investment value in comprehensive assessments and adequate follow-up in these settings in order to draw conclusions about treatment effectiveness.

Acknowledgments

This research was supported by National Institute on Drug Abuse Grant R01DA017507-05 (to Andrew Morral, principal investigator). The authors thank Mike Dennis, and other members of the RAND Casemix Advisory Board for comments they provided on earlier drafts of this research. Preliminary results based on this research were presented on October 28, 2009, at the Addiction Health Services Research Conference in San Francisco, California.

References

- Battjes RJ, Gordon MS, O'Grady KE, Kinlock TW, Katz EC, Sears EA. Evaluation of a group-based substance abuse treatment program for adolescents. *Journal of Substance Abuse Treatment*. 2004; 27:123–134. [PubMed: 15450645]
- Battjes, RJ.; Sears, EA.; Katz, EC.; Kinlock, TW.; Gordon, M. The Epoch Project Team. Evaluation of a group-based outpatient adolescent substance abuse treatment program. In: Stevens, SJ.; Morral, AR., editors. *Adolescent substance abuse treatment in the United States*. Binghamton, NY: The Haworth Press, Inc.; 2003. p. 81-101.
- Child Trends. Marijuana Use. 2010. Downloaded 4/19/11 from: www.childtrendsdatabank.org/?q=node/278
- Cohen JA. A power primer. *Psychological Bulletin*. 1992; 112:155–159. [PubMed: 19565683]
- Dennis, M.; Babor, T.; Diamond, G.; Donaldson, J.; Godley, SH.; Titus, JC. The Cannabis Youth Treatment (CYT) experiment: Preliminary findings. Rockville, MD: Center for Substance Abuse Treatment, Substance Abuse and Mental Health Services Administration, Department of Health and Human Services; 2000.
- Dennis M, Babor TF, Roebuck MC, Donaldson J. Changing the focus: The case for recognizing and treating cannabis use disorders. *Addiction*. 2002; 97:4–15. [PubMed: 12460125]
- Dennis M, Godley SH, Diamond G, Tims FM, Babor T, Donaldson J, et al. The Cannabis Youth Treatment (CYT) study: Main findings from two randomized trials. *Journal of Substance Abuse Treatment*. 2004; 27:197–213. [PubMed: 15501373]

- Dennis, ML. Global appraisal of individual needs (GAIN) Administration guide for the GAIN and related measures (Version 1299). Bloomington, IL: Chestnut Health Systems; 1999.
- Dennis, ML.; Dawud-Noursi, S.; Muck, RD.; McDermeit, M. The need for developing and evaluating adolescent treatment models. In: Stevens, SJ.; Morral, AR., editors. Adolescent substance abuse treatment in the United States. New York: The Haworth Press, Inc.; 2003. p. 3-34.
- Desmond DP, Maddux JF, Johnson TH, Confer BA. Obtaining follow-up interviews for treatment evaluation. *Journal of Substance Abuse Treatment*. 1995; 12(2):95-102. [PubMed: 7623396]
- Diamond G, Godley SH, Liddle HA, Sampl S, Webb C, Tims FM, et al. Five outpatient treatment models for adolescent marijuana use: A description of the cannabis youth treatment interventions. *Addiction*. 2002; 97:70-83. [PubMed: 12460130]
- Godley SH, Garner BR, Smith JE, Meyers RJ, Robert J, Godley MD. Meyers and Associates. A large scale dissemination and implementation model for Evidence-based treatment and continuing care. *Clinical Psychology Science and Practice*. 2011; 18:67-83. [PubMed: 21547241]
- Godley SH, Garner BR, Passetti LL, Funk RR, Dennis ML, Godley MD. Adolescent outpatient treatment and continuing care: Main findings from a randomized clinical trial. *Drug and Alcohol Dependence*. 2010; 110:44-54. [PubMed: 20219293]
- Godley SH, Jones N, Funk R, Ives M, Passetti LL. Comparing outcomes of best-practice and research-based outpatient treatment protocols for adolescents. *Journal of Psychoactive Drugs*. 2004; 36:35-48. PMID: 15152708. [PubMed: 15152708]
- Godley, SH.; Risberg, R.; Adams, L.; Sodetz, A. Chestnut Health Systems' Bloomington Outpatient and Intensive Outpatient Program for adolescent substance abusers. In: Stevens, SJ.; Morral, AR., editors. Adolescent substance abuse treatment in the United States. Binghamton, NY: The Haworth Press, Inc.; 2003. p. 57-80.
- Griffin BA, McCaffrey DF, Morral AR. Evaluation of principle stratification with an application to modelling treatment outcomes for adolescents in substance abuse treatment programs. *Annals of Applied Statistics*. 2008; 2(3):1034-1055. [PubMed: 19779599]
- Griffin BA, McCaffrey DF, Ramchand R, Hunter SB, Suttrop MJ. Assessing the sensitivity of treatment effect estimates to differential follow-up rates: Implications for translational research. Under review at *Drug and Alcohol Dependence*.
- Hirano K, Imbens G, Ridder G. Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*. 2003; 71(4):1161-1189.
- Imbens GW. Nonparametric estimation of average treatment effects under exogeneity: A review. *The Review of Economics and Statistics*. 2004; 86(1):4-29.
- Institute of Medicine (U.S.). Committee on Crossing the Quality Chasm. Washington, DC: National Academies Press; 2006. Improving the quality of health care for mental and substance-use conditions: Quality chasm series.
- Johnston, LD.; O'Malley, PM.; Bachman, JG.; Schulenberg, JE. Monitoring the future national results on adolescent drug use: Overview of key findings. Vol. 2010. Ann Arbor, MI: Institute for Social Research, The University of Michigan; 2011.
- Lee BK, Lessler J, Stuart EA. Improving propensity score weighting using machine learning. *Statistics in Medicine*. 2009; 29:337-346. [PubMed: 19960510]
- Liddle HA, Rowe CL, Quille TJ, Dakof GA, Mills DS, Sakran E, et al. Transporting a research-based adolescent drug treatment into practice. *Journal of Substance Abuse Treatment*. 2002; 22:231-243. [PubMed: 12072167]
- McCaffrey DF, Morral AR, Ridgeway G, Griffin BA. Interpreting treatment effects when cases are institutionalized after treatment. *Drug and Alcohol Dependence*. 2007; 89:126-138. [PubMed: 17275215]
- McCaffrey DF, Ridgeway G, Morral AR. Propensity score estimation with boosted regression for evaluating causal effects in observational studies. *Psychological Methods*. 2004; 9:403-425. [PubMed: 15598095]
- Mee-Lee, D.; Shulman, GD.; Fishman, M.; Gastfriend, DR.; Griffith, JH., editors. ASAM patient placement criteria for the treatment of substance-related disorders: (ASAM PPC-2R). Rev. 2nd ed.. Chevy Chase, MD: American Society of Addiction Medicine, Inc.; 2001.

