Published in final edited form as:

J Subst Abuse Treat. 2019 October ; 105: 64–70. doi:10.1016/j.jsat.2019.03.003.

Understanding site variability in a multisite clinical trial of a technology-delivered psychosocial intervention for substance use disorders

Leila M. Vaezazizi^{a,b}, Aimee N.C. Campbell^{a,b}, Martina Pavlicova^c, Mei-Chen Hu^b, Edward V. Nunes^{a,b,*}

^aDivision on Substance Use Disorders, New York State Psychiatric Institute, 1051 Riverside Drive, New York, NY 10032, USA

^bDepartment of Psychiatry, Columbia University Medical Center, 1051 Riverside Drive, New York, NY 10032, USA

^cDepartment of Biostatistics, Columbia University Mailman School of Public Health, 722 West 168th St, New York, NY 10032, USA

Abstract

Background: Significant fixed effects of site (main effects of site and/or site by treatment interactions) on primary outcome have been identified in the majority of studies performed by NIDA's National Drug Abuse Treatment Clinical Trials Network. While rarely explored, identifying patient- and site-level variables that are associated with site effects can provide information about the context in which outcome is optimized.

Methods: In a 10-site clinical trial that evaluated the effectiveness of a web-based psychosocial intervention compared to usual treatment of patients (N = 507) with substance use disorders, the primary outcome analysis revealed significant main effect of site, modeled as a fixed effect, on the outcome of abstinence (Campbell et al., 2014). In the current analysis, we use a two-level, hierarchical generalized linear model (HGLM) to identify patient- and site-level variables associated with abstinence outcome, while modeling site as a random factor.

Results: The site-specific percentage of patients abstinent in the last 4 weeks of the study varied from 6.1% to 40%. However, only 6.7% ($p = 0.08$) of variability in end-of-study abstinence was accounted for by site, indicating a small-moderate effect. Among patient-level predictors, older age (OR = 1.40; 95% CI = 1.15, 1.71; $p = 0.0009$), abstinence at baseline (OR = 2.77; 95% CI = 1.73, 4.45; $p < 0.0001$), and among site-level predictors, higher annual clinic admissions (OR = 1.28; 95% CI = 1.03, 1.59; $p = 0.0251$) were significantly associated with increased likelihood of abstinence. When controlling for these three variables in a HGLM, only patient age and abstinence

*Corresponding author at: Division on Substance Use Disorders, New York State Psychiatric Institute, 1051 Riverside Drive, New York, NY 10032, USA. Edward.nunes@nyspi.columbia.edu (E.V. Nunes).

Declaration of conflicts of interest

Dr. Edward Nunes has received medications for research studies from Alkermes, is an investigator for a Braeburn-funded study, and has provided unpaid consultation for Alkermes and Braeburn. All other authors have no conflicts of interest to declare.

at baseline remained significant, and random factor site explained only 1.4% of variability in end-of-study abstinence, a 79% reduction in magnitude.

Conclusions: The findings suggest that only some amount of variability in abstinence outcomes among sites can be explained by a combination of patient- and site-level variables. Our findings support the case that variability between sites is a natural phenomenon, and our methodological recommendation is that site be modeled as a random factor when analyzing multi-site clinical trials.

Keywords

Multicenter studies; Research design; Substance-related disorders; Treatment outcome; Community mental health services

1. Introduction

Primary outcomes in multisite clinical trials can vary considerably between sites. In more than half of the National Drug Abuse Clinical Trials Network (CTN) multi-site studies, significant site fixed effects (main effect of site and/or site by treatment interaction on the primary outcome) were identified (Nunes et al., 2010). Whether and how sites are accounted for in the model can affect the estimation of treatment effect and its standard error. Routinely, the effect of site is modeled either as fixed or random factor. Significant main effects of fixed factor site can indicate that some aspect of the clinical context, unique to the site, is driving the observed variability in outcome. For example, if patients with higher socioeconomic status and clinicians with more positive attitudes towards evidence-based treatments are associated with improved outcome, then sites with a higher proportion of these types of patients and clinicians may perform significantly better than other sites. Site by treatment interactions, when present, draw attention to other potential variables that influence treatment efficacy. Multi-site clinical trials do not routinely explore the association of site-level and patient-level variables with outcome measures of interest, but doing so can provide critical information about the context in which treatment efficacy and outcome are measured.

Substance use disorders are among the costliest public health problems and carry high morbidity and mortality nationwide (NDIC, 2011). Among substance abuse treatment centers, there is wide variability in treatment attitudes among clinicians who provide direct services to patients, which can influence the implementation of evidence-based treatments (Knudsen, Abraham, & Oser, 2011; Miller, Sorensen, Selzer, & Brigham, 2006). Technology-delivered interventions built on effective, evidence-based approaches, such as the community reinforcement approach, contingency management, and cognitive behavior therapy show potential for accelerating the uptake of evidence-based treatment (Carroll et al., 2014; Higgins et al., 1994; Higgins et al., 2003; Marsch, 2011). In a recently completed CTN study that measured the effectiveness of an evidence-based web-delivered intervention called Therapeutic Education System (TES), TES plus treatment as usual (TAU) was found to be significantly superior to TAU alone (Campbell et al., 2014). The site by treatment interaction was not significant, but there was a significant main fixed effect of site on the primary outcomes of abstinence and of retention in treatment. The current study aims

to identify any site- or patient-level variables are associated with treatment outcome, and discuss its implications for multi-site trial design.

