



Published in final edited form as:

*Drug Alcohol Depend.* 2014 April 1; 137: 3–19. doi:10.1016/j.drugalcdep.2014.01.012.

## Towards empirical identification of a clinically meaningful indicator of treatment outcome: Features of candidate indicators and evaluation of sensitivity to treatment effects and relationship to one year follow up cocaine use outcomes

Kathleen M. Carroll<sup>1</sup>, Brian D. Kiluk<sup>1</sup>, Charla Nich<sup>1</sup>, Elise E. DeVito<sup>1</sup>, Suzanne Decker<sup>1,2</sup>, Donna LaPaglia<sup>1</sup>, Dianne Duffey<sup>1</sup>, Theresa A. Babuscio<sup>1</sup>, and Samuel A. Ball<sup>1,3</sup>

<sup>1</sup>Department of Psychiatry, Yale University School of Medicine, 950 Campbell Avenue, 151D, West Haven, CT 06516

<sup>2</sup>New England Mental Illness, Research Education and Clinical Center, VA Connecticut Healthcare System

<sup>3</sup>The APT Foundation, New Haven Connecticut, 1 Long Wharf, New Haven CT 06511

### Abstract

**Background**—Selection of an appropriate indicator of treatment response in clinical trials is complex, particularly for the various illicit drugs of abuse. Most widely-used indicators have been selected based on expert group recommendation or convention rather than systematic empirical evaluation. Absence of an evidence-based, clinically meaningful index of treatment outcome hinders cross-study evaluations necessary for progress in addiction treatment science.

**Method**—Fifteen candidate indicators used in multiple clinical trials as well as some proposed recently are identified and discussed in terms of relative strengths and weaknesses (practicality, cost, verifiability, sensitivity to missing data). Using pooled data from five randomized controlled

---

© 2014 Elsevier Ireland Ltd. All rights reserved.

Corresponding author: Kathleen M Carroll, Department of Psychiatry, Yale University School of Medicine, VA CT Healthcare Center (151D), West Haven, Connecticut 06516.

#### Contributors

Authors are required to declare their individual contribution to the manuscript under a subheading Contributors. All authors must have materially participated in the research and/or manuscript preparation, so roles for all authors should be described. The statement that all authors have approved the final manuscript should be true and included in the disclosure.

Author Carroll designed the 5 studies which contributed data to this article. Authors Kilik, Nich, and Carroll designed the study and planned the analyses. Authors Nich, Babuscio, and Kilik conducted the statistical analyses; authors Ball, DeVito, Decker and Duffey contributed to the analyses and their interpretation..

Author Carroll wrote the initial draft of the manuscript and all authors contributed to and have approved of the final manuscript.

#### Conflict of Interest

Author Carroll is a member of CBT4CBT LLC, which makes CBT4CBT available to qualified clinical providers and organizations on a commercial basis. Dr. Carroll works with Yale University to manage any potential conflicts of interest. All other authors declare that they have no conflicts of interest.

A version of this report was presented at the Annual Meeting of CPDD in June, 2012.

**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.

trials of cocaine dependence (N = 434), the indicators were compared in terms of sensitivity to the effects of treatment and relationship to cocaine use and general functioning during follow-up.

**Results**—Commonly used outcome measures (percent negative urine screens; percent days of abstinence) performed relatively well in that they were sensitive to the effects of the therapies evaluated. Others, including complete abstinence and reduction in frequency of use, were less sensitive to effects of specific therapies and were very weakly related to cocaine use or functioning during follow-up. Indicators more strongly related to cocaine use during follow-up were those that reflected achievement of sustained periods of abstinence, particularly at the end of treatment.

**Conclusions**—These analyses did not demonstrate overwhelming superiority of any single indicator, but did identify several that performed particularly poorly. Candidates for elimination included retention, complete abstinence, and indicators of reduced frequency of cocaine use.

## Keywords

Treatment outcome indicators; randomized trials; cocaine use disorders; follow-up studies

---

## 1. INTRODUCTION

### 1.1 Potential advantages of a commonly accepted outcome indicator

The field of treatment of illicit drug use has not yet achieved consensus on a practical, psychometrically sound, clinically significant, and broadly accepted indicator of treatment success or response (Donovan et al., 2012). In areas outside the addictions, identification and adoption of standard measures of treatment response and clinical significance has facilitated several important advances. These include the ability to more easily accrue and compare outcomes for treatments and trials in meta analyses, benchmark clinical research and clinical outcomes, set and monitor performance standards, compare outcomes among different subgroups and samples, clearly convey the magnitude of treatment effects to clinicians and policy makers, inform planning for equivalence trials, and facilitate comparisons against a common meaningful outcome standard. Moreover, adoption of common outcome measures to benchmark intervention trials does not constrain investigators for selecting additional outcomes theoretically linked to the treatments putative mechanisms of action.

The fields of alcohol and nicotine intervention research have made significant inroads in wider adoption of common outcome measures in clinical trials as well as moving towards defining benchmark indicators of treatment response. In nicotine research, prolonged abstinence and point-prevalence abstinence, with biochemical validation, are now standard measures for outcome trials (Hughes, 2007; Hughes et al., 2010; West et al., 2005). These measures have been subjected to extensive psychometric evaluation establishing reliability and validity, linked to multiple outcome domains, and shown to be sensitive to the effects of a range of behavioral and pharmacological therapies (Hughes et al., 2003, 2010).

For alcohol use disorders, percent days of abstinence and indices of heavy drinking (e.g., percent heavy drinking days) have become standard measures of outcome for treatment

trials (Allen, 2003; Anton and Randall, 2005; Finney et al., 2003; Sobell et al., 2003). These indices have been demonstrated to be associated with a range of drinking-related risks and consequences (Falk et al., 2010). Efforts have also focused on identifying clinically meaningful and reliable dichotomous measures of treatment success, or good clinical outcome (Anton et al., 2006; Cisler et al., 2005; Cisler and Zweben, 1999; Sobell et al., 2003). These have included rates of individuals who are completely abstinent, as well as rates of individuals with no heavy drinking days (Falk et al., 2010). Recognizing that many treatment effects take some time to emerge and the course of improvement is not always linear, some indicators include a ‘grace period,’ for example, by not including heavy drinking days that occur early in treatment or allowing a limited number of heavy drinking days during treatment in the dichotomous definition of good outcome (Falk et al., 2010; Winchell et al., 2012).

For illicit drug use disorders, a common definition of treatment response has been elusive, in part because of the complexity of determining what constitutes ‘clinically significant’ change (Winchell et al., 2012). Common conventions for evaluating clinical significance (Jacobson and Truax, 1991; Kendall et al., 1999; Kraemer et al., 2003b; Lambert and Bailey, 2012), such as normative comparisons or reliable change indices are not easily generalized to drug use disorders, e.g., because illicit drug use is essentially non-normative, definitions based on the concept of ‘return to normative levels’ generally imply full sustained abstinence. Although full abstinence is a relatively unambiguous positive outcome, limiting ‘success’ to complete abstinence may be an overly exclusive criterion, as the process of change is inherently dynamic and thus complete abstinence may not be an appropriate indicator of meaningful cessation behavior (Ciraulo et al., 2003; DeBusk et al., 1994; Hser et al., 2008; Hughes et al., 2010; Velicer and Prochaska, 2004). Moreover, given that brief slips are common, a dichotomous measure of prolonged abstinence may be a relatively insensitive measure in that (1) few individuals may meet this standard, and (2) individuals who use drugs once or twice would be considered treatment ‘failures’, even though they may have made considerable improvement with respect to individuals whose drug use is unimproved.

## 1.2 Defining meaningful change

Alternative strategies for quantifying ‘clinically meaningful change’ include defining a threshold of meaningful change, typically at least two standard deviations on a symptom/problem index (Jacobson et al., 1999; Jacobson and Truax, 1991). This strategy has proven useful in defining positive outcomes for alcohol use disorders (e.g., ‘no heavy drinking days’). Reduction of alcohol use to normative levels is now a viable goal for treatment outcome studies, as it has been related to measurably reduced health risks and drinking-related consequences (Cisler et al., 2005; Cisler and Zweben, 1999). Significant reduction in alcohol use also appears to be an acceptable goal to clinical providers, as many clinicians would see a 50% reduction in drinking as clinically meaningful (Miller and Manuel, 2008).

Conversely, ‘significant reduction in tobacco use’ has not been adopted as an indicator of significant change, as even relatively low levels of regular smoking potentially carry significant health risks. Similarly, in the field of drug use, there are few data from well-

controlled trials linking reductions in frequency/level of drug use to meaningful improvement in important areas of functioning (e.g., medical, legal, employment, social). That is, there are few data suggesting that reductions in drug use are associated with meaningful benefits in health or functioning, nor data regarding what might indicate ‘safe’ levels of illicit opioid, cocaine, or marijuana use. Furthermore, patterns of drug use and consequences tend to vary widely both within and across types of illicit drug use (‘chippers’ versus daily users of opioids, binge use of cocaine versus regular low-level use). This variability generally results in very large standard deviations on most indices of drug use such that reduction of two or more standard deviations from baseline frequency/intensity measures typically translates to complete abstinence.

Developing a criterion for ‘treatment success’ based on change from baseline is complex, as it would result in differing criteria for success across study populations and may not yield an outcome that is clinically meaningful despite its statistical significance. For example, in a large population of treated cocaine users in the UK, Marsden and colleagues (Marsden et al., 2009) calculated the reliable change index (Jacobson and Truax, 1991), which they determined to be roughly 12 days of cocaine use or less per month—a definition of ‘successful outcome’ which would seem unacceptable to most clinicians.

### 1.3 Dichotomous versus continuous indicators

Dichotomous (e.g., success/fail, improved/not improved) indicators of outcome are fairly easy to interpret, and thus may be attractive to clinicians or policy makers. However, because of the relative statistical and power advantages of continuous over categorical measures of outcome, many clinical trials have adopted continuous measures of frequency. At the same time, continuous measures such as indices of change in quantity are also much more complex for illicit drug use relative to alcohol and nicotine use. Standard units for alcohol and nicotine are well defined, facilitating evaluations in terms of reductions in quantity of use (Devos-Comby and Lange, 2008; Miller et al., 1991). For illicit drugs, there are no standard units, potency varies widely, adulterants are uncontrolled, and frequently used terms such as ‘dime bags,’ ‘hits,’ and ‘joints’ are anything but standard. This variation makes it extremely difficult to quantify use across individuals and samples. The strategy of transposing quantity of drug use to a standard such as cost in monetary currency is undercut by both the illicit (drug commerce, trading sex for drugs, etc., makes it difficult to estimate monetary value) as well as the social nature of drug use (difficulty of estimating value of shared joints or cocaine in social situations).

### 1.4 Rationale for the current study

Multiple expert groups have sought and failed to reach consensus on a more standardized approach to selection of primary outcomes in the field of drug dependence (Donovan et al., 2012). These panels’ recommendations have been informed by extensive experience in conducting trials, but seldom guided by empirical evaluation and systematic comparison of the candidate outcome variables. McKay and colleagues’ report (2001) on continuous, categorical, and time-to-event indices of cocaine use, drawn from a clinical trial of behavioral interventions involving 132 male cocaine-dependent veterans, remains the most comprehensive empirical evaluation of cocaine treatment outcomes to date. Although none

of the measures was sensitive to treatment effects, McKay reported strong correlations among the indicators and advocated use of two general types of outcomes: a measure of frequency of cocaine use and a measure of severity. Since that report, multiple trials have used these and similar measures to describe outcomes and to evaluate predictors of longer term outcome but there have been no published reports of systematic empirical comparison of outcome indicators that include data on relationships between within-treatment outcomes and drug use during follow-up. Moreover, many trials, particularly those evaluating pharmacotherapies, focus solely on drug use within the treatment protocol (Donovan et al., 2012; Wells et al., 2010). However, because addiction is widely regarded as a chronically relapsing disorder (McLellan, 2002; McLellan et al., 2000; Volkow, 2005), the degree to which within-treatment indicators are associated with follow-up outcomes is a critical but understudied aspect of the outcomes literature.