- Miller W, Sorensen J, Selzer J, Brigham G. Disseminating evidence-based practices in substance abuse treatment: A review with suggestions. *Journal of Substance Abuse Treatment*. 2006; 31(1):25–39. [PubMed: 16814008]
- Morral AR, McCaffrey DF, Ridgeway G. Effectiveness of community-based treatment for substance-abusing adolescents: 12-month outcomes of youths entering Phoenix Academy or alternative probation dispositions. *Psychology of Addictive Behaviors*. 2004; 18:257–268. [PubMed: 15482081]
- Morral, AR.; McCaffrey, DF.; Ridgeway, G.; Mukherji, A.; Beighley, C. The relative effectiveness of 10 adolescent substance abuse treatment programs in the United States (RAND/TR-346-CSAT). Santa Monica, CA: RAND; 2006.
- R Development Core Team. Vienna, Austria: R Foundation for Statistical Computing; 2011. R: A language and environment for statistical computing. Available at: <http://www.R-project.org>
- Ramchand R, Griffin BA, Suttrop M, Harris KM, Morral AR. Using cross-study design to assess the efficacy of MET/CBT5 in treating adolescents with cannabis-related disorders. *Journal of Studies on Alcohol and Drugs*. 2011; 72(3):380–389. PMID: 19737487. [PubMed: 21513674]
- Ridgeway G, McCaffrey DF, Morral AR. Twang: Toolkit for weighting and analysis of nonequivalent groups (Version 1.2–2). 2010 [Computer software and reference manual]. Available at: <http://cran.r-project.org/web/packages/twang/index.html>.
- Rosenbaum, PR. *Observational Studies*. New York: Springer-Verlag; 1995.
- Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983; 70:41–55.
- Sampl, S.; Kadden, R. Center for Substance Abuse Treatment (U.S.). Motivational enhancement therapy and cognitive behavioral therapy for adolescent cannabis users: 5 sessions. 2001. Available at <http://purl.access.gpo.gov/GPO/LPS49183>
- Scott CK. A replicable model for achieving over 90% follow-up rates in longitudinal studies of substance abusers. *Drug and Alcohol Dependency*. 2004; 74:21–36.
- Stevens, SJ.; Estrada, BD.; Carter, T.; Reinardy, L.; Seitz, V.; Swartz, T. The Teen Substance Abuse Treatment Program: Program design, treatment issues, and client characteristics. In: Stevens, SJ.; Morral, AR., editors. *Adolescent substance abuse treatment in the United States*. Binghamton, NY: The Haworth Press, Inc.; 2003. p. 37–56.
- Stevens SJ, Schwebel R, Ruiz B. The seven challenges: An effective treatment for adolescents with co-occurring substance abuse and mental health problems. *Journal of Social Work Practice in the Addictions*. 2007; 7:29–49.
- Substance Abuse and Mental Health Services Administration. Office of Applied Studies. Treatment episode data set (TEDS). 1997–2007. National admissions to substance abuse treatment services, DASIS series: S-47 (09-4379). Rockville, MD: Author; 2009.
- Tims FM, Dennis ML, Hamilton N, Buchan BJ, Diamond G, Funk R, et al. Characteristics and problems of 600 adolescent cannabis abusers in outpatient treatment. *Addiction*. 2002; 97:46–57. [PubMed: 12460128]
- Walton MA, Ramanathan CS, Reischl TM. Tracking substance abusers in longitudinal research: understanding follow-up contact difficulty. *American Journal of Community Psychology*. 1998; 26:233–253. [PubMed: 9693691]
- Weisz JR, Donenberg GR, Han SS, Weiss B. Bridging the gap between laboratory and clinic in child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology*. 1995; 63:688–701. [PubMed: 7593861]
- Weisz JR, Weiss B, Donenberg GR. The lab versus the clinic: Effects of child and adolescent psychotherapy. *American Psychologist*. 1992; 47:1578–1585. [PubMed: 1476328]
- Wooldridge, JM. *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press; 2002.

