

Dennis, M. L., Godley, S.H., Diamond, G., Tims, F.M., Babor, T., Donaldson, J., Liddle, H., Titus, J.C., Kaminer, Y., Webb, C., Hamilton, N., and Funk, R. (in press). Appendix for - The Cannabis Youth Treatment (CYT) Study: Main Findings from Two Randomized Trials. *Journal of Substance Abuse Treatment*. Retrieved from www.chestnut.org/li/downloads/Dennis_et_al_CYT_MF_Appendix.pdf .

This appendix provides site and cross site level data to document the results reported in the main findings paper, a discussion of several concerns and alternative analyses that have been conducted to address common questions about the findings, and the results of more complicated analyses that differ somewhat from what is reported in the main paper.

A.1 Documentation of Site Level Analysis of Clinical Outcomes and Cost Effectiveness

Table A1 at the end of this appendix displays the results of the cost, clinical outcome, and cost effectiveness analyses for each condition across sites (the focus of the main paper) and within each site. As shown in Table 3 in the main findings paper, the treatment costs are in 1999 dollars and have been previously reported (French et al., 2002). The next two columns show the results of the two clinical outcomes: the average days abstinent over the 12 month follow-up period and percent of participants in recovery at the end of the study with the across site data matching what was summarized in Figure 2 in the main paper. In the last two columns are the two economic measures: the cost per day of abstinence over the 12 month follow-up and the cost per person in recovery at the end of the study. For each column, the table shows the mean, Cohen's f , and the probability of this degree of group differences occurring by chance at $\alpha < .05$ overall, as well as for site effects.

Notice the large site difference in the cost per day abstinent in Trial 2 (Site 3=\$5.15 vs. Site 4=\$12.23) and differences in the pattern by site. The Site 4 (CHOP) results parallel the cross-site findings, with ACRA having a lower average cost per day of abstinence than MET/CBT5 or MDFT (\$8.09 vs. \$15.83 vs. \$12.79; $f=0.23$, $p<.05$). The differences between the ACRA and MDFT were significant in pair-wise Tukey range testing, but MET/CBT5 was between them in terms of cost effectiveness and was not significantly different than either. In Site 3, MET/CBT5 was less expensive than ACRA or MDFT in pair-wise comparisons (\$839 vs. \$1,237 vs. \$1,428), but also had a (non-significant) trend to be less effective (257 vs. 281 vs. 271 days abstinent). While it had a lower average cost per day of abstinence (\$3.86 vs. \$5.36 vs. \$5.94; $f=0.21$, $p<.05$), when we controlled for the average performance for the site, MET/CBT5 was actually less cost-effective (i.e., [condition cost-average cost] / [condition effect-average effect]) than ACRA (\$26.34 vs. \$4.10 per additional day of abstinence over average).

A.2 Alternative Analyses

Several additional analyses were conducted to examine alternative explanations that have been raised as possible alternative explanations for the findings. One possible explanation was that there might have been inadequate power to detect greater differences among the

interventions. All of the main analyses conducted for the main findings paper had over 90% power, though several of the within site and pair-wise comparison analyses as part of Tukey range testing dropped down to 80%.

A second explanation that has been raised regarding the finding that overall participants reported an increase in days of abstinence and after treatment a higher percent of adolescents were in recovery is that drug substitution may have occurred. Both the days of abstinence and percent in recovery variables considered use of alcohol, cannabis and other drugs. We also evaluated cannabis, alcohol and cocaine individually, with no change in the findings. We did find some small differences (2-3 days per quarter) in the days in a controlled environment, but these differences varied by site and condition and we believe this is better considered in terms of changes in the total cost to society of multiple high cost services simultaneously as reported in the cost-benefit paper (see French et al., 2003).

A third explanation that has often been suggested is that examining other outcome variables would reveal differences among treatment models. Across interventions, there were significant reductions from intake to 12 months in days of behavioral problems (-58%), family problems (-56%), arguing/ violence (-66%), illegal activity (-70%), and missing school (-40%). These changes were not significantly different by condition in 9 of 10 comparisons. While there was a significant difference by condition in Trial 1, it was primarily due to MET/CBT5 (-54%) having a quicker impact in the first three months of treatment and holding its gains, FSN (-63%) having a slower impact initially, but eventually doing better and MET/CBT12 (-43%) having the least impact initially, but eventually showing an impact. Even in this one case, however, the effect was still small (Cohen's $f = .1$).