To address these issues, and as an initial step towards identifying variables that might serve as common outcome indicators in treatment studies for illicit drug use, we identified a range of continuous and dichotomous indicators that have been used as primary outcome measures in clinical trials of treatments for illicit drug use (Donovan and Marlatt, 2005; Dutra et al., 2008; McKay et al., 2001; Peters et al., 2011). We then review them briefly in terms of the extent to which they are (1) easily collected and computed, (2) verifiable via biological measures, (3) sensitive to the effects of missing data, and (4) cost effective. Next, we use pooled data from a series of randomized controlled trials of cocaine dependence to compare these measures in terms of criteria that can be evaluated empirically (sensitivity to treatment effects, relationship to follow-up outcomes). We limited candidate indicators to those that focus on drug use, as alternate outcomes measures (e.g., craving, quality of life) have recently been reviewed elsewhere (Tiffany et al., 2012a).

#### 1.4.1 Continuous indicators

- 1 *Days of retention in the treatment protocol.* Retention is frequently used as a process and outcome variable. It can be an indicator of the acceptability of treatments and can signal the possibility of differential data availability across conditions. Multiple reports have linked treatment retention to better outcomes (Ciraulo et al., 2003; NIDA, 2007; Simpson et al., 1999).
- 2 *Percent of cocaine-negative urine specimens.* Results based on urine toxicology screens are among the most widely used in the field of treatment of illicit drugs, as they reflect biological verification of recent use. Other sources of biological data (e.g., hair, nails, skin, saliva) were not included here because they are not yet commonly used in clinical trials (Donovan et al., 2012) and were not collected in the studies included in this report.
- 3 *Longest period of consecutive abstinence during treatment.* A measure of the extent to which the participant attains a stable period of abstinence within treatment has frequently been linked to longer-term outcome (Higgins et al., 2000). It is often based on self-report via a Timeline Follow Back (TLFB) method (Robinson et al., in press; Sobell and Sobell, 1992), and can be verified via biological measures, providing they are collected at appropriate intervals.

- 4 *Percent days of abstinence based on participant self-report (PDA)*. This is a widely-used indicator in both the illicit drug and alcohol literature. Based on self-reports of frequency of use (typically drawn from TLFB), it has the advantage of typically being available for all protocol days for all participants, as missing information can be back-filled when the participant is reached at subsequent assessment points.
- 5 *Maximum days of abstinence during last two weeks of treatment*. Similar to maximum consecutive days of abstinence, this is an indicator of the extent to which the participant attained stable abstinence during the two-week interval prior to completing or dropping out of treatment. This indicator might be useful when it is recognized that a treatment may take several weeks to exert its effects, and a continuous measure of end-of-treatment abstinence is desired.

**1.4.2. Dichotomous indicators**—Dichotomous indicators were drawn from the treatment outcome literature as well as in response to calls for novel indicators tapping reductions in frequency of use and included:

- 6 *Abstinence during the final two weeks of treatment*. This is a dichotomous version of ‘maximum days of abstinence during the last two weeks of treatment’ described above. The modest success of many agents evaluated as pharmacotherapy for stimulant dependence has led to proposals to consider alternative measures of outcome such as these, in that they may be sensitive to more subtle medication effects (McCann and Li, 2011).
- 7 *Abstinence for 3 or more weeks*. This measure has been used frequently since the earliest randomized clinical trials for cocaine dependence (Gawin et al., 1989) as an indicator of positive treatment response (Ehrman et al., 2001). It measures whether or not the participant attained a stable period of 21 or more consecutive days of abstinence at any time during the course of treatment, rather than just at the end of treatment.
- 8 and 9 *Abstinence of 2 or 1 weeks duration*. As the ‘three or more weeks of abstinence’ was adopted based on clinical intuition and convention rather than empirical evaluation, these two dichotomous indicators were included in order to evaluate whether briefer sustained periods of abstinence may be associated with meaningful changes in drug use over time
- 10 *Complete abstinence*. This is usually considered the most stringent and unambiguous indicator of treatment success. It has been criticized as insensitive to smaller levels of change that might be meaningful. When using this measure it is important to specify when the beginning of abstinence is to be measured, as many participants enter treatment while still using illicit drugs, and thus a ‘grace’ or ‘washout’ period might be used. In addition, use of this variable requires specification of whether this refers to complete verified abstinence during the full *intended* length of the trial or simply during the period of time the participant provided data.

- 11 *Treatment completion with end-of-treatment abstinence.* This is a composite measure indicating the participant completed treatment *and* was abstinent at the end of treatment. Focus on ‘success at the finish line’ has recently received support within the alcohol literature (Falk et al., 2010), and is frequently utilized in clinical practice.

**1.4.3. Reduction indicators**—Both continuous and categorical indicators of reduction were included, given recent interest in this approach for evaluating novel medications (see McCann and Li, 2011).

- 12 *Percent reduction in frequency of drug use.* Given the difficulty of measuring changes in quantity of illicit drug use, the outcome literature has typically focused on changes in frequency of use from baseline to endpoint (Crits-Christoph et al., 1999; Simpson et al., 1999). An alternate approach to evaluating reduction in use that incorporates biological validation would include changes in quantitative urine measures over time. This approach is promising (Huestis et al., 2000, 2006; Preston et al., 1997) but not included here as quantitative urine toxicology was not conducted in these studies.
13. and 14 *50% or 75% reduction in frequency of drug use.* These are dichotomized versions of the continuous reduction variable described above. They were selected for this report in order to compare several dichotomous indicators tapping reduction in frequency of use to those focused on abstinence. The 50% reduction variable was selected based on Miller’s recommendation that reduction of frequency of use by half would be seen as meaningful by clinicians (Miller and Manuel, 2008). The 75% reduction variable was selected as it would be a conservative estimate of a reliable change index and consistent with indices used in previous health system evaluation studies (Marsden et al., 2009).
- 15 *Proxy measure of ‘good functioning’ at the end of treatment.* Definitions of ‘success’ in treating drug users are not always confined to substance use alone; but may include other variables related to employment status, criminal justice behavior, psychological functioning, health and quality of life (Humphreys and McLellan, 2012; Miller and Miller, 2009; Tiffany et al., 2012a, b; Winchell et al., 2012; Witkiewitz, 2013). Thus, in order to allow some comparison of the candidate indices included here, which focus only on drug use, to a single measure indicating relatively unambiguous positive outcome on a number of dimensions, we developed a proxy dichotomous measure for good functioning, based on Addiction Severity Index (ASI; McLellan et al., 1985). It reflects complete abstinence in the past 28 days as well as 0 days of problems during that period in the legal, psychological, employment, and family domains.

## 1.5 General characteristics of the candidate indicators

**1.5.1 Ease of measurement**—Broadly speaking, the various dichotomous measures (complete abstinence and/or abstinence for defined intervals) are fairly straightforward and easily computed. That is, the participant either completes treatment or not, provides drug-negative urine specimens and self-report data supporting the defined length of abstinence, or

not. Furthermore, the dichotomous variables can usually be computed for all participants randomized, as individuals who do not complete treatment or do not demonstrate abstinence of the defined length are assumed to have not meet the criterion.

The reduction measures are more complex, as they require the investigator to operationalize the time period of interest for measurement of baseline use. Moreover, operationalization of reduction is complicated (e.g., is reduction defined in terms of frequency or quantity of use of a given drug? How are *increases* in drug use handled in the analyses?). Reduction measures are also complex in terms of selecting the period over which reduction is to be measured (i.e., the entire course of treatment, last month?). Finally, estimates of the frequency, or quantity of use during the period preceding treatment entry are typically based on self-report and cannot be verified independently.

The continuous indicators based on self-report (longest period of abstinence, percent days of abstinence) or urine toxicology screens (percent drug-positive or negative urines) are fairly straightforward in terms of calculation when trial data are complete. However, as is the case of virtually all trials involving users of illicit drugs, some level of missing data is inevitable. When data are incomplete (due to attrition from the trial, missed assessment visits, missing urines), careful consideration regarding specification of the outcome indicator is needed as it can influence study findings: For example, does one consider percent drug-positive urines in terms of all urine specimens *expected* or those actually *collected*? If based only on urine specimens actually collected, results can be highly variable: Consider the case of an individual who submitted one urine specimen and then dropped out of a hypothetical cocaine trial. There are multiple ways of calculating a “percent urine positive/negative variable”: For example, if the specimen was positive for cocaine, the participant would be considered to have 100% positive urines; if the trial reported percent negative, rather than positive specimens, the variable would indicate 0%. However, if the single urine submitted was positive and the variable was operationalized based on number of *expected* specimens, in a 12-week trial collecting specimens weekly the percentage would be 8%. If the trial collected urines 3x weekly, the percentage would be 3%. The widely varying interpretation based on how these seemingly simple variables are calculated underscores the need for clarity in clinical trials reporting regarding how missing data is handled as well as thoughtful sensitivity analyses.

Finally, calculation of maximum days of consecutive abstinence or consecutive number of negative urines is straightforward when missing data are minimal, but extremely complex in studies where the level of missingness is substantial. For example, a participant who was completely abstinent during an 84 day treatment protocol and who provided 81 days of self-report data, but was missing 3 data points, could have a value ranging from 28 days of consecutive abstinence to 81 days of abstinence, depending on the pattern of missingness.

**1.5.2 Verifiability**—In contrast to many other disorders, the addictions have the advantage of reliable biological indicators of use that detect substance use over varying periods, depending on the source of the sample, type of assay, and type of drug use targeted (Huestis et al., 1995; Schwartz, 1988). In trials of cocaine use, quantitative urines collected 2–3 times per week can provide accurate estimates of use, providing the analytic method controls for



possible carry-over of use (Preston et al., 1997). However, frequent specimen collection with quantitative analysis is costly, and in many cases it is necessary to provide incentives to participants in order to sustain collection of multiple urine specimens per week over the course of a several-month trial. Moreover, in the interest of independence of observations, the urine collection schedule must be designed to minimize effects of such overlap (Preston et al., 1997, 1999; Schuler et al., 2009; Schwilke et al., 2010).

The reliability and accuracy of self-report measures of illicit substance use remains a controversial topic, but there is general agreement that self-reports are more accurate when collected in conjunction with a biological indicator (Del Boca and Noll, 2000; Vocci, 2008). While most consensus panels advocate combining urine and self-report data to enhance accuracy (Donovan et al., 2012), doing so is not at all straightforward: In order to reconcile the two sources of data, it is important that the urine specimens be collected at optimum frequency (at least twice weekly). Ability to combine sources of data is constrained by the analytical method and the type of drug considered, as most assays can identify only fairly recent drug use, thus weekly or monthly urine specimen collection leaves ‘gaps’ in periods of confirmability, while bi- and tri-weekly collection can overestimate use due to carry-over effects (Preston et al., 1997; Somoza et al., 2008).