Figure 1. 12-month dichotomous outcomes (recovery and institutionalization) for the EAT (MET/CBT-5) and ATM (control group) samples, after propensity score weighting.

Figure 2. 12-month continuous outcomes (emotional problem, substance frequency, substance problem, and illegal activities scales) for the EAT (MET/CBT-5) and ATM (control group) samples, after propensity score weighting.

Figure 3.

Bolded lines show the range of effect size differences on outcomes that would be required between simulated ATM responders and non-responders for the 95% simulation bands to contain zero as a possible mean difference between the ATM and EAT (MET/CBT-5) groups (note that for Recovery and IAS, the simulation band extends beyond 2, which represents a very large effect size). The circles denote the exact effect size difference on an outcome for which the mean treatment effect estimate comparing the ATM and EAT (MET/CBT-5) groups across the 100 simulations is zero (i.e., null or no treatment effect).

Note: SPS=Substance Problem Scale; SFS=Substance Frequency Scale; EPS=Emotional Problem Scale, IAS=Illegal Activity Scale 38

Table 1

Mean values on pre-treatment characteristics of the EAT (MET/CBT-5) sample and the weighted and unweighted standard outpatient treatment sample

	EAT n=2293	ATM unweighted N=431	weighted ESS=47
Continuous outcomes (Baseline Values)			
Emotional Problems Scale-Proportional	0.206	0.278 *	0.240
Illegal Activity Scale	0.087	0.226 *	0.109
Substance Problem Scale (past month)	2.691	3.146	3.032
Substance Frequency Scale	0.108	0.147 *	0.114
Other client characteristics at baseline			
<i>Demographics/Living Conditions</i>			
Female	0.317	0.213	0.200 *
White/Caucasian	0.498	0.671 *	0.573
Black/African-American	0.082	0.132	0.073
Hispanic	0.248	0.100 *	0.241
Other	0.172	0.095	0.111
Current Living Situation: House/apartment	0.889	0.849	0.942
Friend/relatives house/unsupervised dorm	0.050	0.104	0.033
Other	0.060	0.044	0.025
<i>Substance, mental health, and assoc. behaviors</i>			
Internal Mental Distress Scale--4 factors	7.761	7.733	8.452
Crime & Violence Scale	6.110	9.069 *	6.370
Days (Past 90) drunk or high most of day	11.557	14.019	11.253
Substance Problem Scale (past year)	6.278	7.884 *	7.224
Substance Dependence Scale (past month)	0.877	1.032	0.861
Substance Dependence Scale (past year)	2.314	3.079 *	2.87 *
When last time gambled for money, etc	0.863	1.140	1.037
<i>Environmental Risk</i>			
Environment Risk Scale	35.608	31.399 *	33.927
Training Activity Scale	0.687	0.557 *	0.672
Social Environmental Risk Index	12.970	14.452 *	13.784
<i>Treatment/Institutionalization</i>			
Ever attended self-help group	0.116	0.313 *	0.174
Mental health treatment	0.359	0.444	0.412
Maximum number of days in a controlled environment	4.465	13.179 *	3.799

Note:

* absolute SMD>0.25; ESS=Effect Sample Size