We performed a secondary data analysis to determine whether the effect of site on the abstinence outcome remain significant when site is modeled as a random factor. Additionally, we investigated the influence of random factor site on abstinence outcome while accounting for patient- and site-level variables. In order to answer these questions, we used a two-level hierarchical generalized linear model (HGLM), a statistical model appropriate for nested data (patients are nested within sites), and utilized a modified, dichotomous outcome measure suitable for this model. We hypothesized that (1) patient-level variables including age, race/ethnicity, abstinence at baseline, and opioid use, and (2) site-level variables including counselor education, counselor recovery status, counselor attitudes towards new and evidence-based technology, average caseload, and annual number of clinic admissions would be associated with the modified primary outcome of end-of-study abstinence when site is modeled as random factor.

2. Material and methods

2.1. Design overview

The CTN-044 study, Web Delivery of Evidence-Based Treatment for Substance Use Disorders, was a randomized controlled trial of 12 weeks of treatment as usual (TAU) versus TAU plus the Therapeutic Education System (TES), whereby TES substituted for approximately 2 h of TAU. TAU varied across sites but on average was approximately 4–6 h of face-to-face group and individual counseling per week. Randomization was stratified by treatment site, patient's primary substance of abuse (stimulant versus nonstimulant), and patient abstinence status at study entry indicated by negative urine drug screen and breath alcohol screen results. In parallel with CTN-044, an ancillary study (Site Influences on Treatment Effects or CTN-044-A1) collected organizational and provider level data from program directors and counselors prior to patient randomization consisting of workforce characteristics and attitudes towards implementation of a new technology. Additional details of the methodology, design, and findings of CTN-044 and CTN-044-A1 have been published elsewhere (Buti et al., 2013; Campbell et al., 2014; Campbell et al., 2012). All participants read and signed an informed consent after receiving a complete description of the study. The protocols were approved by the institutional review board of New York State Psychiatric Institute, Oregon Health and Science University, and all participating clinical sites.

2.2. Sample

2.2.1. Sites—Of 41 nominated sites affiliated with the CTN, 10 were selected for participation in the study with the goal of maximizing patient diversity with respect to sex, race/ethnicity, and types of substances treated, as well as geographic diversity, consistent with the aims of an effectiveness trial (range of enrolled participants = 38–60). Sites selected to participate provided intensive outpatient treatment that consisted of at least 2 h of group or individual sessions weekly. Sites were excluded if they provided contingency management or the community reinforcement approach (both treatment components of

the experimental arm) as part of treatment as usual, or were opioid treatment programs. As described in the original study design, opioid treatment programs were excluded due to inherent differences in client population, treatment approach, and more homogeneous population as compared to programs that provide intensive outpatient substance use disorder treatment (Campbell et al., 2012).

2.2.2. Patients—Patients (N = 507) were eligible for the study if they were 18 years or older and had used illicit substances in the 30 days prior to study entry (or 60 days if coming out of a controlled environment such as a detoxification unit, hospital, or correctional facility), and were excluded if they were receiving medications for opioid use disorder. Patients with a primary alcohol use disorder were eligible as long as they used an illicit substance in the required timeframe.

2.2.3. Counselors and program directors—Counselors (n = 135) and program directors in the 10 participating program sites completed surveys assessing clinic and workforce characteristics, and counselor attitudes and beliefs towards new technology.

2.3. Measures

2.3.1. Patient-level measures—Patient-level variables used in this analysis include age, race, ethnicity (Hispanic vs. Non-Hispanic), opioid use (defined as primary opioid use, past-year opioid abuse or dependence), and abstinence at baseline. Patients were considered abstinent at baseline if their urine drug and alcohol breath test were negative at the time of randomization.

2.3.2. Site-level measures—Site-level variables included both counselor and organizational factors. Annual number of clinic admissions was the one organizational variable included in the analysis. Counselor variables include counselor education (% with Master's Degree or higher), recovery status, 12-step influence (not at all, somewhat, or very much), and average caseload. Attitudes towards contingency management were assessed by examining counselors' (1) perceived efficacy of the intervention ("Contingency management is effective in treatment substance abuse"), (2) intention to implement ("My program intends to use Contingency Management"), and (3) social norms around the intervention ("People who are important to me think that my outpatient program should use Contingency Management"). Attitudes were assessed on a 7-point scale with higher scores indicating less support for an intervention. The survey questions were similar to those used in other adaptations of Fishbein and Ajzen's theory of reasoned action, which posits that an individual will be more likely to perform a behavior if attitudes, social norms, and intentions are supportive of that behavior (Fishbein & Ajzen, 1975; Rieckmann, Daley, Fuller, Thomas, & McCarty, 2007). Contingency management has a wide evidence base and is a component of TES, the treatment intervention (Prendergast, Podus, Finney, Greenwell, & Roll, 2006). Counselor attitudes towards contingency management were selected as a proxy measure for the willingness to utilize an evidence-based, novel treatment in clinical practice.