In addition, there is very little consensus regarding how self-report and urine data should be combined when the two sources of information do not agree: The recommendation “believe the source that indicates use” is sound but difficult to operationalize, particularly for continuous variables (Korte et al., 2011). Consider the case in which a participant reports using no marijuana for 2 consecutive weeks. With a cannabis- negative toxicology result at week 1 and positive report at week 2, the self-report of abstinence is clearly not accurate, but it is not clear how many days of use occurred. In the absence of a correction (e.g., through re-interviewing the participant), the investigator would need to make assumptions regarding how many days of marijuana use should be interpolated (one, two, or all seven?). Dichotomous variables are somewhat less complex in these cases of discrepancy, as a single positive urine sample will result in a clear ‘no’ for an indicator such as complete abstinence. Similarly, the ‘abstinent at least X continuous days’ indicators are fairly straightforward, as the presence of a positive urine during the time period in question will yield ‘no’.

**1.5.3 Sensitivity to missing data**—As noted earlier, the validity of many of the candidate outcome indicators is to a large extent dependent on the availability of a full data set, and their sensitivity and vulnerability to missing data is largely related to the assumptions made regarding how missing data are handled. The implications of missing data in analysis and reporting of clinical trials has been covered elsewhere (Lavori et al., 2008; Siddique et al., 2008); hence we focus on the relative effects of missing data primarily as they affect the indicators discussed herein. The key consideration underlying the “intention to treat principle” is that data be collected regardless of the participant’s level of adherence, retention, or dropout from a clinical trial (Nich and Carroll, 2002). In practice, however, it is generally difficult to locate and obtain data from dropouts after they leave the trial.

Thus, a critical issue for most trials becomes clarifying the assumptions made regarding the indicator after the participant drops out or is withdrawn, because most strategies used for

estimating use after dropout (in the absence of data from the participant) are of questionable validity. In particular, the frequently used strategy of ‘carrying forward’ the last value recorded from the participant is untenable and highly problematic (Lavori, 1992; Lavori et al., 2008; Mallinckrodt et al., 2004). Similarly, coding all data points from the point of dropout as indicating drug use is unlikely to be accurate, as drug use trajectories after treatment cessation vary widely (Hser et al., 2001, 2006; Morral et al., 1997). This approach is particularly problematic when there is differential attrition across treatment conditions. Statistical approaches such as multiple imputation strategies address these issues to some extent, but remain constrained by the level of missing data (Mackinnon, 2010; Spratt et al., 2010).

Regarding continuous variables, as noted earlier, use of different assumptions regarding missing data can yield widely different estimates of use. Consider a participant who dropped out of a 12-week trial after 4 weeks, with 2 of 4 of the weekly urine specimens collected indicating use, and who reported using once a week. Percent days abstinent could be calculated at least 2 different ways (7/28 or 25% if based on within treatment data; but  $7+56/84$ , or 75%, if consistent use was assumed after dropout). Percent negative urine specimens could also be calculated multiple ways (50% if based on urines given, 16% if based on samples expected).

Reduction measures are also highly vulnerable to missing data, as they tend to require full data over the time period of interest. For example, in order to evaluate the reduction in cocaine use outcome, both the baseline “days of cocaine use in the past 28” and the treatment endpoint “days of cocaine use in the past 28” must both be complete. The dichotomous measures (continuous abstinence, continuous abstinence over a given period) are less vulnerable to missing data, as they are generally based on the assumption of use when data are missing.

**1.5.4 Cost of measurement**—The relative costs of different approaches to assessment in randomized clinical trials in the addictions have been reviewed at length elsewhere (Babor et al., 2000). In general terms, self-report is generally seen as the least costly, followed by collateral reports and then biological indicators. However, costs of the various methods are in reality much more complex, and one often “gets what one pays for” in that obtaining accurate data often increases cost, regardless of the method used. Quantitative urinalyses, which allow estimation of new episodes of drug use, are much more costly than qualitative analyses. Similarly, self-report is generally seen as inexpensive and flexible relative to urinalysis, but virtually all procedures used to increase accuracy of self-reports also increase their costs. This includes the costs of more frequent data collection and with it, larger incentives to participants to reimburse the time spent completing assessments, costs of more experienced and highly trained research staff, and costs of collecting and analyzing urine toxicology data against which to check accuracy of self-reports. Thus, even among self-report methods, there are gradients of cost based on frequency of collection, whether the measure is obtained via completing a form or through interview, and the frequency and type of biological specimens required for validation. Variables such as retention in treatment and reduction in frequency of use are probably the least expensive to obtain, as they take comparatively little time to assess and cannot yet be verified biologically.

## 1.6 Summary

Characteristics of the 15 candidate indicators across these criteria are summarized in Table 1, which also includes details on how they were operationalized. Overall, each indicator has strengths and weaknesses, and no single one emerges as clearly superior based on the features proposed here. The continuous variables are highly complex, as are the assumptions underlying them, but they are most frequently used in reports of clinical trials and have the benefit of sensitivity and added statistical power (Deyi et al., 1998; Snapinn and Jiang, 2007). However, continuous indicators of outcome are often not seen as particularly compelling or informative to clinicians or policy makers (Kraemer et al., 2003a; Miller and Manuel, 2008; Winchell et al., 2012). In contrast, the dichotomous variables have several strengths when evaluated in this respect, but tend to be less sensitive and reduce power for most analyses.

As noted earlier, few of these indicators have been evaluated empirically. Hence, a number of important questions regarding key properties of these variables have rarely if ever been studied. To address these issues, we evaluated these 15 candidate indicators using a dataset compiled from 5 independent trials of behavioral and pharmacologic treatments for cocaine dependence in terms of two key characteristics that can be evaluated empirically: First, *sensitivity to the effects of a range of interventions, both behavioral and pharmacological*. While in any single study it is crucial to include outcome indicators that are theoretically linked to the goals and putative mechanisms of the intervention(s) studied, lack of a widely used common indicator as basis on which to compare outcomes across studies/clinics hampers progress in the field. At least some of the resistance to adoption of common outcome indicators may be concern that some measures may not be sensitive to effects of particular treatments. Thus, indicators that detect significant effects of multiple forms of intervention may be more likely to be accepted by clinicians and researchers and therefore adopted more broadly. Second, *association with longer-term outcomes*. Arguably the most important characteristic of an outcome indicator, we evaluated the degree to which each indicator was associated with cocaine use and general functioning during follow-up.

## 2. METHODS

### 2.1 Overview of the five trials

Rather than compare these indicators' performance across multiple independent studies, data from 5 randomized clinical trials of a range of treatments for cocaine dependence were pooled in order to describe relative performance of the indicators in a large, heterogeneous sample. These five trials (Carroll et al., 1998, 2004, 2008, 2012, under review) shared a number of common characteristics: All were RCTs for cocaine dependent outpatients and used similar inclusion/exclusion criteria. All but one trial was of 12 weeks in duration and included a one-year follow-up with in-person interviews and urine specimen collection at one, three-, six- and 12 months. All used a common assessment battery which included the ASI (McLellan et al., 2006), administered monthly throughout treatment and at each follow-up. Urine specimens were collected at least weekly during treatment and at each follow-up. The Substance Use Calendar, an adaptation of the TLFB (Robinson et al., 2012; Sobell and Sobell, 1992), provided self-reports of cocaine use on a daily basis for the 30 days preceding

baseline assessment, daily throughout treatment, and through the terminal follow-up interview. Urine data were generally consistent with self-report; rates of cocaine-positive specimens in cases where the participant had denied use in the past three days were low and consistent across studies (range 8.8 to 16.7%). Study designs, sample sizes, frequency of urine sample collection and citations for the original trial reports and follow-up studies are provided in Figure 1.

## 2.2 Calculation of within-treatment and follow-up outcome indicators

Descriptions of how the 15 candidate indicators were operationalized for these analyses are provided in Table 1. Follow-up outcomes included days of cocaine use in the 28 days preceding the 1, 3, 6, and 12 month follow-up, as well as complete abstinence from cocaine for the entire follow-up period. Urine specimens were collected at each follow-up interview; however, as these reflect only very recent cocaine use, these were used primarily as a check of self-reports. The proxy measure of ‘good functioning’ within treatment was also calculated over the course of follow-up as a dichotomous variable derived from the ASI as described above.

## 2.3 Data analyses

The primary purpose of this report was to provide a descriptive comparison of the 15 candidate indicators in terms of sensitivity to treatment effects and prediction of cocaine use and functioning during follow-up in a large heterogeneous sample. Hence, because we were not testing specific hypotheses regarding efficacy of specific treatments or conducting a meta-analysis (and thus simply replicating the original study reports), corrections for multiple comparisons or use of inferential statistics were not made. Sensitivity of the candidate outcomes to the effects of specific treatments, with separate analyses for the behavioral versus pharmacologic approaches, was evaluated with specific contrasts for the comparisons of interest (medication versus no medication; behavior therapy versus control) using chi-square and ANOVA. Baseline severity was included as a covariate in those analyses; because results were similar and would not alter interpretation, these are not described here. Pearson product-moment and point biserial correlations were used to describe relationships between the indicators and the follow-up outcomes. It should be noted that this approach differs from those used in the original study reports, which used longitudinal GLM models (e.g., days of cocaine use by week) and the full randomized samples; the data described here are limited to the cross-sectional candidate indicators and the subset of participants in each trial who contributed follow-up data.

## 3. RESULTS

### 3.1 Study samples, within treatment and follow-up outcomes

Descriptive data on participants’ demographic characteristics as well as substance use, psychiatric and general functioning at baseline across the five studies are presented in Supplemental Table 1<sup>1</sup>. The pooled sample (N=434) had a mean age of 36.5; about one-

---

<sup>1</sup>Three Supplemental tables are included: Supplemental Table 1 provides demographic data by study and for the total sample; Table 2 provides intercorrelations among the 15 candidate indicators, and Table 3 provides detailed information on medication effects by study; they can be found by accessing the online version of this paper at <http://dx.doi.org> and by entering doi:...

third were female, one-half Caucasian and 39% African American, and 78% completed high school. About half were unemployed, and 16% were referred by the criminal justice system. Participants reported using cocaine an average of 13.4 days of the last 28, and had used cocaine regularly for about 9 years prior to treatment. The majority (76%) had a lifetime history of alcohol abuse/dependence; a smaller percentage met DSM-IV criteria for lifetime depression (19%) or anxiety (10%) disorders. In terms of non-substance related problems on the ASI, the participants reported most severe problems in employment (mean composite score .61), followed by the family (.19) and psychological (.18) domains. Other than Study C, which was composed of cocaine-dependent methadone-maintained individuals, opioid use was minimal. As expected, there was substantial variability across studies in terms of most baseline characteristics.

The 15 candidate outcome indicators by study are shown in Table 2. The retention indicator (days in treatment) for the 84 day studies (Studies A, B, C, and E) ranged from 41 to 69 days. Higher retention in Study C likely reflects it being conducted in the context of a methadone maintenance program. Percent of cocaine-negative urine specimens ranged widely (23% to 54% across studies). Maximum days of consecutive abstinence ranged from a mean of 17 to 28 days. Self-reported percent days of abstinence ranged from 60% to 85%. Maximum days of consecutive abstinence during the last 2 weeks of treatment ranged from 7 to 10.

For the dichotomous indicators, percentages of participants who reported being completely abstinent for the duration of treatment ranged from 5% to 24% across the studies. Percentage reporting abstinence for 3 or more continuous weeks ranged from 23 to 65%; percentage of those reporting abstinence for one or more weeks ranged from 63 to 81%. Percentage of participants who completed treatment and were abstinent during the last week of treatment ranged from 22% to 55%. For the reduction variables, the mean percent reduction in frequency of cocaine use from the baseline month to the final month of treatment ranged from 48% to 70%. Rates of participants who achieved 50% or more reduction in cocaine use ranged from 16 to 33%, and for 75% reduction or more, rates ranged from 3 to 24%.