A fourth explanation suggested is that certain interventions might be more effective for certain adolescents due to treatment by subject matching effects. No treatment matching effects were found on common classification schemes (gender, onset age, family history, externalizing disorders, internalizing disorders and temperament) in terms of substance use frequency, substance abuse problems, social support for substance use, family conflict, school problems and negative peer associations (see Babor et al., 2002).

A fifth explanation suggested that differential outcomes were obscured because the analyses did not adequately control for individual differences in characteristics, trajectory, and treatment dosage received. To address this concern, we conducted a more complicated analysis and did find some differences that are summarized below. JSAT's reviewers considered these analyses to be more complicated than warranted and revealing little additional information that was of clinical significance, but they are included here for those who are interested.

A.3 Results of Alternative Mixed Model Analysis

To better model individual differences we used a SPSS (2001) mixed-effects model allowing the a-intercept to be a random factor. To control for differences in the quantity of different treatment services received, a term for the number of days of therapy sessions -- nested within site and condition -- was included. To address the statistical "hinge" in the trend line of Figure 1, we modeled time effects with two orthogonal contrasts: a) a "treatment outcome" effect calculated by contrasting the intake value with the average value of a measure across all 4 follow-up waves (i.e., a contrast of -4 +1 +1 +1 +1 by observation wave); and, b) an "outcome stability" effect calculated by comparing the average values from early (3-months and 6-months) with later (9-months and 12-months) follow-up interviews (i.e., a contrast of 0 -1 -1 +1 +1 by

observation wave). This approach increases the observed eta square and effect size by 25-50% (depending on the variable) instead of attempting to fit the data to a single linear trend that ignores this hinge. Within each trial, site differences were modeled with a dummy variable. Reflecting the randomized block design, conditions were modeled as nested within site. To model dosage effects, we used the days attending therapy sessions, nested within condition and site (i.e., does more dosage in any model predict better outcomes?). Missing data were estimated in the mixed effects analyses using the restricted maximum likelihood (REML) method recommended by Little and Rubin (1989) for randomized trials. To reduce measurement error in the dependent variables, we switched from days abstinent to the GAIN's Substance Frequency Scale (SFS) and Substance Problem Scale (SPS). The SFS is based on the average percent of days during a 90 day period that an adolescent reports each of the following: days of "any" substance use, days of heavy substance use, days of problems from substance use, days of alcohol, cannabis, crack/cocaine, and heroin/opioid use. It has good internal consistency ($\alpha=.76$ to $.85$), test-retest reliability ($\rho=.94$) and is sensitive to change (Dennis et al., 2003; Dennis, Titus et al., 2002; Shane et al., 2003). The SPS is based on recency ratings (e.g., past month, 2-12 months ago, more than 12 months ago, never) of 16 symptoms: 7 corresponding to DSM-IV criteria for dependence, 4 for abuse, 2 for substance-induced health and psychological problems, and 3 that correspond to lower severity symptoms of use (hiding use, people complaining about use, weekly use). The past month SPS symptom count has good internal consistency (Cronbach $\alpha=.85$ to $.92$), test-retest reliability ($\rho=.70$ to $.81$), and has also been demonstrated to be sensitive to change (Dennis et al., 2003; Dennis, Titus et al., 2002; Shane et al., 2003). While this analysis produced similar effect sizes between conditions to those reported in the paper, the increased power of the above approach led to more "statistical significant" (i.e., reliably measured) differences being found. Tables A2 (SFS) and A3 (SPS), at the end of this document, summarizes the results over time by trial, site and condition, as well as the outcome and stability effects in terms of raw change, relative change, Cohen's effect size d for pre to post change and Cohen's effect size f for comparing change by therapy condition and is summarized below.