2.3.3. Outcome measure—The primary abstinence outcome measure is based on self-reported drug and alcohol use assessed using the time line follow back method (TLFB;

Sobell and Sobell, 1992), and confirmed via urine drug and breath alcohol screens. Self-report TLFB and biological samples were obtained during twice-weekly study visits, allowing for biologically confirmed abstinence for each half-week of the 12-week treatment phase. Abstinence in the last 4 weeks of the treatment phase was thus based on the last 8 half-weeks. Abstinence was dichotomized (yes or no) where patients are categorized as abstinent only if they are abstinent for all 8 half-weeks, allowing for two missing half-weeks of data.

2.4. Data analysis

To answer the study question, we utilized a two-level hierarchical generalized linear model (HGLM) with a dichotomous abstinence outcome that analyzes hierarchically organized data (patients are nested within sites). Ignoring nesting in data can impact estimated variance, power needed to detect treatment effects, Type I error, statistical significance testing, and introduce bias; for these reasons, multilevel modeling is widely used in the social sciences, where data and observed variables are often hierarchically organized.

Whether any variable should be modeled in a statistical analysis as a random or fixed factor is a decision that should be based on the natural characteristics of the observed variable. Here we provide a simple illustration to demonstrate the differences between a variable modeled as a fixed or random factor. When making multiple measurements of a single subject's weight at one time point, the variation in measurements is assumed to be only due to measurement error; that is, under the null hypothesis it is assumed that the observed weight values are combinations of the true weight (the 'precise' value) plus or minus measurement error. However, when thinking about weight observations across multiple subjects in a study, the null hypothesis should not assume that the observed weights are just a combination of one true weight plus or minus measurement error. The correct null hypothesis in this case should assume that the observed differences in weight between subjects are partially due to measurement error and partially due to the natural variation among humans. Subjects are naturally different from each other, some have near average weights, while others have more extreme weights. It is common to assume human weight follows some (typically normal) distribution. Thus, the variation in observed weights across all subjects is a combination of natural variation in population weights plus measurement error. In this case, the additional variation in weights due to natural distribution of humans is typically modeled by including subject as a random factor.

In this paper, we assume that the variability across sites is not only due to measurement error but because sites (similar to example above) form a distribution in which some sites are extreme and most are average, and so we model site as a random factor. This allows us to better understand the nature of the site variability. Site modeled as a random factor assumes that the observed sites are a random sample from a distribution of all possible sites, giving a different meaning to site variability. Site modeled as a fixed factor assumes that any residual variability after controlling for site is due to measurement error. While site is commonly regarded as a fixed factor covariate in many clinical trial analyses (often because a small number of participating sites prevents researchers from properly modeling the random distribution of sites), we choose to model site as a random factor to appropriately

model and explain site variability (10 sites is adequate to include site as random factor). For an excellent overview of understanding variability and random vs. fixed effects, please refer to Gelman (2005).

In the first step of the model-building process, we modeled end-of-study abstinence using an HGLM with no fixed predictors and only site as a random factor. In addition to providing an overall estimate of the odds of abstinence for a typical patient, the model also provided an intra-class correlation (ICC) that indicates how much of the total variation in the abstinence is accounted for by site random factor. In the second step, we queried each of the putative patient- and site-level variables and their relationship to the primary abstinence outcome one by one in the HGLM. In the final step, all patient- and site-level variables that were found to be significant were included in a multiple variable HGLM. Because the primary analysis did not find a significant site by treatment interaction and treatment assignment was stratified by site, treatment assignment (TAU vs. TAU + TES) was not included in the HGLM models. For further details regarding HGLM, refer to Ene, Elizabeth, Blue, and Bell (2015).

All missing data was analyzed as missing. All hypothesis tests were two-sided and performed of 5% level of significance. The data were analyzed using SAS® 9.4.

3. Results

Ten community-based substance abuse treatment sites participated in the study, with 507 patients enrolled and randomized to receive TAU vs. TAU + TES. The 10 sites represented diverse patient populations; the percentage of Black patients ranged from 0 to 63%, male sex from 34% to 75%, and problem opioid use from 5% to 66%. Clinic characteristics also varied widely; the percent of counselors with a master's degree or greater ranged from 13% to 100%, counselors in recovery from 0% to 80%, and annual clinic admission from 300 to 5000. Patient-level and site-level baseline characteristics for each of the ten sites are shown in Tables 1 and 2, respectively.

3.1. Patient abstinence

The primary outcome of proportion of abstinence during the last 4 weeks of the study treatment phase ranged from 6.1% to 40% per site, as shown in Table 3. The HGLM model with only site as a random factor (see Table 4, Model 1) estimated that 6.7% of the variability in the abstinence outcome is due to differences between study sites, leaving 93.3% of the variability to be due to either individual patient characteristics or other unknown factors. Though not significant ($p = 0.0792$), the estimate of the amount of variability in the outcome explained by site ($ICC = 6.7\%$) is consistent with a Cohen's f of 0.26, which could be considered a small-moderate effect size, and the analysis may be underpowered to detect an effect of this size (Schoeneberger, 2016). This finding complements the primary outcome analysis in the primary paper (Campbell et al., 2014) where site was significantly associated with abstinence outcome when modeled as fixed factor ($OR = 0.37\text{--}1.02$; 95% $CI = 1.12\text{--}2.35$; $p = 0.01$).