Follow-up outcomes are also presented in Table 2. Across studies, participants reported using cocaine about 5 days each month throughout the follow-up period. For the four studies with a one year follow-up, percentages of participants who reported abstinence for the full duration ranged from 7 to 19%. The rate of complete abstinence for the study that followed participants only up to 6 months post treatment (Study D) was 32%. Finally, a minority of participants met criteria for the proxy indicator of 'good functioning' in the last month of treatment (ranging from 0 to 25% across studies). Rates of individuals who met this definition increased through follow-up; at the end of 6 months, between 13 and 27% met this criterion; rates ranged from 11 to 39% for the four studies that followed participants out to one year.

Simple correlations among the candidate outcomes are shown in Supplemental Table 2<sup>2</sup>; gradations in shading are used in order to simplify interpretation (darker shades of grey

---

<sup>2</sup>Supplementary material can be found by accessing the online version of this paper at <http://dx.doi.org> and by entering doi:...

reflect higher correlations/stronger relationships). As expected, the majority of variables were correlated with each other at a statistically significant level. Correlations were particularly strong among the variables tapping similar dimensions (e.g., maximum days of consecutive abstinence; 21, 14, or 7 days of abstinence; and abstinence at the end of treatment). Variables that tended to have comparatively lower correlations with other indicators, therefore suggesting greater independence, included retention as well as the two dichotomous indicators of reduction in frequency of cocaine use (50%/75% reduction).

### 3.2 Sensitivity to effects of pharmacologic and behavioral therapies

**3.2.1 Effects of medication**—As four of the five studies evaluated disulfiram (Studies A, B, C, and E), it was possible to explore outcomes by condition (disulfiram, placebo, no medication). For the pooled sample in these four studies (N=396), six of the 15 indicators suggested significant main effects; four of which indicated better outcomes for those assigned to disulfiram as shown in Table 3. Three of these were continuous (retention, maximum consecutive days of abstinence, and maximum days of abstinence during the last two weeks of treatment) and one was dichotomous (abstinent and completed treatment). Percent of participants who attained three or more weeks of abstinence approached but did not reach statistical significance. The reduction variables, as well as complete abstinence, did not indicate significant effects by medication condition.

Note that the original studies did not all report significant effects for disulfiram with respect to the control condition (Studies A and B reported main effects for disulfiram over the control, while in Studies C and E findings were weaker or confined to subgroups; Supplemental Table 3<sup>3</sup> presents results across the outcome indicators by study). Considering each study separately, these results replicate those described in each primary study report in terms of the outcomes that indicated statistically significant effects. That is, the original reports describing Studies A and B indicated a significant effect for disulfiram on maximum days of abstinence and attaining three or more weeks of continuous weeks of abstinence; in the current dataset there was a medication effect for number and percent completely abstinent the last 2 weeks of treatment for these studies. Conversely, in the single protocol that did not indicate a main effect of disulfiram (Study C), there were no significant differences for any indicator.

**3.2.2 Effects of behavioral therapies**—The five RCTs evaluated three different behavioral therapies (TSF, CBT, and CM), alone and in various combinations, and used a range of designs and control groups, as shown in Figure 1. This complicated evaluation of the 15 candidate indicators' ability to detect effects of specific behavioral therapies, in a manner similar to the difficulty of conducting meta-analyses for behavioral therapies where different types of control groups are used. That is, magnitude of effects in behavioral therapies research is to a large extent dependent upon the nature of the condition to which it is compared (Kazdin and Bass, 1989; i.e., no treatment/wait list controls are more likely to result in larger effect sizes than attention control conditions or active controls). Therefore,

---

<sup>3</sup>Supplementary material can be found by accessing the online version of this paper at <http://dx.doi.org> and by entering doi:...

we evaluated the 15 indicators across the full pooled sample (N=434), contrasting each of the three therapies of interest with one pooled control condition.

Results are shown in Table 4, which suggests that among the six continuous variables, significant effects were seen for percentage of cocaine-negative urine specimens and self-reported percent days of abstinence, two of the most widely-used outcome indicators. Maximum days of abstinence approached statistical significance ( $p=.06$ ), as did the continuous measure of reduction ( $p=.09$ ). Among the nine dichotomous variables, five indicated statistically significant differences between the experimental therapy and control: abstinence in the last two weeks of treatment, treatment completion with abstinence at the end of treatment, attaining 2/3 or more weeks of continuous abstinence, reduction of frequency of cocaine use by 75% or more, and the variable indicating good general functioning. Those that appeared least sensitive included treatment retention and complete abstinence.

For those variables where there were significant differences, *post hoc* pairwise comparisons were consistent with previous systematic reviews and meta-analyses (Dutra et al., 2008) which have indicated particularly strong within-treatment effects for CM, generally followed by CBT, with less consistent effects for TSF. As with the medication analyses above, selection of different primary indicators would lead to different conclusions regarding efficacy of treatments. For example, CM, which specifically rewards submission of consecutive drug-free urine specimens, had better outcomes than CBT, TSF, and the pooled control condition, on variables reflecting those outcomes (percent negative urine specimens, and various durations of consecutive abstinence). CBT generally appears strongest for outcomes reflecting abstinence in the last 2 weeks of treatment, consistent with the literature suggesting delayed emergence of effects (Carroll et al., 1994).

### 3.3 Association with follow-up outcomes

Relationships between the outcome indicators and follow-up outcomes (frequency of cocaine use and 'good functioning' by follow-up) are shown in Table 5. To simplify interpretation, strength of relationships is indicated by gradations in shading. Among the continuous variables, the indicators most consistently/strongly related to frequency of cocaine use and abstinence across the follow-up included maximum length of abstinence in the last 2 weeks of treatment, maximum days of abstinence during treatment, and percent days of cocaine use, all commonly used indicators of outcome in multiple trials. Among the dichotomous indicators, those most strongly and consistently related to cocaine use during follow-up included whether the participant attained 3 or more weeks of abstinence. Within this set of candidate indicators, retention in treatment, complete abstinence, and the dichotomous reduction indicators (50/75% reduction) were weakly and inconsistently related to cocaine use during follow-up. The proxy indicator of 'good functioning' within treatment showed moderate relationship to cocaine use during follow-up.

Regarding correlation of the indicators with the proxy measure of 'good functioning' during follow-up, the three variables most strongly and consistently related to better outcomes during follow-up were all dichotomous: three or more weeks of continuous abstinence, 2 or more weeks of continuous abstinence, and good functioning at the end of treatment.

Retention in treatment, the two dichotomous reduction variables and complete abstinence within treatment were very weakly related to good functioning during follow-up.

## 4. DISCUSSION

With the broad aim of informing efforts to adopt a more standardized and clinically relevant approach to reporting outcome measures in the addictions, this report describes a systematic empirical comparison of commonly-used and novel indicators of cocaine treatment outcome measures. In terms of sensitivity to the effects of pharmacologic and behavioral therapies, the most commonly used outcome measures (percent negative urine screens; percent days of abstinence), performed relatively well in that they were sensitive to the effects of behavioral and pharmacologic treatments evaluated (and were consistent with findings reported for the original trials). However, many of the other indicators were less sensitive to the effects of these treatments. Finally, the indicators that had the highest and most consistent relationships with cocaine use frequency and general functioning during follow-up were those that reflected achievement of sustained periods of abstinence, particularly at the end of treatment.

The primary aim of this report was to compare multiple outcome indicators across multiple dimensions to inform efforts to identify clinically useful indicators of outcome for intervention studies in the addictions. These data underscore that all indicators have some strengths and weaknesses and suggest why adoption of a single indicator has been so elusive: Even among these 5 trials, which were conducted by the same research group, with largely parallel methods and assessment batteries, there was marked variability in baseline characteristics and outcome across the candidate indicators. This level of heterogeneity is inherent in the field as well, and underscores why investigators often have widely varying opinions as to selection of appropriate indicators (and hence the limits of expert consensus panels alone in selecting one common indicator). The variability in this pooled dataset also highlights the limitations of using data from any single trial to identify robust predictors or standard outcome indicators.

### 4.1 Poorer performing candidate indicators

On the basis of dimensions evaluated here, our data suggest several indicators as candidates for elimination from the potential pool of common outcome indicators: First, retention in treatment, while arguably the most attractive based on practical considerations (very inexpensive to obtain, simple to verify, easy to calculate, and relatively impervious to missing data), performed poorly. Retention also had relatively low levels of relationships to the other indicators. Although it was one of the few variables demonstrating significant differences between disulfiram and the control conditions, it did not discriminate among the behavioral therapies tested and was a particularly weak predictor of cocaine use or functioning in the year following treatment. Although retention has been frequently used as an outcome indicator in important treatment trials (Simpson et al., 1997, 1999), these data are consistent with recent reports that suggest retention may have limited utility as an outcome measure in terms of its relationship with cocaine use and functioning during follow-up (Hien et al., 2012; Walker, 2009).



Another indicator which, although relatively easy to calculate and straightforward to interpret, fared relatively poorly in this set of analyses was complete abstinence within treatment. Although achieved by a reasonable minority of the sample (14%), complete abstinence was not sensitive to the effects of the behavioral and pharmacologic treatments evaluated here and was a relatively weak predictor of post-treatment cocaine use and general functioning. Given the emphasis placed on complete abstinence in many clinical treatment programs, the weak performance of this indicator is surprising and worthy of further investigation. In contrast, the variables indicating abstinence achieved at the end of treatment (abstinence in the last two weeks of treatment, completing treatment and abstinent in the last week, maximum days of abstinence in the last two weeks) were more sensitive to the effects of treatment and more consistently related to less cocaine use and better functioning in follow-up. This may suggest some inherent value in striving to achieve abstinence, or learning from episodes of relapse within treatment.

Other weaker indicators included the variables associated with reduction in use. These not only were cumbersome to calculate, highly susceptible to missing data, and difficult to verify biologically (unless quantitative urinalysis with an appropriate collection schedule is available) but also proved insensitive to the effects of the therapies evaluated here. Moreover, they had relatively weak relationships to indicators of cocaine use and functioning during follow-up. While the concept of focusing on reduction, rather than abstinence, in intervention trials is relatively novel and has the benefit of being tied to traditional approaches of measuring clinical significance (Jacobson et al., 1999; Jacobson and Truax, 1991; Marsden et al., 2009), reduction in frequency of cocaine use was, as calculated here, of limited use in discriminating treatments or predicting cocaine use during follow-up.

#### 4.2 Better-performing candidates

Several variables performed comparatively well, including the continuous indicators that have been used most frequently in the treatment outcome literature (maximum days of consecutive abstinence, percent days abstinence, abstinence at the end of treatment, and percent of negative urine specimens). These were generally sensitive to the effects of at least some of the treatments evaluated and were relatively good predictors of follow-up. Although these measures have limitations, these data should be of some reassurance to the field. It therefore seems reasonable to continue to use them, provided they are consistently defined and missing data are minimized.

Regarding the dichotomous indicators, attaining a period of abstinence for three or more weeks at any time during the treatment episode performed relatively well. It was consistently significantly related to cocaine use and 'good functioning' during follow-up, perhaps the most important of our criteria. Although not sensitive to effects of disulfiram for the pooled sample, it did replicate findings from the individual studies. In terms of the behavioral therapies, it was also sensitive to the effects of CM. Finally, as one of the aims of this report was to use a combined dataset and empirically based criteria to identify a dichotomous indicator as a candidate standard outcome of clinical significance, it is notable

that ‘three or more weeks of abstinence’, a ‘success indicator’ identified very early on in cocaine treatment research (Gawin et al., 1989) performed relatively well.

In contrast to variables focused solely on cocaine use, this report also introduced a proxy measure of ‘good functioning’, operationalized as a composite of no cocaine use with no reported legal, employment, psychological or family problems in the 28 days prior to interview. This variable, while requiring validation and replications, also suggested some promise, as it identified a reasonable proportion of individuals at each assessment point (0% of the pooled sample at baseline, 11% at end of treatment, and 21% at one year follow up).