Across sites and conditions in Trial 1 (see top section of Tables), there was a significant, moderate-sized effect of treatment on reducing substance use (-34%, $d= -0.39$, $p<.05$) that was stable over the follow-up period (-1%, $d= -0.01$, n.s.d.). UCHC adolescents had higher rates of substance use across waves, but also had a larger treatment effect than PAR adolescents (-38% vs. -28%; $d= -0.51$ vs. -0.30 , $p<.05$) and had further reductions (vs. increases in Site 2) over the follow-up period (-15% vs. +15%; $d= -0.13$ vs. $+0.11$, $p<.05$). There were no significant differences by conditions in substance use frequency for either the treatment effect or stability effect analyses across the PAR sites. Across sites and conditions in Trial 1, there was a significant, moderate-sized effect of treatment on reducing substance problems (-46%, $d= -0.50$, $p<.05$) that showed further improvement over the course of the follow-up period (-25%, $d= -0.17$, $p<.05$). There were no significant "site differences" in the overall level of problems, treatment effects or the stability of the effects. There were significant differences by condition in the treatment effects for substance related problems in the treatment effects analysis. FSN and MET/CBT5 participants reported greater decreases in problems than MET/CBT12 participants (-51% vs. -50% vs. -33%; $d= -0.62$ vs. -0.53 vs. -0.35 ; $f=0.15$, $p<.05$). Though the pattern of these differences and effect sizes ($f=0.15$, $p<.05$; $f=0.14$, $p<.05$) was consistent within each site, the condition differences were not significant due to the smaller sample size/power for the within-site analyses.

Across sites and conditions in Trial 2 (see bottom section of Tables), there was a significant, moderate-sized effect of treatment on reducing substance use (-35%, $d = -0.47$, $p < .05$) that was stable over the follow-up period (-6%, $d = -0.05$, n.s.d.). CHS participants had higher rates of substance use across waves, but had similar treatment effects to CHOP participants in terms of relative change (-34% vs. -36%, $p < .05$), better treatment outcomes in terms of effect sizes ($d = -0.52$ vs. -0.42 , $p < .05$) and continued gains (vs. deterioration in CHS) over the follow-up period (-17% vs. +11%; $d = -0.18$ vs. $+0.07$, $p < .05$). While there were no significant differences in overall treatment outcomes by condition, there were moderately sized and statistically significant differences by conditions in terms of stability ($f = 0.19$, $p < .05$). Over the course of follow-up waves, substance use was further reduced for both ACRA (-10%; $d = -0.10$, $p < .05$) and MDFT participants (-11%; $d = -0.07$, $p < .05$), but remained unchanged for MET/CBT5 participants (+4%; $d = +0.02$, n.s.d.). Increased treatment dosage (within condition and site) was also significantly related to better outcomes. Though varying in magnitude, these patterns were replicated in each of the sites. Across sites and conditions in Trial 2, there was a significant, moderate-sized effect of treatment on substance related problems (-43%, $d = -0.49$, $p < .05$) that was stable over the follow-up period (-8%, $d = -0.05$, n.s.d.). CHOP adolescents had higher rates of substance related problems across waves, but similar treatment effects and outcome stability to those in the CHS site. While there were no significant differences by condition in treatment effects or outcome stability across sites, there were differences within site. At the CHS site, the treatment effect analysis revealed that ACRA participants reported the largest reductions in substance related problems compared to MDFT or MET/CBT5 (-54% vs. -38% vs. -28%; $d = -0.67$ vs. -0.38 vs. -0.28 ; $f = 0.18$, $p < .05$). At the CHOP site, the stability analysis revealed that MDFT participants reported further reductions in substance related problems during follow-up compared to ACRA or MET/CBT5 (-36% vs. +5% vs. +10%; $d = -0.29$ vs. $+0.04$ vs. $+0.06$; $f = 0.24$, $p < .05$). Within the site and intervention conditions, adolescents receiving more than the average number of sessions generally reduced their substance problems more than those receiving less than the average number of sessions. This generic dosage effect was significant across sites, at the CHS site, and had a trend toward significance in the CHOP site.

Table A1. Cost, Effectiveness, and Efficiency Analysis by Trial, Site and Condition^a