3.2. Effect of patient- and site-level factors on abstinence outcome

Table 4 Model 2 shows the results of unadjusted effects of each patient- or site-level variable on the abstinence outcome. Patient age (OR = 1.40; 95% CI = 1.15, 1.71; $p = 0.0009$) and abstinence at baseline (OR = 2.77; 95% CI = 1.73, 4.45; $p < 0.0001$) were significantly associated with the abstinence outcome; older patients and those who were abstinent at baseline were more likely to be abstinent at end of study. Among site-level variables, higher annual clinic admissions (OR = 1.28; 95% CI = 1.03, 1.59; $p = 0.0251$) was significantly associated with abstinence. Of note, higher average counselor caseload (OR = 0.88; 95% CI = 0.74, 1.04; $p = 0.1408$) was associated with less abstinence but did not reach significance; however, when entered into the model, it reduced the amount of variance explained by site the most (from 6.7% to 1.9%, a 72% reduction in magnitude), suggesting that counselor caseload may be an important characteristic in differentiating between sites.

In the final, two-level HGLM, patient- and site-level variables that were found to be significant in unadjusted models were entered together into one adjusted model: patient age, patient abstinence at baseline, and annual clinic admissions. Because one site was missing data for its annual clinic admissions, the final model included only 9, instead of 10 sites. In this model, the site-level ICC reduced from 6.7% (model 1) to 1.4% (a 79% reduction in magnitude), suggesting that site differences in these variables may also largely account for the site effects on abstinence. The patient level characteristics of age (OR = 1.39; 95% CI = 1.12, 1.72; $p = 0.0027$) and baseline abstinence (OR = 2.30; 95% CI = 1.39, 3.82; $p = 0.0013$) remained significant while annual clinic admissions did not (OR = 1.12; 95% CI = 0.94, 1.35; $p = 0.2144$).

4. Discussion

Site variation in multi-site trials is common, though generally not investigated fully. This study examined site-specific effects on abstinence outcome rates, and contextual factors that could account for why some sites appeared to perform better than others. Multiple patient- and site-level characteristics including abstinence outcome rates varied across sites, and we found that the site-specific effect on abstinence outcome is small-moderate, though not significant when site was modeled as a random factor (i.e., when accounting for its natural variation). This complements the findings in the primary outcome analysis paper (Campbell et al., 2014) where site was included as a fixed effect covariate, and was significantly associated with abstinence outcome.

When site is modeled as a fixed factor, under the null hypothesis it is assumed that all sites behave the same and any observed variability across sites is only due to measurement error. In this clinical trial, when site was modeled as a fixed factor in the primary outcome analysis, the natural variability between sites were too large to be explained by measurement error alone, and site as a fixed factor was found to be significantly associated with the outcome. When we modeled site as a random factor, the large but natural variability among sites was only modestly related to the variation in the abstinence outcome, and the variation that it accounted for was diminished even further when two patient-level and one site-level variable (patient age, patient abstinence at baseline, and annual clinic admissions)

were included in the model. This may indicate that site differences in these variables are responsible for much of the site differences in abstinence.

The decision to model site as a fixed or random factor depends on the philosophical assumptions that one is willing to adopt about the nature of site variability and the characteristics of the data, but the decision has major ramifications for the number of sites and sample size required, the measurement of between-site variance, and implications for generalizability. Two models, one where site is included as a fixed factor and one where site is included as a random factor, treat the same site variability in a different way, with resultant differences in the conclusions that can be drawn from significance testing. When site is modeled as a random factor, it is assumed that the sites included in the study were randomly selected from a larger population of sites with a natural distribution and inherent variation, with some types of sites being more common and others more extreme. Site as a random factor then accounts for and models this natural variability and provides a statistical basis for generalizing results to the whole distribution of sites, including other sites. In a model where site is included as a fixed factor, one assumes that all sites in principle are the same and that the observed variability (after controlling for site) is due to measurement error. In such a model, it is hard to generalize findings to the whole population. It is important to note that in the clinical trial presented here, that though 10 sites were chosen from 41 candidate sites to maximize patient, program, and geographic diversity, they were not chosen randomly. A detailed statistical overview of the differences between fixed and random effect models in the context of clinical trial design are discussed at length elsewhere (Feaster, Mikulich-Gilbertson, & Brincks, 2011).

The second aim of the study was to identify the context of care driving site variability. We hypothesized that the clinical context, which includes a number of patient- and site-level characteristics, would account for variability between sites in the abstinence outcome. The amount of site variability associated with the abstinence outcome was greatly diminished when two patient-level and one site-level variable were included in the model: patient age, patient abstinence at baseline, and annual clinic admissions. Patients abstinent at study entry had much higher odds of achieving abstinence at the end of the treatment phase, a finding often replicated (Bisaga et al., 2005; Kampman et al., 2002). Given that all patients had begun to receive treatment (TAU) prior to study enrollment (patients had to be within the first 30 days of the current treatment episode), clinics providing better care as part of TAU may have had a higher percentage of patients abstinent in the first week. Therefore, this measure may be a proxy for overall better clinical care. Older patients had higher odds of abstinence, consistent with a number of studies showing that younger patients have lower retention rates in substance abuse treatment (Evans, Padwa, Li, Lin, & Hser, 2015; Weinstein et al., 2017). Higher annual clinic admission rates were associated with abstinence; clinics that have a higher volume of patients may have more experience and success working with this population, and may have greater resources. The relationship between higher patient volume and improved outcome is well-established in the medical, surgical, and mental health fields (Druss, Miller, Pincus, & Shih, 2004; Halm, Lee, & Chassin, 2002).