### 4.3 Strengths and limitations

As one of the first systematic empirical comparisons of commonly-used and more novel outcome indicators for cocaine intervention trials, this report had multiple strengths. Data were drawn from randomized, well-controlled clinical trials that used a common assessment battery and identical approach to measurement, with substantial representation of women (33%) and ethnic minorities (48%). It included a variety of empirically-validated behavioral therapies and one pharmacotherapy, disulfiram, which is one of a small field of agents demonstrating promise in the treatment of cocaine dependence. Unlike many pharmacotherapy trials, these trials included one-year follow-up data on both cocaine use and general functioning. Finally, it was the first to systematically compare outcome indicators based on standard definitions of clinical significance (e.g., reductions in symptoms of 2 standard deviations or more; 75% reduction) to other indicators. Moreover, the estimated level of reduction in frequency of cocaine use in the pooled sample (58% overall) is comparable to another recently reported estimate from another large dataset (Marsden et al., 2009).

There were some weaknesses as well; the dataset, while comparatively large, was drawn from a single research group and focused only on cocaine treatment trials; hence validation by other groups and with datasets focused on other types of illicit drug use is necessary. The trials included multiple empirically-validated behavioral treatments, but only a single medication disulfiram, that has not been uniformly effective. As different indicators may be differentially sensitive to the effects of different medications, similar analyses may be useful across other promising mediations. We did not include some candidate indicators that have substantial support (craving, quality of life; Tiffany et al., 2000; Tiffany and Wray, 2012), as these variables were not collected identically across all 5 trials. We operationalized and defined each variable carefully; however, there are multiple, reasonable ways to calculate many of these variables (e.g., percent days of abstinence, percent positive/negative urine specimens). Alternate ways of calculating these indicators may have led to different conclusions, underscoring the importance of sensitivity analyses in clinical trials and clear specification regarding handling of missing data.

### 4.4 Summary

As the first empirical comparison of multiple widely-used outcome indicators across multiple trials with well-characterized treatments, the data presented here represent a modest proposal and challenge to the field move toward use and reporting of a smaller empirically

strong set of common outcome measures. The outcome indicators emerging from these analyses with the most support were closely related to those that have already been adopted in the fields of tobacco and alcohol use, that is, abstinence-based outcomes, such as abstinence at the end of treatment or a durable period of continuous abstinence. While variability across samples and designs dictates continued selection of outcomes which are matched to the theoretical mechanisms of action of the treatment and design of the trial, consistent reporting of one common outcome measure would greatly accelerate comparison across trials and benchmarking with clinical outcomes and hence are essential to systematic progress within our field.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

### Role of Funding Source

Support for this study was provided by a supplement to National Institute on Drug Abuse grant R01 DA015969-09S1 (Carroll, PI), as well as grants P50-DA09241 and U10 DA015831 (Carroll, PI). Dr. DeVito was supported by K12DA031050 (BIRCWH, CM Mazure, PI). Dr. Decker was supported by the Office of Academic Affiliations, Advanced Fellowship in Mental Illness Research and Treatment, Department of Veterans Affairs, the Veterans Affairs Connecticut Healthcare System, and the Department of Veterans Affairs New England Mental Illness Research, Education, and Clinical Center (M. Sofuoglu, Director). NIDA and the VA had no further role in study design; in the collection, analysis and interpretation of data; in the writing of the report; or in the decision to submit the paper for publication.

We gratefully acknowledge Tami Frankforter and Karen Hunkele for their invaluable help with assembling the dataset and the participants, research staff and colleagues who participated in the original clinical trials.

## References

- Allen JP. Measuring outcome in interventions for alcohol dependence and problem drinking: executive summary of a conference sponsored by the national institute on alcohol abuse and alcoholism. *Alcohol Clin Exp Res.* 2003; 27:1657–1660. [PubMed: 14574237]
- Anton RE, Randall CL. Measurement and choice of drinking outcome variables in the COMBINE Study. *J Stud Alcohol Suppl.* 2005:104–109. [PubMed: 16223062]
- Anton RF, O'Malley SS, Ciraulo DA, Cisler RA, Couper D, Donovan DM, Gastfriend DR, Hosking JD, Johnson BA, LoCastro JS, Longabaugh R, Mason BJ, Mattson ME, Miller WR, Pettinati HM, Randall CL, Swift RM, Weiss RD, Williams LD, Zweben A. for the COMBINE Study Research Group. Combined pharmacotherapies and behavioral interventions for alcohol dependence: The COMBINE Study. *JAMA.* 2006; 2006:2003–2017. [PubMed: 16670409]
- Babor TF, Steinberg K, Anton RF, Del Boca FK. Talk is cheap: measuring drinking outcomes in clinical trials. *J Stud Alcohol.* 2000; 61:55–63. [PubMed: 10627097]
- Carroll KM, Ball SA, Martino S, Nich C, Babuscio T, Gordon MA, Portnoy GA, Rounsaville BJ. Computer-assisted cognitive-behavioral therapy for addiction. A randomized clinical trial of 'CBT4CBT'. *Am J Psych.* 2008; 165:881–888.
- Carroll KM, Ball SA, Martino S, Nich C, Babuscio TA, Rounsaville BJ. Enduring effects of a computer-assisted training program for cognitive-behavioral therapy: a six-month follow-up of CBT4CBT. *Drug Alcohol Depend.* 2009; 100:178–181. [PubMed: 19041197]
- Carroll KM, Fenton LR, Ball SA, Nich C, Frankforter TL, Shi J, Rounsaville BJ. Efficacy of disulfiram and cognitive behavior therapy in cocaine-dependent outpatients: a randomized placebo-controlled trial. *Arch Gen Psychol.* 2004; 61:264–272.

- Carroll KM, Nich C, Ball SA, McCance E, Frankforter TL, Rounsaville BJ. One-year follow-up of disulfiram and psychotherapy for cocaine-alcohol users: sustained effects of treatment. *Addiction*. 2000; 95:1335–1349. [PubMed: 11048353]
- Carroll KM, Nich C, Ball SA, McCance E, Rounsaville BJ. Treatment of cocaine and alcohol dependence with psychotherapy and disulfiram. *Addiction*. 1998; 93:713–727. [PubMed: 9692270]
- Carroll KM, Nich C, Petry NM, Eagan D, Shi J, Ball SA, Herman A, Sofuoglu M. Disulfiram and contingency management to enhance CBT for cocaine dependence: effects on cocaine use and preliminary evidence for interactions with dopamine B-hydroxylase polymorphism. 2013 Under review.
- Carroll KM, Nich C, Shi JM, Eagan D, Ball SA. Efficacy of disulfiram and Twelve Step Facilitation in cocaine-dependent individuals maintained on methadone: a randomized placebo-controlled trial. *Drug Alcohol Depend*. 2012; 126:224–231. [PubMed: 22695473]
- Carroll KM, Rounsaville BJ, Nich C, Gordon LT, Wirtz PW, Gawin F. One-year follow-up of psychotherapy and pharmacotherapy for cocaine dependence. Delayed emergence of psychotherapy effects. *Arch Gen Psychol*. 1994; 51:989–997.
- Ciraulo DA, Piechniczek-Buczek J, Iscan EN. Outcome predictors in substance use disorders. *Psychiatr Clin North Am*. 2003; 26:381–409. [PubMed: 12778840]
- Cisler RA, Kowalchuk RK, Saunders SM, Zweben A, Trinh HQ. Applying clinical significance methodology to alcoholism treatment trials: determining recovery outcome status with individual- and population-based measures. *Alcohol Clin Exp Res*. 2005; 29:1991–2000. [PubMed: 16340456]
- Cisler RA, Zweben A. Development of a composite measure for assessing alcohol treatment outcome: operationalization and validation. *Alcohol Clin Exp Res*. 1999; 23:263–271. [PubMed: 10069555]
- Crits-Christoph P, Siqueland L, Blaine JD, Frank A, Luborsky L, Onken LS, Muenz L, Thase ME, Weiss RD, Gastfriend DR, Woody G, Barber JP, Butler S, Dale D, Salloum I, Bishop S, Najavits L, Lis J. Psychosocial treatments for cocaine dependence: results of the National Institute on Drug Abuse Collaborative Cocaine Study. *Arch Gen Psychol*. 1999; 56:495–502.
- DeBusk RF, Miller NH, Superko HR, Dennis CA, Thomas RJ, Lew HT, Berger WE 3rd, Heller RS, Rompf J, Gee D, Kraemer HC, Bandura A, Ghandour G, Clark M, Shah RV, Fisher L, Taylor CB. A case-management system for coronary risk factor modification after acute myocardial infarction. *Ann Intern Med*. 1994; 120:721–729. [PubMed: 8147544]
- Del Boca FK, Noll JA. Truth or consequences: the validity of self-report data in health services research on addictions. *Addiction*. 2000; 95:S347–360. [PubMed: 11132362]
- Devos-Comby L, Lange JE. “My drink is larger than yours?”. A literature review of self-defined drink sizes and standard drinks. *Curr Drug Abuse Rev*. 2008; 1:162–176. [PubMed: 19630715]
- Deyi BA, Kosinski AS, Snapinn SM. Power considerations when a continuous outcome variable is dichotomized. *J Biopharm Stat*. 1998; 8:337–352. [PubMed: 9598427]
- Donovan DM, Bigelow GE, Brigham GS, Carroll KM, Cohen AJ, Gardin JG, Hamilton JA, Huestis MA, Hughes JR, Lindblad R, Marlatt GA, Preston KL, Selzer JA, Somoza EC, Wakim PG, Wells EA. Primary outcome indices in illicit drug dependence treatment research: systematic approach to selection and measurement of drug use end-points in clinical trials. *Addiction*. 2012; 107:694–708. [PubMed: 21781202]
- Donovan, DM.; Marlatt, GA., editors. *Assessment of Addictive Behaviors*. 2. The Guilford Press; New York: 2005.
- Dutra L, Stathopoulou G, Basden SL, Leyro TM, Powers MB, Otto MW. A meta-analytic review of psychosocial interventions for substance use disorders. *Am J Psychol*. 2008; 165:179–187.
- Ehrman RN, Robbins SJ, Cornish JW. Results of a baseline urine test predict levels of cocaine use during treatment. *Drug Alcohol Depend*. 2001; 62:1–7. [PubMed: 11173162]
- Falk D, Wang XQ, Liu L, Fertig J, Mattson M, Ryan M, Johnson B, Stout R, Litten RZ. Percentage of subjects with no heavy drinking days: evaluation as an efficacy endpoint for alcohol clinical trials. *Alcohol Clin Exp Res*. 2010; 34:2022–2034. [PubMed: 20659066]
- Finney JW, Moyer A, Swearingen CE. Outcome variables and their assessment in alcohol treatment studies: 1968–1998. *Alcohol Clin Exp Res*. 2003; 27:1671–1679. [PubMed: 14574240]