Site	Condition (n in analysis)	Episode Cost			Outcome Measures						Efficiency Measures					
		Cost (1999 Dollars)			Days Abstinent			% in Recovery at			Cost Per Days			Cost Per Person in		
		Mean	f	p	Mean	f	p	Mean	f	p	Mean	f	p	Mean	f	p
Trial 1																
Across Sites 1 & 2 (n=299)		\$ 1,861	0.78	\i	262	0.06		0.23	0.12	\i	\$ 8.79	0.48	\i	\$ 8,846	0.72	\h,i
	MET/ CBT5 (n=102)	\$ 1,113			269			0.28			\$ 4.91			\$ 3,958		
	MET/ CBT12 (n=95)	\$ 1,185			256			0.17			\$ 6.15			\$ 7,377		
	FSNM (n=102)	\$ 3,246			260			0.22			\$ 15.13			\$ 15,116		
Site 1 - UCHC (n=131)		\$ 1,800	0.72	\g	245	0.03		0.21	0.18	\g	\$ 9.97	0.40	\g	\$ 10,034	0.67	\g
	MET/CBT5 (n=48)	\$ 1,112			249			0.32			\$ 5.75			\$ 3,495		
	MET/CBT12 (n=41)	\$ 1,187			242			0.13			\$ 7.67			\$ 9,257		
	FSNM (n=42)	\$ 3,200			244			0.18			\$ 17.04			\$ 18,284		
Site 2 - PAR (n=168)		\$ 1,909	0.84	\g	275	0.11		0.24	0.06		\$ 7.88	0.63	\g	\$ 7,912	0.81	\g
	MET/CBT5 (n=54)	\$ 1,114			287			0.25			\$ 4.17			\$ 4,369		
	MET/CBT12 (n=54)	\$ 1,183			266			0.20			\$ 5.00			\$ 5,914		
	FSNM (n=60)	\$ 3,279			271			0.25			\$ 13.80			\$ 12,899		
Trial 2																
Across Sites 3 & 4 (n=298)		\$ 1,655	0.54	\h	258	0.06		0.25	0.16		\$ 8.65	0.22	\h,i	\$ 7,615	0.78	\h,i
	MET/ CBT5 (n=99)	\$ 1,558			251			0.23			\$ 9.00			\$ 6,611		
	ACRA (n=100)	\$ 1,408			265			0.34			\$ 6.62			\$ 4,460		
	MDFT (n=99)	\$ 2,002			257			0.19			\$ 10.38			\$ 11,775		
Site 3 - CHS (n=150)		\$ 1,194	0.57	\g	271	0.10		0.27	0.20	\g	\$ 5.15	0.21	\g	\$ 4,769	0.61	\g
	MET/CBT5 (n=42)	\$ 839			257			0.18			\$ 3.86			\$ 4,673		
	ACRA (n=54)	\$ 1,237			281			0.40			\$ 5.36			\$ 3,123		
	MDFT (n=54)	\$ 1,428			271			0.22			\$ 5.94			\$ 6,490		
Site 4 - CHOP (n=148)		\$ 2,118	0.52	\g	244	0.03		0.23	0.11		\$ 12.23	0.23	\g	\$ 10,462	0.83	\g
	MET/CBT5 (n=57)	\$ 2,078			247			0.26			\$ 12.79			\$ 8,016		
	ACRA (n=46)	\$ 1,608			245			0.27			\$ 8.09			\$ 6,029		
	MDFT (n=45)	\$ 2,691			240			0.15			\$ 15.83			\$ 17,979		
Average Across Trials (n=597)		\$ 1,758	0.66		260	0.06		0.24	0.14		\$ 8.72			\$ 8,231		

^a Predicted as dependent variable at follow-up or $DV_{(post)} = DV_{(pre)} + SITE + COND(SITE)$

^b Cost in 1999 dollars estimates from French et al 2002; prorated based on % of mean days of formal treatment for site and condition.

^c Summed across follow-up waves 3 to 12 (with mean replacement within individual for missing waves).

^d Recovery is a dichotomous variable, so significance tested by logistic regression.

having zero cost; people with 0 days of abstinence being dropped; low correlation ($r = -.03$, n.s.d.) between cost and days abstinent; and rounding.

^f Individual cost divided by % in recovery for site & condition (a constant within cell); Estimates vary from simple division of means due to the low correlation between cost and recovery ($r = -.02$, n.s.d.) and rounding.

^g Significant difference ($p < .05$) between conditions within site

^h Significant difference ($p < .05$) between sites (Trial 1: UCHC=0, PAR=1; Trial 2: CHOP=0, CHS=1)

ⁱ Significant differences ($p < .05$) by condition (nested within site) across sites.