A number of other hypothesized patient- and site-level variables queried in the current study were not associated with abstinence. Patient opioid use, race and ethnicity, and

counselor education, recovery status, and attitudes towards evidence-based practices were not significantly associated with outcome. Expert opinion from co-investigators in the clinical trial included the observation that sites with younger, white, opioid using patients appeared to perform more poorly in the study, but this clinical observation was not supported by our results. Average counselor caseload diminished the most variance explained by site, though it was not significantly associated with outcome. Average counselor caseload is generally a function of budget and staffing limitations, and clinics with fewer resources may require counselors to see more patients. Substance abuse treatment counselors with higher caseloads may spend less time with patients and be more likely to experience burnout, which in turn may detrimentally affect quality of care (Shoptaw, Stein, & Rawson, 2000). However, there is a dearth of research on the impact of average counselor caseload on performance differences between sites or treatment outcome, an area which warrants further attention.

One of the major strengths of this study is that it incorporated rich patient, organizational and clinician data, which were measured at the outset of the trial and hence were available to analyze with respect to their association with site-level patient outcomes. This study also presents a statistical approach and rationale that can be adapted for future studies aiming to understand the role of site when aggregated in clinical trials. One limitation of this study is the small number of sites which may be too few to understand the overall distribution of sites, and that the final model included only nine sites due to missing data. Our study was likely underpowered to detect a small-moderate effect size of site on abstinence outcome, with site modeled as a random factor. Another limitation is that the outcome measure used in the site analysis is a modified version of what was used in the original analysis. The original abstinence outcome measure was not dichotomous but longitudinal dichotomous that included data for each half-week of the last 4 weeks of treatment (up to 8 time points). Thus, while the site-level outcome measures are not directly comparable, the modified outcome measure presented in this study allows for a more precise answer to questions regarding the nature of site variability. Lastly, the majority of the variance in abstinence outcome was not accounted for by site, leaving the possibility that treatment assignment combined with dropout, individual patient characteristics or other unassessed factors could be important in fully understanding its variance, e.g. patient substance use severity, or counselor fidelity to the treatment provided.

Because a significant fixed factor of site was identified in the primary analysis paper (Campbell et al., 2014) and there was large observed variability in outcome between sites, we hypothesized that site would remain significant when modeled as a random effect, and, further, that a combination of patient-level and site-level factors would best account for outcome variability across sites. We discovered was that site, though no longer significant, yielded a small-moderate effect size when modeled as a random effect, and that a model combining patient- and site-level factors best explained the site-specific variance. A site-level variable, average counselor caseload, emerged as the most promising proxy variable accounting for site variability, a finding which warrants further research. It has been recommended that both site and site by treatment interactions be routinely explored in multisite analysis (Kraemer & Robinson, 2005). This study further supports this recommendation, as well as including site as a random effect where possible.

Comprehensively exploring site effects will help to define the contexts of care optimizing treatment effect and outcome.

Acknowledgments

We would like to acknowledge Tse-Hwei Choo for providing statistical consultation and reviewing the manuscript.

Role of funding source

This work was supported by the CTN Greater NY Node (UG1 DA013035; PIs: Rotrosen, Nunes), Research in Addiction Medicine Scholars Program (4R25DA033211-04; Vaezazizi), Substance Abuse and Mental Health Services Administration Substance Abuse Minority Fellows Program (5T06sM060562-05; Vaezazizi), and Columbia University/New York State Psychiatric Institute T32DA00729.