- Gawin FH, Kleber HD, Byck R, Rounsaville BJ, Kosten TR, Jatlow PI, Morgan C. Desipramine facilitation of initial cocaine abstinence. *Arch Gen Psychol.* 1989; 46:117–121.
- Hien DA, Morgan-Lopez AA, Campbell AN, Saavedra LM, Wu E, Cohen L, Ruglass L, Nunes EV. Attendance and substance use outcomes for the Seeking Safety program: sometimes less is more. *J Consult Clin Psychol.* 2012; 80:29–42. [PubMed: 22182262]
- Higgins ST, Wong CJ, Badger GJ, Haug-Ogden DE, Dantona RL. Contingent reinforcement increases cocaine abstinence during outpatient treatment and one year follow-up. *J Consult Clin Psychol.* 2000; 68:64–72. [PubMed: 10710841]
- Hser Y, Hoffman V, Grella CE, Anglin MD. A 33-year follow-up of narcotics addicts. *Arch Gen Psychol.* 2001; 58:503–508.
- Hser YI, Huang D, Brecht ML, Li L, Evans E. Contrasting trajectories of heroin, cocaine, and methamphetamine use. *J Addict Dis.* 2008; 27:13–21. [PubMed: 18956525]
- Hser YI, Stark ME, Paredes A, Huang D, Anglin MD, Rawson RA. A 12-year follow-up of a treated cocaine-dependent sample. *J Subst Abuse Treat.* 2006; 30:219–226. [PubMed: 16616166]
- Huestis MA, Cone EJ, Wong CJ, Umbricht A, Preston KL. Monitoring opiate use in substance abuse treatment patients with sweat and urine drug testing. *J Anal Toxicol.* 2000; 24:509–521. [PubMed: 11043653]
- Huestis MA, Elsohly M, Nebro W, Barnes A, Gustafson RA, Smith ML. Estimating time of last oral ingestion of cannabis from plasma THC and THCCOOH concentrations. *Ther Drug Monit.* 2006; 28:540–544. [PubMed: 16885722]
- Huestis MA, Mitchell JM, Cone EJ. Detection times of marijuana metabolites in urine by immunoassay and GC-MS. *J Anal Toxicol.* 1995; 19:443–449. [PubMed: 8926739]
- Hughes JR. A quantitative estimate of the clinical significance of treating tobacco dependence. *Am J Prev Med.* 39:285–286. [PubMed: 20709263]
- Hughes JR. Measurement of the effects of abstinence from tobacco: a qualitative review. *Psychol Addict Behav.* 2007; 21:127–137. [PubMed: 17563132]
- Hughes JR, Carpenter MJ, Naud S. Do point prevalence and prolonged abstinence measures produce similar results in smoking cessation studies? A systematic review. *Nicotine Tob Res.* 2010; 12:756–762. [PubMed: 20504946]
- Hughes JR, Keely JP, Niaura RS, Ossip-Klein DJ, Richmond RL, Swan GE. Measures of abstinence in clinical trials: issues and recommendations. *Nicotine Tob Res.* 2003; 5:13–25. [PubMed: 12745503]
- Humphreys K, McLellan AT. A policy-oriented review of strategies for improving the outcomes of services for substance use disorder patients. *Addiction.* 2012; 106:2058–2066. [PubMed: 21631620]
- Jacobson NS, Roberts LJ, Berns SB, McGlinchey JB. Methods for defining and determining the clinical significance of treatment effects: description, application and alternatives. *J Consult Clin Psychol.* 1999; 67:300–307. [PubMed: 10369050]
- Jacobson NS, Truax PA. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol.* 1991; 59:12–19. [PubMed: 2002127]
- Kazdin AE, Bass D. Power to detect differences between alternative treatments in comparative psychotherapy outcome research. *J Consult Clin Psychol.* 1989; 57:138–147. [PubMed: 2925964]
- Kendall PC, Marrs-Garcia A, Nath SR, Sheldrick RC. Normative comparisons for the evaluation of clinical significance. *J Consult Clin Psychol.* 1999; 67:285–299. [PubMed: 10369049]
- Korte JE, Magruder KM, Chiuhan CC, Logan SL, Killeen T, Bandyopadhyay D, Brady KT. Assessing drug use during follow-up: direct comparison of candidate outcome definitions in pooled analyses of addiction treatment studies. *Am J Drug Alcohol Abuse.* 2011; 37:358–366. [PubMed: 21854278]
- Kraemer HC, Morgan GA, Leech NL, Gliner JA, Vaske JJ, Harmon RJ. Measures of clinical significance. *J Am Acad Child Adolesc Psychol.* 2003a; 42:1525–1549.
- Kraemer HC, Morgan GA, Leech NL, Gliner JA, Vaske JJ, Harmon RJ. Measures of clinical significance. *J Am Acad Child Adolesc Psychol.* 2003b; 42:1524–1529.

- Lambert, MJ.; Bailey, RJ. Measures of clinically significant change. In: Cooper, H.; Camic, PM.; Long, DL.; Panter, AT.; Rindskopf, D.; Sher, KJ., editors. *APA Handbook of Research Methods in Psychology*. American Psychological Association; Washington, D.C: 2012. p. 147-160.
- Lavori PW. Clinical trials in psychiatry: should protocol deviation censor patient data? *Neuropsychopharmacol*. 1992; 6:39-48.
- Lavori PW, Brown CH, Duan N, Gibbons RD, Greenhouse J. Missing data in longitudinal clinical trials Part A: design and conceptual issues. *Psychiatr Ann*. 2008; 38:784-792. [PubMed: 19756211]
- Mackinnon A. The use and reporting of multiple imputation in medical research - a review. *J Intern Med*. 2010; 268:586-593. [PubMed: 20831627]
- Mallinckrodt CH, Watkin JG, Molenberghs G, Carroll RJ. Choice of the primary analysis in longitudinal clinical trials. *Pharm Stat*. 2004; 3:161-169.
- Marsden J, Eastwood B, Bradbury C, Dale-Perera A, Farrell M, Hammond P, Knight J, Randhawa K, Wright C. Effectiveness of community treatments for heroin and crack cocaine addiction in England: a prospective, in-treatment cohort study. *Lancet*. 2009; 374:1262-1270. [PubMed: 19800681]
- McCann DJ, Li SH. A novel, nonbinary evaluation of success and failure reveals bupropion efficacy versus methamphetamine dependence: reanalysis of a multisite trial. *CNS Neurosci Thera*. 2011; 18:414-418.
- McKay JR, Alterman AI, Koppenhaver JM, Mulvaney FD, Bovasso GB, Ward K. Continuous, categorical and time to event cocaine use outcome variables: degree of intercorrelation and sensitivity to treatment group differences. *Drug Alcohol Depend*. 2001; 62:19-31. [PubMed: 11173164]
- McLellan AT. Have we evaluated addiction treatment correctly? Implications from a chronic care perspective. *Addiction*. 2002; 97:249-252. [PubMed: 11964098]
- McLellan AT, Arndt IO, Metzger DS, Woody GE, O'Brien CP. The effects of psychosocial services in substance abuse treatment. *JAMA*. 1993; 269:1953-1959. [PubMed: 8385230]
- McLellan AT, Cacciola JC, Alterman AI, Rikoon SH, Carise D. The Addiction Severity Index at 25: origins, contributions and transitions. *Am J Addict*. 2006; 15:113-124. [PubMed: 16595348]
- McLellan AT, Lewis DC, O'Brien CP, Kleber HD. Drug dependence, a chronic medical illness: implications for treatment, insurance, and outcomes evaluation. *JAMA*. 2000; 284:1689-1695. [PubMed: 11015800]
- McLellan AT, Luborsky L, Cacciola J, Griffith J, Evans F, Barr HL, O'Brien CP. New data from the Addiction Severity Index. Reliability and validity in three centers. *J Nerv Ment Dis*. 1985; 173:412-423. [PubMed: 4009158]
- Miller PG, Miller WR. What should we be aiming for in the treatment of addiction? *Addiction*. 2009; 104:685-686. [PubMed: 19413777]
- Miller WR, Heather N, Hall W. Calculating standard drink units: international comparisons. *Addiction*. 1991; 86:43-47.
- Miller WR, Manuel JK. How large must a treatment effect be before it matters to practitioners? An estimation method and demonstration. *Drug Alcohol Rev*. 2008; 27:524-528. [PubMed: 18608445]
- Morrall AR, Iguchi MY, Belding MA, Lamb RJ. Natural classes of treatment response. *J Consult Clin Psychol*. 1997; 65:673-685. [PubMed: 9256569]
- National Institute on Drug Abuse. *Principles of Drug Addiction Treatment: A Research Based Guide*. NIDA; Bethesda, MD: 2007.
- Nich C, Carroll KM. Intention-to-treat meets missing data: implications of alternate strategies for analyzing clinical trials data. *Drug Alcohol Depend*. 2002; 68:121-130. [PubMed: 12234641]
- Peters EN, Nich C, Carroll KM. Primary outcomes in two randomized controlled trials of treatments for cannabis use disorders. *Drug Alcohol Depend*. 2011; 118:408-416. [PubMed: 21620591]
- Peters SA, Bots ML, den Ruijter HM, Palmer MK, Grobbee DE, Crouse JR 3rd, O'Leary DH, Evans GW, Raichlen JS, Moons KG, Koffijberg H. Multiple imputation of missing repeated outcome measurements did not add to linear mixed-effects models. *J Clin Epidemiol*. 2012; 65:686-695. [PubMed: 22459429]

- Preston KL, Huestis MA, Wong CJ, Umbricht A, Goldberger BA, Cone EJ. Monitoring cocaine use in substance-abuse-treatment patients by sweat and urine testing. *J Anal Toxicol*. 1999; 23:313–322. [PubMed: 10488917]
- Preston KL, Silverman K, Schuster CR, Cone EJ. Assessment of cocaine use with quantitative urinalysis and estimation of new uses. *Addiction*. 1997; 92:717–727. [PubMed: 9246799]
- Robinson SM, Sobell LC, Sobell MB, Leo GI. Reliability of the Timeline Followback for cocaine, cannabis, and cigarette use. *Psych Addict Behav*. in press.
- Schuler MS, Lechner WV, Carter RE, Malcolm R. Temporal and gender trends in concordance of urine drug screens and self-reported use in cocaine treatment studies. *J Addict Med*. 2009; 3:211–217. [PubMed: 20209029]
- Schwartz RH. Urine testing in the detection of drugs of abuse. *Arch Int Med*. 1988; 148:2407–2412. [PubMed: 3056313]
- Schwilke EW, Gulberg RG, Darwin WD, Chiang CN, Cadet JL, Gorelick DA, Pope HG, Huestis M. Differentiating new cannabis use from residual urinary cannabinoid excretion in chronic daily cannabis users. *Addiction*. 2010; 106:499–506. [PubMed: 21134021]
- Siddique J, Brown CH, Hedeker D, Duan N, Gibbons RD, Miranda J, Lavori PW. Missing data in longitudinal trials - part B, analytic issues. *Psychiatr Ann*. 2008; 38:793–801. [PubMed: 19668352]
- Simpson DD, Joe GW, Brown BS. Treatment retention and follow-up outcomes in the Drug Abuse Treatment Outcome Study (DATOS). *Psychol Addict Behav*. 1997; 11:294–307.
- Simpson DD, Joe GW, Fletcher BW, Hubbard RL, Anglin MD. A national evaluation of treatment outcomes for cocaine dependence. *Arch Gen Psychol*. 1999; 56:507–514.
- Snapinn SM, Jiang Q. Responder analyses and the assessment of a clinically relevant treatment effect. *Trials*. 2007; 8:31. [PubMed: 17961249]
- Sobell, LC.; Sobell, MB. Timeline followback: a technique for assessing self-reported alcohol consumption. In: Litten, RZ.; Allen, J., editors. *Measuring Alcohol Consumption: Psychosocial and Biological Methods*. Humana Press; New Jersey: 1992. p. 41-72.
- Sobell LC, Sobell MB, Connors GJ, Agrawal S. Assessing drinking outcomes in alcohol treatment efficacy studies: selecting a yardstick of success. *Alcohol Clin Exp Res*. 2003; 27:1661–1666. [PubMed: 14574238]
- Somoza E, Somoza P, Lewis D, Li SH, Winhusen T, Chiang N, Vocci F, Horn P, Elkashef A. The SRPHK1 outcome measure for cocaine-dependence trials combines self-report, urine benzoylecgonine levels, and the concordance between the two to determine a cocaine-use status for each study day. *Drug Alcohol Depend*. 2008; 93:132–140. [PubMed: 18029115]
- Spratt M, Carpenter J, Sterne JA, Carlin JB, Heron J, Henderson J, Tilling K. Strategies for multiple imputation in longitudinal studies. *Am J Epidemiol*. 2010; 172:478–487. [PubMed: 20616200]
- Tiffany ST, Carter BL, Singleton EG. Challenges in the manipulation, assessment and interpretation of craving relevant variables. *Addiction*. 2000; 95:S177–187. [PubMed: 11002913]
- Tiffany ST, Friedman L, Greenfield SF, Hasin DS, Jackson R. Beyond drug use: a systematic consideration of other outcomes in evaluations of treatments for substance use disorders. *Addiction*. 2012a; 107:709–718. [PubMed: 21981638]
- Tiffany ST, Friedman L, Greenfield SF, Hasin DS, Jackson R. Other outcomes in treatments for substance-use disorders: a call for action. *Addiction*. 2012b; 107:725–726. [PubMed: 22372702]
- Tiffany ST, Wray JM. The clinical significance of drug craving. *Ann NY Acad Sci*. 2012; 1248:1–17. [PubMed: 22172057]
- Velicer WF, Prochaska JO. A comparison of four self-report smoking cessation outcome measures. *Addict Behav*. 2004; 29:51–60. [PubMed: 14667420]
- Vocci FJ. Cognitive remediation in the treatment of stimulant abuse disorders: a research agenda. *Exp Clin Psychopharmacol*. 2008; 16:484–497.
- Volkow ND. What do we know about drug addiction? *Am J Psychol*. 2005; 162:1401–1402.
- Walker R. Retention in treatment--indicator or illusion: an essay. *Subst Use Misuse*. 2009; 44:18–27. [PubMed: 19137480]