Table A2. Effects on Substance Frequency Scale (SFS) over Time by Site and Condition^a

Site	Condition (n in analysis)	Follow-up Wave				Averages			Treatment (Tx) Effects				Stability (St) Effects				
		Intake	3	6	9	12	3 to 12	3 & 6	9 & 12	Chng	R.C.	d _{Tx}	f _{Tx}	Chng	R.C.	d _{St}	f _{St}
Trial 1																	
Across Sites 1 & 2 (n=299) \b,d,e,f		0.15	0.10	0.10	0.10	0.10	0.10	0.10	0.10	-0.05	-34%	-0.39	0.09	0.00	-1%	-0.01	0.10
	MET/CBT5 (n=102)	0.15	0.09	0.11	0.09	0.09	0.09	0.10	0.09	-0.05	-35%	-0.40		-0.01	-10%	-0.06	
	MET/CBT12 (n=95)	0.15	0.11	0.12	0.10	0.11	0.11	0.10	0.10	-0.04	-26%	-0.30		-0.01	-7%	-0.07	
	FSNM (n=102)	0.16	0.10	0.09	0.10	0.11	0.10	0.09	0.11	-0.06	-38%	-0.47		0.01	14%	0.10	
Site 1 - UCHC (n=131) \b,g		0.20	0.13	0.12	0.11	0.11	0.12	0.13	0.11	-0.07	-38%	-0.51	0.10	-0.02	-15%	-0.13	0.13
	MET/CBT5 (n=48)	0.20	0.12	0.13	0.11	0.09	0.11	0.13	0.10	-0.08	-41%	-0.57		-0.03	-22%	-0.17	
	MET/CBT12 (n=41)	0.19	0.15	0.14	0.10	0.12	0.13	0.15	0.11	-0.05	-28%	-0.37		-0.03	-22%	-0.23	
	FSNM (n=42)	0.20	0.13	0.08	0.12	0.11	0.11	0.11	0.11	-0.09	-44%	-0.60		0.00	3%	0.05	
Site 2 - PAR (n=168) \b,g		0.12	0.07	0.09	0.09	0.10	0.09	0.08	0.09	-0.03	-28%	-0.30	0.07	0.01	15%	0.11	0.06
	MET/CBT5 (n=54)	0.10	0.05	0.09	0.07	0.09	0.08	0.07	0.08	-0.03	-26%	-0.26		0.01	8%	0.08	
	MET/CBT12 (n=54)	0.12	0.08	0.10	0.10	0.10	0.10	0.09	0.10	-0.03	-24%	-0.24		0.01	11%	0.09	
	FSNM (n=60)	0.14	0.08	0.09	0.10	0.11	0.09	0.08	0.10	-0.05	-33%	-0.38		0.02	24%	0.15	
Trial 2																	
Across Sites 3 & 4 (n=298) \b,d,e,f,i,k		0.17	0.12	0.11	0.11	0.11	0.11	0.11	0.11	-0.06	-35%	-0.47	0.12	-0.01	-6%	-0.05	0.19
	MET/CBT5 (n=99)	0.18	0.12	0.10	0.12	0.10	0.11	0.11	0.11	-0.06	-36%	-0.49		0.00	4%	0.02	
	ACRA (n=100)	0.16	0.11	0.11	0.10	0.10	0.10	0.11	0.10	-0.06	-36%	-0.46		-0.01	-10%	-0.10	
	MDFT (n=99)	0.18	0.14	0.10	0.11	0.11	0.12	0.12	0.11	-0.06	-32%	-0.45		-0.01	-11%	-0.07	
Site 3 CHS (n=150) \b,g,k		0.15	0.09	0.09	0.10	0.09	0.09	0.09	0.10	-0.05	-36%	-0.42	0.11	0.01	11%	0.07	0.15
	MET/CBT5 (n=42)	0.14	0.10	0.10	0.12	0.10	0.10	0.10	0.11	-0.03	-23%	-0.28		0.01	14%	0.09	
	ACRA (n=54)	0.14	0.09	0.10	0.08	0.08	0.09	0.09	0.08	-0.05	-36%	-0.38		-0.01	-11%	-0.09	
	MDFT (n=54)	0.17	0.08	0.08	0.11	0.10	0.09	0.08	0.11	-0.07	-45%	-0.56		0.03	35%	0.22	
Site 4 CHOP (n=148) \b,g,i		0.20	0.16	0.12	0.12	0.12	0.13	0.14	0.12	-0.07	-34%	-0.52	0.12	-0.02	-17%	-0.18	0.23
	MET/CBT5 (n=57)	0.20	0.14	0.11	0.13	0.11	0.12	0.12	0.12	-0.09	-43%	-0.67		0.00	-1%	-0.03	
	ACRA (n=46)	0.19	0.13	0.13	0.11	0.12	0.12	0.13	0.12	-0.07	-35%	-0.56		-0.01	-10%	-0.12	
	MDFT (n=45)	0.19	0.21	0.14	0.10	0.13	0.15	0.18	0.11	-0.04	-19%	-0.31		-0.07	-40%	-0.42	
Average Across Trials (n=597)		0.16	0.11	0.11	0.10	0.10	0.11	0.11	0.10	-0.06	-34%	-0.43		0.00	-4%	-0.03	