References

- Bisaga A, Aharonovich E, Garawi F, Levin FR, Rubin E, Raby WN, & Nunes EV (2005). Utility of lead-in period in cocaine dependence pharmacotherapy trials. *Drug and Alcohol Dependence*, 77(1), 7–11. 10.1016/j.drugalcdep.2004.06.007. [PubMed: 15607836]
- Buti AL, Eakins D, Fussell H, Kunkel LE, Kudura A, & McCarty D (2013). Clinician attitudes, social norms and intentions to use a computer-assisted intervention. *Journal of Substance Abuse Treatment*, 44(4), 433–437. 10.1016/j.jsat.2012.08.220. [PubMed: 23021495]
- Campbell AN, Nunes EV, Matthews AG, Stitzer M, Miele GM, Polsky D, & Ghitza UE (2014). Internet-delivered treatment for substance abuse: A multisite randomized controlled trial. *The American Journal of Psychiatry*, 171(6), 683–690. doi 10.1176/appi.ajp.2014.13081055. [PubMed: 24700332]
- Campbell AN, Nunes EV, Miele GM, Matthews A, Polsky D, Ghitza UE, & Crowell AR (2012). Design and methodological considerations of an effectiveness trial of a computer-assisted intervention: An example from the NIDA Clinical Trials Network. *Contemporary Clinical Trials*, 33(2), 386–395. doi 10.1016/j.cct.2011.11.001. [PubMed: 22085803]
- Carroll KM, Kiluk BD, Nich C, Gordon MA, Portnoy GA, Marino DR, & Ball SA (2014). Computer-assisted delivery of cognitive-behavioral therapy: Efficacy and durability of CBT4CBT among cocaine-dependent individuals maintained on methadone. *The American Journal of Psychiatry*, 171(4), 436–444. doi 10.1176/appi.ajp.2013.13070987. [PubMed: 24577287]
- Druss BG, Miller CL, Pincus HA, & Shih S (2004). The volume-quality relationship of mental health care: Does practice make perfect? *The American Journal of Psychiatry*, 161(12), 2282–2286. 10.1176/appi.ajp.161.12.2282. [PubMed: 15569901]
- Ene ML, Elizabeth A, Blue GL, & Bell BA (2015). Multilevel models for categorical data using SAS PROC GLIMMIX: The basics. *SAS Global Forum 2015 Proceedings*.
- Evans E, Padwa H, Li L, Lin V, & Hser YI (2015). Heterogeneity of mental health service utilization and high mental health service use among women eight years after initiating substance use disorder treatment. *Journal of Substance Abuse Treatment*, 59, 10–19. 10.1016/j.jsat.2015.06.021. [PubMed: 26321439]
- Feaster DJ, Mikulich-Gilbertson S, & Brincks AM (2011). Modeling site effects in the design and analysis of multi-site trials. *The American Journal of Drug and Alcohol Abuse*, 37(5), 383–391. doi 10.3109/00952990.2011.600386. [PubMed: 21854281]
- Fishbein M, & Ajzen I (1975). *Belief, attitude, intention and behaviour: An introduction to theory and research*. Reading, MA: Addison-Wesley.
- Gelman A (2005). Analysis of variance-why it is more important than ever. *The Annals of Statistics*, 33(1), 1–53. 10.1214/009053604000001048.
- Halm EA, Lee C, & Chassin MR (2002). Is volume related to outcome in health care? A systematic review and methodologic critique of the literature. *Annals of Internal Medicine*, 137(6), 511–520. 10.7326/0003-4819-137-6-200209170-00012. [PubMed: 12230353]

- Higgins ST, Budney AJ, Bickel WK, Foerg FE, Donham R, & Badger GJ (1994). Incentives improve outcome in outpatient behavioral treatment of cocaine dependence. *Archives of General Psychiatry*, 51(7), 568–576. 10.1001/archpsyc.1994.03950070060011. [PubMed: 8031230]
- Higgins ST, Sigmon SC, Wong CJ, Heil SH, Badger GJ, Donham R, & Anthony S (2003). Community reinforcement therapy for cocaine-dependent out-patients. *Archives of General Psychiatry*, 60(10), 1043–1052. 10.1001/archpsyc.60.9.1043. [PubMed: 14557150]
- Kampman KM, Volpicelli JR, Mulvaney F, Rukstalis M, Alterman AI, Pettinati H, & O'Brien CP (2002). Cocaine withdrawal severity and urine toxicology results from treatment entry predict outcome in medication trials for cocaine dependence. *Addictive Behaviors*, 27(2), 251–260. 10.1016/S0306-4603(01)00171-X. [PubMed: 11817766]
- Knudsen HK, Abraham AJ, & Oser CB (2011). Barriers to the implementation of medication-assisted treatment for substance use disorders: The importance of funding policies and medical infrastructure. *Evaluation and Program Planning*, 34(4), 375–381. 10.1016/j.evalprogplan.2011.02.004. [PubMed: 21371752]
- Kraemer HC, & Robinson TN (2005). Are certain multicenter randomized clinical trial structures misleading clinical and policy decisions? *Contemporary Clinical Trials*, 26(5), 518–529. 10.1016/j.cct.2005.05.002. [PubMed: 15993653]
- Marsch LA (2011). Technology-based interventions targeting substance use disorders and related issues: An editorial. *Substance Use & Misuse*, 46(1), 1–3. 10.3109/10826084.2011.521037.
- Miller WR, Sorensen JL, Selzer JA, & Brigham GS (2006). Disseminating evidence-based practices in substance abuse treatment: A review with suggestions. *Journal of Substance Abuse Treatment*, 31(1), 25–39. 10.1016/j.jsat.2006.03.005. [PubMed: 16814008]
- NDIC (2011). National Drug Intelligence Center: Economic impact of illicit drug use on American society. (Washington, DC).
- Nunes EV, Ball S, Booth R, Brigham G, Calsyn DA, Carroll K, ... Woody G (2010). Multisite effectiveness trials of treatments for substance abuse and co-occurring problems: Have we chosen the best designs? *Journal of Substance Abuse Treatment*, 38(Suppl. 1), S97–112. 10.1016/j.jsat.2010.01.012. [PubMed: 20307801]
- Prendergast M, Podus D, Finney J, Greenwell L, & Roll J (2006). Contingency management for treatment of substance use disorders: A meta-analysis. *Addiction*, 101(11), 1546–1560. 10.1111/j.1360-0443.2006.01581.x. [PubMed: 17034434]
- Rieckmann T, Daley M, Fuller BE, Thomas CP, & McCarty D (2007). Client and counselor attitudes toward the use of medications for treatment of opioid dependence. *Journal of Substance Abuse Treatment*, 32(2), 207–215. 10.1016/j.jsat.2006.09.002. [PubMed: 17306729]
- Schoeneberger JA (2016). The impact of sample size and other factors when estimating multilevel logistic models. *Journal of Experimental Education*, 84, 373–397. 10.1080/00220973.2015.1027805.
- Shoptaw S, Stein JA, & Rawson RA (2000). Burnout in substance abuse counselors. Impact of environment, attitudes, and clients with HIV. *Journal of Substance Abuse Treatment*, 19(2), 117–126. 10.1016/S0740-5472(99)00106-3. [PubMed: 10963923]
- Sobell LC, & Sobell MB (1992). In Litten RZ, & Allen JP (Eds.). *Measuring alcohol consumption: psychological and biochemical methods*. (pp. 41–72). New Jersey: Humana Press.
- Weinstein ZM, Kim HW, Cheng DM, Quinn E, Hui D, Labelle CT, & Samet JH (2017). Long-term retention in Office Based Opioid Treatment with buprenorphine. *Journal of Substance Abuse Treatment*, 74, 65–70. 10.1016/j.jsat.2016.12.010. [PubMed: 28132702]