- Wells EA, Saxon AJ, Calsyn DA, Jackson TR, Donovan DM. Study results from the Clinical Trials Network's first 10 years: where do they lead? *J Subst Abuse Treat.* 2010; 38(Suppl 1):S14–30. [PubMed: 20307792]
- West R, Hajek P, Stead L, Stapleton J. Outcome criteria in smoking cessation trials: proposal for a common standard. *Addiction.* 2005; 100:299–303. [PubMed: 15733243]
- Winchell C, Rappaport BA, Roca R, Rosebraugh CJ. Reanalysis of methamphetamine dependence treatment trial. *CNS Neurosci Thera.* 2012; 18:367–368.
- Witkiewitz K. “Success” following alcohol treatment: moving beyond abstinence. *Alcohol Clin Exp Res.* 2013; 37:S1, E9–E13.



| Study | Behavioral Treatment Conditions | Medication Treatment Conditions | Design  |          |         | Randomized and follow up study size and sample | Urine collection schedule  | Citation for primary/follow-up trial findings |                              |
|-------|---------------------------------|---------------------------------|---------|----------|---------|--|--|---|------------------------------|
| A     | CBT v TSF v Clinical Management | disulfiram v no medication      | Disulf  | CBT      | TSF     | ClinM  | 121 cocaine and alcohol dependent outpatients, 91 (75%) reach for follow-up              | Weekly  | Carroll et al., 1998; 2000   |
|       |                                 |                                 | No med  |          |         |  |  |   |                              |
| B     | CBT v IPT                       | disulfiram v placebo            | Disulf  | CBT      | IPT     |  | 121 cocaine dependent outpatients, 108 (89%) reached for follow-up                       | Weekly  | Carroll et al., 2004         |
|       |                                 |                                 | Placebo |          |         |  |  |   |                              |
| C     | TSF +TAU v TSF                  | disulfiram v placebo            | Disulf  | TSF+TAU  | TAU     |  | 112 cocaine dependent methadone-maintained individuals, 112 (100%) reached for follow-up | Three times weekly                            | Carroll et al., 2012         |
|       |                                 |                                 | Placebo |          |         |  |  |   |                              |
| D     | CBT4CBT + TAU v TAU             | --                              | TAU     | CBT4CBT  | Control |  | 45 cocaine using outpatients, 38 (84%) reached for follow-up                             | Two times weekly                              | Carroll et al., 2008, 2009   |
|       |                                 |                                 |         |          |         |  |  |   |                              |
| E     | CM + CBT v CBT                  | disulfiram v placebo            | Disulf  | CM + CBT | CBT     |  | 99 cocaine-dependent outpatients, 85 (86% reached for follow-up)                         | Three times weekly                            | Carroll et al., under review |
|       |                                 |                                 | Placebo |          |         |  |  |   |                              |

**Figure 1.**

Overview of designs and samples of randomized clinical trials included in dataset

*Note.* CBT = Cognitive-behavioral therapy. TSF = Twelve-step facilitation. Disulf = disulfiram. ClinM = Clinical management. CM = Contingency management.

**Table 1**

Candidate outcome indicators: Comparison across general features

| Indicator   | Type | Ease of computation  | Verifiability   | Vulnerability to missing data  | Relative cost   | Operationalization for these analyses  |
|---|------|--|---|--|---|--|
| 1<br>Days retained in treatment protocol                                      | C    | Easy   | No  | Low  | Low   | Frequent   |
| 2<br>Percentage of cocaine urine specimens testing negative                   | C    | Easy for complete data   | Yes, by definition  | Assumes independence of urine specimens (denominator), assumes numerator is unbiased by collection schedule or missing data. | High  | Number of cocaine-negative urine specimens collected/all specimens collected             |
| 3<br>Maximum consecutive days abstinent                                       | C    | Easy for complete data   | Yes, provided appropriate schedule of data/urine collection | Likely to result in case-wise missingness or reduced sample size   | Moderate, due to biological verification and derivation from TLFB | Longest continuous cluster of self-reported abstinence within treatment                  |
| 4<br>Percent days of abstinence   | C    | Depends on treatment duration, level of missing data, and intermittent missingness | Yes, provided appropriate schedule of data/urine collection | Likely to result in case-wise missingness or reduced sample size   | Moderate, due to biological verification and derivation from TLFB | Number of self-reported days of abstinence from cocaine/days in treatment (retention)    |
| 5<br>Maximum days of continuous abstinence during last two weeks of treatment | C    | Complex for intermittent and monotone, dropouts                                    | Yes, provided appropriate schedule of data/urine collection | Low  | Moderate, due to biological verification and derivation from TLFB | For those retained 14+ days, longest cluster of abstinence in final 2 weeks; otherwise 0 |
| 6<br>Completely abstinent last two weeks of treatment                         | D    | Easy   | Yes, provided appropriate schedule of data/urine collection | Low  | Moderate, due to biological verification and derivation from TLFB | For those retained 14+ days, 0 days of use in last 14 days, otherwise 0                  |
| 7<br>3 or more weeks of continuous abstinence                                 | D    | Easy   | Yes, provided appropriate schedule of data/urine collection | Low  | Moderate, due to biological verification and derivation from TLFB | “Yes” if participant retained 21+ days, max days abstinent > 20. Otherwise No            |
| 8<br>2 or more weeks of continuous abstinence                                 | D    | Easy   | Yes, provided appropriate schedule of data/urine collection | Low  | Moderate, due to biological verification and derivation from TLFB | “Yes” if participant retained 14+ days, max days abstinent > 13. Otherwise No            |
| 9<br>1 or more weeks of continuous abstinence                                 | D    | Easy   | Yes, provided appropriate schedule of data/urine collection | Low  | Moderate, due to biological verification and derivation from TLFB | “Yes” if participant retained 7+ days, max days abstinent > 6. Otherwise No              |
| 10<br>Completely abstinent from cocaine during treatment                      | D    | Easy   | Same  | Low  | Moderate, due to biological verification                          | 0 days of use and 0 positive urines  |

| Indicator | Type | Ease of computation                           | Verifiability  | Vulnerability to missing data | Relative cost                                     | Operationalization for these analyses  |
|-----------|------|---|--|-------------------------------|---|--|
|           |      |   |  |                               | and derivation from TLFB and derivation from TLFB |  |
| 11        | D    | Easy  | Yes  | Low                           | Low   | Completion of treatment, 0 days of use in final week   |
| 12        | C    | Complex, baseline definition can be arbitrary | No, relies on accurate baseline/pre-treatment assessment         | Moderate                      | Low   | Percent days of use in final 28 days of treatment/percent days of use in 28 days prior to baseline |
| 13        | D    | Complex, baseline definition can be arbitrary | Relies on access to accurate baseline/pre-treatment level of use | Moderate                      | Low   | % reduction is 50% or higher   |
| 14        | D    | Complex, baseline definition can be arbitrary | Same   | Moderate                      | Low   | % reduction is 75% or higher   |
| 15        | D    | Easy  | Partial  | Low                           | Low   | Completes treatment, 0 days of problems in drug, legal, employment and psych ASI in past 28 days   |

Note. C=continuous, D=Dichotomous, TLFB=Timeline Followback method

**Table 2**  
Candidate indicators (within treatment outcomes) and follow up outcomes across studies, N=424

| Outcome indicator   | Type | Study A<br>n=91 |         | Study B<br>n=108 |         | Study C<br>n=112 |         | Study D<br>n=38 |         | Study E<br>n=85 |         | Total<br>n=434 |         |
|---|------|-----------------|---------|------------------|---------|------------------|---------|-----------------|---------|-----------------|---------|----------------|---------|
|   |      | N/mean          | Prop/SD | N/mean           | Prop/SD | N/mean           | Prop/SD | N/mean          | Prop/SD | N/mean          | Prop/SD | N/mean         | Prop/SD |
| 1 Days retained in treatment protocol, mean and SD <sup>2</sup>                                 | C    | 50.2            | 31.6    | 53.5             | 34.2    | 69.6             | 26.9    | 40.8            | 17.8    | 41.4            | 35.1    | 53.5           | 32.6    |
| 2 Percent cocaine negative urine specimens <sup>3</sup>   | C    | 26.3            | 30.2    | 53.7             | 37.9    | 23.3             | 26.5    | 51.6            | 42.3    | 44.3            | 40.0    | 39.0           | 37.1    |
| 3 Maximum consecutive days abstinent  | C    | 27.0            | 24.7    | 28.1             | 27.7    | 17.3             | 19.2    | 22.0            | 18.5    | 18.4            | 25.3    | 22.6           | 24.2    |
| 4 Percent days of abstinence  | C    | 85.5            | 15.4    | 80.4             | 24.9    | 60.3             | 25.7    | 82.4            | 19.3    | 76.6            | 26.3    | 75.6           | 25.1    |
| 5 Max. days continuous abstinence during participant's last two weeks of treatment <sup>4</sup> | C    |                 |         | 9.1              | 4.2     | 7.2              | 4.7     | 10.5            | 4.2     | 8.8             | 5.0     | 8.6            | 4.6     |
| 6 Number, proportion completely abstinent last 2 weeks of treatment                             | D    | 37              | .46     | 33               | .39     | 21               | .19     | 20              | .53     | 23              | .42     | 134            | .36     |
| 7 Number, proportion attaining 3+ weeks of continuous abstinence                                | D    | 37              | .41     | 45               | .42     | 26               | .23     | 17              | .45     | 55              | .65     | 180            | .42     |
| 8 Number, proportion attaining 2+ weeks of continuous abstinence                                | D    | 51              | .56     | 52               | .48     | 35               | .31     | 20              | .53     | 58              | .68     | 216            | .50     |
| 9 Number, proportion, attaining 1+ weeks of continuous abstinence                               | D    | 68              | .75     | 78               | .72     | 71               | .63     | 24              | .63     | 69              | .81     | 310            | .71     |
| 10 Number, proportion of participants completely abstinent from cocaine during treatment        | D    | 17              | .19     | 12               | .11     | 5                | .05     | 9               | .24     | 17              | .24     | 60             | .14     |