^a **Chng.** is Change=post-pre; **R.C.** is Relative Change calculated as (post-pre)/pre; **d_{Tx}** is Cohen's effect size d for within condition/site "treatment effect"; **f_{Tx}** is Cohen's effect size f for differences in treatment effects by condition(within site); **d_{St}** is Cohen's effect size d for within condition/site outcome "stability effect"; **f_{St}** is Cohen's effect size f for differences in stability effects by condition(within site).

^b Treatment effect (intake vs. average follow up) significant (p<.05)

^c Outcomes changing significantly (p<.05) over time (3 & 6 vs.9& 12)

^d Significant difference (p<.05) between site (Trial 1: UCHC=0, PAR=1; Trial 2: CHOP=0, CHS=1).

^e Treatment effects vary significantly (p<.05) by site.

^f Stability of outcomes vary significantly (p<.05) by site.

^g Significant baseline differences (p<.05) by condition (within site)

^h Treatment effects vary significantly (p<.05) by condition (within site)

ⁱ Stability of outcomes vary significantly (p<.05) by condition (within site).

^j Days of formal treatment (within condition and site) significantly (p<.05) related to baseline rates of use.

^k Treatment effect varies significant (p<.05) by days of formal treatment (within condition and site).

^l Stability of outcomes varies significant (p<.05) by days of formal treatment (within condition and site).

Table A3. Effects on Substance Problem Scale (SPS) over Time by Site and Condition^a