Table 1

Select patient-level variables by site (n = 507 patients over ten sites).

Site	1	2	3	4	5	6	7	8	9	10
Number of patients	49	51	50	38	53	50	53	48	60	55
Abstinence at baseline (%)	46.94	25.49	48.00	47.37	62.26	32.00	49.06	64.58	76.67	81.82
Male (%)	57.14	35.29	72.00	34.21	73.58	66.00	75.47	58.33	75.00	62.96
Age (years, mean)	27.96	38.06	32.74	32.50	33.47	34.44	35.77	33.67	36.13	42.40
Race (%) ^b										
Asian	0.00	0.00	2.00	0.00	0.00	0.00	0.00	2.08	18.33	0.00
Black	8.16	62.75	10.00	0.00	5.66	16.00	62.26	0.00	1.67	54.55
Chose not to answer	2.04	0.00	0.00	0.00	0.00	2.00	0.00	0.00	0.00	0.00
Multiracial	8.16	5.88	6.00	7.89	0.00	2.00	11.32	6.25	48.33	3.64
Native American	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	20.00	0.00
Other	4.08	0.00	12.00	2.63	3.77	10.00	1.89	0.00	1.67	9.09
Pacific Islander	0.00	0.00	0.00	7.89	0.00	0.00	0.00	0.00	0.00	0.00
White	77.55	31.37	70.00	81.58	90.57	70.00	24.53	91.67	10.00	32.73
Ethnicity: Hispanic/Latino (%)	12.50	7.84	14.00	5.26	5.66	18.00	1.96	6.38	18.33	16.36
Education (%)										
Less than high school	12.24	35.29	14.00	23.68	24.53	24.00	22.64	33.33	16.67	27.27
High school	63.27	52.94	70.00	71.05	58.49	48.00	71.70	58.33	68.33	50.91
AA	16.33	5.88	4.00	2.63	13.21	14.00	5.66	6.25	11.67	16.36
BA	8.16	5.88	8.00	0.00	1.89	12.00	0.00	2.08	3.33	5.45
MA	0.00	0.00	4.00	2.63	1.89	2.00	0.00	0.00	0.00	0.00
Primary substance (%)										
Alcohol	10.20	39.22	12.00	21.05	28.30	28.00	18.87	16.67	13.33	18.18
Opioids	38.78	1.96	26.00	15.79	50.94	18.00	18.87	31.25	5.00	9.09
Cocaine	22.45	37.25	30.00	0.00	11.32	18.00	13.21	8.33	5.00	50.91
Stimulants	28.57	21.57	30.00	26.32	3.77	26.00	49.06	20.83	16.67	5.45
Marijuana	0.00	0.00	0.00	36.84	0.00	4.00	0.00	22.92	58.33	12.73
Other	0.00	0.00	2.00	0.00	5.66	6.00	0.00	0.00	1.67	3.64
Opioid use (%) ^a	57.14	13.73	36.00	34.21	66.04	30.30	24.53	56.25	5.00	21.82

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Percent of patients with primary opioid use, past-year opioid abuse, or past-year opioid dependence.
Four patients were missing data on race, which was analyzed as missing.

Table 2

Select site-level variables by site (n = 135 counselors over ten sites).

Site	1	2	3	4	5	6	7	8	9	10
Number of counselors	14	14	12	15	32	11	10	12	7	8
Counselor age (mean)	46.57	47.57	53.00	46.47	53.50	38.09	39.90	48.92	47.14	33.75
Counselor sex (female, %)	78.57	71.43	58.33	66.67	96.88	81.82	80.00	83.33	71.43	75.00
Counselor race (%)										
Asian	0.00	7.14	0.00	0.00	0.00	0.00	10.00	0.00	14.29	12.50
Black	7.14	21.43	0.00	0.00	0.00	18.18	0.00	0.00	0.00	25.00
Chose not to answer	7.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Multiracial	0.00	0.00	0.00	26.67	3.13	0.00	0.00	8.33	28.57	12.50
Native American	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Other	7.14	0.00	0.00	0.00	0.00	9.09	0.00	0.00	14.29	0.00
Pacific Islander	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	28.57	0.00
White	78.57	71.43	100.00	73.33	96.88	72.73	90.00	91.67	14.29	50.00
Counselor education (%)										
No college degree	0.00	7.14	0.00	13.33	15.63	0.00	0.00	0.00	0.00	25.00
AA	0.00	14.29	0.00	53.33	6.25	9.09	20.00	50.00	28.57	12.50
BA/BS	35.71	21.43	0.00	20.00	6.25	9.09	10.00	41.67	14.29	25.00
Masters	57.14	50.00	83.33	13.33	68.75	72.73	60.00	8.33	57.14	37.50
Doctorate	7.14	7.14	16.67	0.00	3.13	9.09	10.00	0.00	0.00	0.00
Counselor work experience > 5 years (%)	64.29	78.57	83.33	60.00	71.88	36.36	50.00	58.33	71.43	12.50
Counselors in recovery (%)	23.08	14.29	0.00	80.00	13.79	9.09	10.00	72.73	57.14	42.86
Contingency management										
Social norms (mean)	5.00	4.64	4.00	5.07	4.16	4.55	5.40	3.33	5.43	5.25
Intention to implement (mean)	5.43	5.07	4.36	4.67	4.44	4.55	6.30	3.25	4.43	4.88
Efficacy attitude (mean)	5.64	5.86	5.45	5.80	4.63	5.45	6.20	4.50	4.86	5.38
Twelve step influence (%)										
Not at all	7.14	0.00	16.67	0.00	3.13	18.18	10.00	0.00	0.00	25.00
Somewhat	35.71	50.00	58.33	66.67	62.50	63.64	70.00	50.00	42.86	37.50
Very much	57.14	50.00	25.00	33.33	34.38	18.18	20.00	50.00	57.14	37.50

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Site	1	2	3	4	5	6	7	8	9	10
Typical caseload per counselor	a	55	25	25	70	25	35	26	15	32
Annual client admissions to clinic	976	720	300	766	752	376	a	682	1861	5000

^bMissing data.

Table 3

Treatment outcome by site (n = 507 patients over ten sites).

Site	1	2	3	4	5	6	7	8	9	10
Number in TAU/TES	24/25	26/25	24/26	19/19	27/26	23/27	26/27	25/23	31/29	27/28
Abstinence last four weeks (%)	Overall	6.12	13.73	16.00	18.42	20.75	24.00	26.42	33.33	40.00
	TAU	0.00	11.54	16.67	5.26	11.11	26.09	26.92	32.00	33.33
	TES	12.00	16.00	15.38	31.58	30.77	22.22	25.93	34.78	44.83
									44.83	46.43

TAU = treatment as usual; TES = Therapeutic Education System.

Table 4

Two-level hierarchical generalized linear models (HGLMs) of binary abstinence outcome in last four weeks of study (weeks 9–12) with patients (n = 507) nested within ten community-based treatment sites. Columns show the effect of site alone without any covariates (Model 1), and adjusting for patient or site-level covariates one at a time (Model 2).^a

	Odds ratio (CI 95%) for effect of covariate	Effect of site as random factor ^b (standard error)	p value for the effect of site as random factor	Percent of variance explained by site ^c
<i>Model 1</i>				
No covariates	N/A	0.2348 (0.1665)	0.0792	6.67%
<i>Model 2</i>				
Level 1: Patient-level covariates				
Age (/10)	1.40 (1.15–1.71) ^d	0.1529 (0.1297)	0.1192	4.44%
Abstinence at baseline	2.77 (1.73–4.45)	0.1211 (0.1172)	0.1506	3.55%
Race	Black 0.75 (0.38–1.47) Other 1.04 (0.58–1.90) White (reference group) ^e	0.2743 (0.1929)	0.0775	7.70%
Hispanic ethnicity	0.54 (0.25–1.16)	0.2481 (0.1727)	0.0754	7.01%
Opioid use	0.93 (0.57–1.50)	0.2287 (0.1648)	0.0826	6.50%
Level 2: Site-level covariates				
Counselors in recovery (%)	2.56 (0.77–8.48)	0.1569 (0.1321)	0.1174	4.55%
12 step influence (%) ^f	1.41 (0.10–20.51)	0.2366 (0.1670)	0.0784	6.71%
Typical caseload per counselor (/10)	0.88 (0.74–1.04)	0.0647 (0.0812)	0.2128	1.93%
Annual client admissions (/1000)	1.28 (1.03–1.59)	0.1223 (0.1231)	0.1603	3.58%
CM social norm, mean (/10)	2.23 (0.01–56.57)	0.2291 (0.1649)	0.0824	6.51%
CM intent to implement, mean (/10)	0.10 (< 0.00–10.2)	0.2066 (0.1520)	0.0870	5.91%
CM efficacy attitude, mean (/10)	0.64 (0.34–1.20)	0.1766 (0.1375)	0.0995	5.09%
Advanced degree (%)	0.56 (0.28–1.12)	0.1533 (0.1286)	0.1165	4.45%

Abbreviations: CM = contingency management.

^aSeparate HGLMs are run for each variable.

^bThe level-2 random intercept term reflects the effect of site in the 2-level HGLM.

^cIntraclass correlation coefficient reflects the percentage of variance explained by site alone (Model 1) and after adjustment for each variable (Models 2).

^dOdds ratios with significant values ($p < 0.05$) are indicated in bold.

^eRace overall was not significant ($F_{2, 496} = 0.43$; $p = 0.6487$).

^fPercent of counselors reporting 12-step influence as “not at all” or “somewhat” in counselor survey.