Site	Condition (n in analysis)	Intake	Follow-up Wave				Averages				Treatment (Tx) Effects				Stability (St) Effects			
			3	6	9	12	3 to 12	3 & 6	9 & 12	Chng	R.C.	d _{Tx}	f _{Tx}	Chng	R.C.	d _{St}	f _{St}	
Trial 1																		
Across Sites 1 & 2 (n=299) ^{b,c,h}		3.7	2.4	2.2	1.7	1.7	2.0	2.3	1.7	-1.7	-46%	-0.50	0.15	-0.6	-25%	-0.17	0.12	
	MET/CBT5 (n=102)	3.7	2.3	2.0	1.4	1.5	1.8	2.2	1.4	-1.8	-50%	-0.53		-0.8	-36%	-0.21		
	MET/CBT12 (n=95)	3.5	2.8	2.7	2.0	1.9	2.3	2.7	1.9	-1.2	-33%	-0.35		-0.7	-26%	-0.23		
	FSNM (n=102)	4.0	2.2	1.9	1.8	1.8	1.9	2.0	1.8	-2.1	-51%	-0.62		-0.2	-11%	-0.07		
Site 1 - UCHC (n=131) ^{b,c,g}		4.7	3.3	2.9	2.3	2.5	2.8	3.1	2.4	-1.9	-41%	-0.50	0.15	-0.7	-22%	-0.19	0.14	
	MET/CBT5 (n=48)	5.0	3.0	2.8	2.0	2.0	2.5	2.9	2.0	-2.5	-49%	-0.62		-0.9	-32%	-0.21		
	MET/CBT12 (n=41)	4.4	4.1	3.7	2.6	2.4	3.2	3.9	2.5	-1.2	-27%	-0.30		-1.2	-30%	-0.35		
	FSNM (n=42)	4.8	3.1	2.4	2.3	3.1	2.7	2.7	2.7	-2.1	-42%	-0.57		0.0	2%	0.00		
Site 2 - PAR (n=168) ^{b,c,g}		3.0	1.7	1.6	1.3	1.2	1.4	1.7	1.2	-1.5	-52%	-0.55	0.14	-0.5	-29%	-0.17	0.06	
	MET/CBT5 (n=54)	2.5	1.8	1.3	0.8	1.1	1.2	1.6	0.9	-1.3	-51%	-0.52		-0.7	-43%	-0.26		
	MET/CBT12 (n=54)	2.9	1.8	1.9	1.6	1.5	1.7	1.9	1.5	-1.2	-41%	-0.43		-0.4	-19%	-0.12		
	FSNM (n=60)	3.5	1.6	1.6	1.4	0.9	1.4	1.6	1.2	-2.1	-60%	-0.68		-0.4	-26%	-0.13		
Trial 2																		
Across Sites 3 & 4 (n=298) ^{b,d,k}		3.9	2.4	2.2	2.2	2.0	2.2	2.3	2.1	-1.7	-43%	-0.49	0.10	-0.18	-8%	-0.05	0.15	
	MET/CBT5 (n=99)	3.7	2.0	2.4	2.5	2.3	2.3	2.2	2.4	-1.4	-39%	-0.41		0.18	8%	0.06		
	ACRA (n=100)	4.4	2.4	2.1	2.2	2.0	2.2	2.3	2.1	-2.2	-50%	-0.64		-0.15	-6%	-0.06		
	MDFT (n=99)	3.5	2.8	2.0	2.0	1.8	2.2	2.4	1.9	-1.4	-40%	-0.43		-0.59	-24%	-0.16		
Site 3 - CHS (n=150) ^{b,g,h,k}		3.9	2.4	2.3	2.3	2.1	2.3	2.4	2.2	-1.7	-42%	-0.46	0.18	-0.19	-8%	-0.05	0.08	
	MET/CBT5 (n=42)	3.8	2.4	3.0	3.1	2.6	2.8	2.7	2.9	-1.1	-28%	-0.28		0.18	7%	0.05		
	ACRA (n=54)	4.7	2.6	2.1	1.8	2.0	2.1	2.4	1.9	-2.5	-54%	-0.67		-0.37	-16%	-0.12		
	MDFT (n=54)	3.2	2.3	2.0	2.2	1.6	2.0	2.1	1.9	-1.2	-38%	-0.38		-0.28	-13%	-0.07		
Site 4 - CHOP (n=148) ^{b,g,i}		3.8	2.4	2.1	2.2	2.0	2.2	2.2	2.1	-1.7	-44%	-0.54	0.01	-0.17	-8%	-0.05	0.24	
	MET/CBT5 (n=57)	3.7	1.8	2.0	2.2	2.0	2.0	1.9	2.1	-1.7	-47%	-0.52		0.18	10%	0.06		
	ACRA (n=46)	4.0	2.3	2.1	2.6	2.0	2.2	2.2	2.3	-1.8	-44%	-0.60		0.11	5%	0.04		
	MDFT (n=45)	3.9	3.6	2.0	1.7	2.0	2.3	2.8	1.9	-1.7	-42%	-0.48		-1.00	-36%	-0.29		
Average Across Trials (n=597)		3.81	2.44	2.19	1.98	1.88	2.12	2.31	1.93	-1.69	-44%	-0.50		-0.37	-16%	-0.11		

^a **Chng.** is Change=post-pre; **R.C.** is Relative Change calculated as (post-pre)/pre; **d_{Tx}** is Cohen's effect size d for within condition/site "treatment effect"; **f_{Tx}** is Cohen's effect size f for differences in treatment effects by condition(within site); **d_{St}** is Cohen's effect size d for within condition/site outcome "stability effect"; **f_{St}** is Cohen's effect size f for differences in stability effects by condition(within site).

^b Treatment effect (intake vs. average follow up) significant (p<.05)

^c Outcomes changing significantly (p<.05) over time (3 & 6 vs.9& 12)

^d Significant difference (p<.05) between site (Trial 1: UCHC=0, PAR=1; Trial 2: CHOP=0, CHS=1).

^e Treatment effects vary significantly (p<.05) by site.

^f Stability of outcomes vary significantly (p<.05) by site.

^g Significant baseline differences (p<.05) by condition (within site)

^h Treatment effects vary significantly (p<.05) by condition (within site)

ⁱ Stability of outcomes vary significantly (p<.05) by condition (within site).

^j Days of formal treatment (within condition and site) significantly (p<.05) related to baseline rates of use.

^k Treatment effect varies significant (p<.05) by days of formal treatment (within condition and site).

^l Stability of outcomes varies significant (p<.05) by days of formal treatment (within condition and site).