| Outcome indicator  | Type I | Study A<br>n=91 |      | Study B<br>n=108 |      | Study C<br>n=112 |         | Study D<br>n=38 |      | Study E<br>n=85 |      | Total<br>n=434 |         |     |      |         |
|--|--------|-----------------|------|------------------|------|------------------|---------|-----------------|------|-----------------|------|----------------|---------|-----|------|---------|
|  |        | N               | mean | Prop/SD          | N    | mean             | Prop/SD | N               | mean | Prop/SD         | N    | mean           | Prop/SD | N   | mean | Prop/SD |
| 11 Completed treatment and abstinent in last week  | D      | 27              | .30  | .37              | 37   | .37              | 33      | .31             | 12   | .55             | 19   | .22            | 128     | .32 |      |         |
| 12 Percent reduction in frequency of cocaine use (28 days prior/ days last 4 weeks)                                    | C      | 70.1            | 32.5 | 37.8             | 63.1 | 48.2             | 39.2    | 59.5            | 38.4 | 50.9            | 36.7 | 57.9           | 37.9    |     |      |         |
| 13 Number, proportion attaining 50% reduction in frequency of cocaine use  | D      | 28              | .31  | .39              | 42   | .35              | .31     | 6               | .16  | 28              | .33  | 139            | .32     |     |      |         |
| 14 Number, proportion attaining 75% reduction in frequency of cocaine use  | D      | 14              | .15  | .22              | 24   | .10              | .09     | 1               | .03  | 20              | .24  | 69             | .16     |     |      |         |
| 15 Number, proportion reporting no cocaine use, legal, employment, or psychological problems last 28 days of treatment | D      | 7               | .18  | .00              | 0    | .11              | .10     | 3               | .08  | 21              | .25  | 42             | .11     |     |      |         |
| Follow-up  |        |                 |      |                  |      |                  |         |                 |      |                 |      |                |         |     |      |         |
| Days of cocaine use, month 1 follow up   | C      | 7.6             | 8.5  | 7.5              | 5.5  | 5.3              | 7.4     | 3.0             | 5.6  | 4.2             | 6.8  | 5.4            | 7.5     |     |      |         |
| Days of cocaine use, month 3 follow up   | C      | 7.6             | 8.5  | 6.9              | 4.8  | 6.0              | 9.0     | 2.5             | 5.3  | 2.8             | 6.2  | 5.1            | 7.8     |     |      |         |
| Days of cocaine use, month 6 follow-up <sup>5</sup>  | C      | 6.3             | 8.0  | 8.0              | 5.8  | 5.2              | 7.7     | 3.9             | 6.5  | 3.6             | 7.5  | 5.2            | 7.7     |     |      |         |
| Days of cocaine use, month 12 follow up  | C      | 6.9             | 8.7  | 6.9              | 3.8  | 4.7              | 6.9     | NA              |      | 2.8             | 5.9  | 4.5            | 7.2     |     |      |         |
| Completely abstinent during full follow up   | D      | 10              | .11  | .07              | 8    | .07              | .07     | 12              | .32  | 16              | .19  | 54             | .13     |     |      |         |
| 6-Number, proportion meeting criteria for 'good functioning' at month 1  | D      | 12              | .14  | .13              | 14   | .18              | .17     | 5               | .13  | 22              | .26  | 71             | .17     |     |      |         |
| Number, proportion meeting criteria for  | D      | 10              | .11  | .18              | 19   | .14              | .13     | 5               | .14  | 25              | .29  | 73             | .17     |     |      |         |

| Outcome indicator  | Type <sup>1</sup> | Study A<br>n=91 |         | Study B<br>n=108 |         | Study C<br>n=112 |         | Study D<br>n=38 |         | Study E<br>n=85 |         | Total<br>n=434 |         |
|--|-------------------|-----------------|---------|------------------|---------|------------------|---------|-----------------|---------|-----------------|---------|----------------|---------|
|  |                   | N/mean          | Prop/SD | N/mean           | Prop/SD | N/mean           | Prop/SD | N/mean          | Prop/SD | N/mean          | Prop/SD | N/mean         | Prop/SD |
| 'good functioning' at month 3  |                   |                 |         |                  |         |                  |         |                 |         |                 |         |                |         |
| 'good functioning' at month 3  |                   |                 |         |                  |         |                  |         |                 |         |                 |         |                |         |
| Number, proportion meeting criteria for 'good functioning' at month 6  | D                 | 11              | .13     | 13               | .13     | 18               | .16     | 8               | .24     | 22              | .27     | 72             | .17     |
| Number, proportion meeting criteria for 'good functioning' at month 12 | D                 | 8               | .11     | 14               | .14     | 22               | .21     | NA              |         | 31              | .39     | 75             | .21     |

Note.

- <sup>1</sup> Indicates whether variable is continuous (C) or dichotomous (D). For continuous variables, values are presented as mean, SD. For dichotomous variables, values presents are number and proportion
- <sup>2</sup> Maximum days of retention is 84 (12 weeks), with the exception of Study D which was 56 days (8 weeks)
- <sup>3</sup> Operationalization of outcome indicators 1–15 is described in Table 1
- <sup>4</sup> Could not be calculated for Study A due to data availability
- <sup>5</sup> Studies A, B, C and E conducted follow-up interviews at 1, 3, 6, and 12 months after treatment. Study D conducted follow-ups at 1, 3, and 6 months.
- <sup>6</sup> Indicates participant reported 0 days of cocaine use and 0 days of problems in past 28 days for psychological, family, legal, and employment domains on ASI.

**Table 3**

Sensitivity to medication effects (disulfiram versus control)

| Outcome indicator  | Type | No medication<br>n=36 |         | Placebo<br>n=148 |         | Disulfiram<br>n=212 |         | F/X <sup>2</sup> | p   | Tukey/phi                |
|--|------|-----------------------|---------|------------------|---------|---------------------|---------|------------------|-----|--------------------------|
|  |      | mean or N             | sd or % | mean or N        | sd or % | mean or N           | sd or % |                  |     |                          |
| 1 Days retained in treatment protocol  | C    | 37.5                  | 32.1    | 53.7             | 35.6    | 58.3                | 32.6    | 6.26             | .00 | Disulf, Placebo > No med |
| 2 Percent cocaine negative urine specimens   | C    | 42.6                  | 41.1    | 37.5             | 36.0    | 38.9                | 36.8    | 0.37             | .70 |                          |
| 3 Maximum consecutive days abstinent   | C    | 19.5                  | 18.6    | 19.1             | 23.4    | 25.7                | 26.2    | 3.2              | .04 | Disulf > Placebo         |
| 4 Percent days of abstinence from cocaine Maximum days of continuous abstinence                  | C    | 83.1                  | 18.0    | 76.4             | 25.3    | 70.9                | 26.9    | 5.61             | .00 | No med > Placebo         |
| 5 during last two weeks of treatment*  | C    |                       |         | 7.6              | 4.5     | 8.9                 | 4.7     | 4.83             | .03 | Disulfiram > Placebo     |
| 6 Completely abstinent last two weeks of treatment   | D    | 10                    | 36%     | 29               | 24%     | 75                  | 24%     | 10.39            | .01 | No med > Disulf, placebo |
| 7 3 or more weeks of continuous abstinence   | D    | 9                     | 25%     | 61               | 41%     | 93                  | 44%     | 4.52             | .10 |                          |
| 8 2 or more weeks of continuous abstinence   | D    | 15                    | 42%     | 68               | 46%     | 113                 | 53%     | 2.86             | .24 |                          |
| 9 1 or more weeks of continuous abstinence   | D    | 22                    | 61%     | 106              | 72%     | 158                 | 75%     | 2.84             | .26 |                          |
| 10 Completely abstinent from cocaine during treatment  | D    | 5                     | 14%     | 17               | 10%     | 29                  | 14%     | 0.26             | .88 |                          |
| 11 Completed treatment and abstinent in last week  | D    | 7                     | 19%     | 35               | 24%     | 74                  | 37%     | 8.27             | .02 | Disulf > No med          |
| 12 Percent reduction (28 days prior/days last 4 weeks)   | C    | 62.2                  | 36.0    | 60.2             | 37.3    | 52.8                | 39.3    | 1.78             | .17 |                          |
| 13 50% reduction in frequency of cocaine use   | D    | 9                     | 25%     | 0.54             | 37%     | 70                  | 33%     | 1.78             | .41 |                          |
| 14 75% reduction in cocaine use  | D    | 4                     | 11%     | 28               | 19%     | 36                  | 17%     | 1.25             | .54 |                          |
| 15 Report no cocaine use, legal, employment, or psychological problems last 28 days of treatment | D    | 2                     | 17%     | 19               | 13%     | 18                  | 19%     | 1.18             | .55 |                          |

Note. N=396, sample represents participants from Studies A, B, C and E only, as Study D did not include medication condition

**Table 4**

Sensitivity of candidate indicators to effects of behavioral therapies, N=434

| Outcome indicator   | Type | Comparison <sup>1</sup> (1)<br>n=195 |         | TSF <sup>2</sup> (2)<br>n=97 |         | CBT (3)<br>n=105 |         | CM (4)<br>n=37 |         | F, T, x <sup>2</sup> | p    | Tukey                |
|---|------|--------------------------------------|---------|------------------------------|---------|------------------|---------|----------------|---------|----------------------|------|----------------------|
|   |      | mean or n                            | sd or % | mean or n                    | sd or % | mean or n        | sd or % | Mean or n      | sd or % |                      |      |                      |
| 1 Days retained in treatment protocol <sup>2</sup>  | C    | 54.2                                 | 33.1    | 56.9                         | 33      | 51.63            | 30.86   | 46.08          | 34.05   | 1.13                 | .34  | ns                   |
| 2 Proportion cocaine negative urine specimens <sup>3</sup>  | C    | .36                                  | .37     | .26                          | .29     | .49              | .38     | .56            | .38     | 8.13                 | 0.00 | 1v3, 1v4<br>2v3, 2v4 |
| 3 Maximum consecutive days abstinent, mean and SD   | C    | 20.1                                 | 24.3    | 23.7                         | 22.97   | 25.7             | 23.58   | 23.76          | 27.72   | 1.26                 | .29  | ns                   |
| 4 Percent days of abstinence  | C    | .72                                  | .27     | .75                          | .24     | .82              | .21     | .81            | .26     | 4.20                 | 0.01 | 1v3                  |
| 5 Maximum days of continuous abstinence during participant's last two weeks of treatment <sup>4</sup>                   | C    | 8.1                                  | 4.7     | 7.9                          | 4.53    | 9.35             | 4.38    | 10.27          | 4.63    | 2.54                 | 0.06 | ns                   |
| 6 Number, percent completely abstinent last 2 weeks of treatment  | D    | 54                                   | 33%     | 24                           | 27%     | 42               | 46%     | 14.00          | 54%     | 11.61                | .009 | 4 v other, 3 v 2     |
| 7 Number, percent attaining 3+ weeks of continuous abstinence   | D    | 74                                   | 38%     | 34                           | 35%     | 45               | 43%     | 27.00          | 73%     | 17.85                | .00  | 4 v other            |
| 8 Number, percent attaining 2+ weeks of continuous abstinence   | D    | 90                                   | 46%     | 40                           | 41%     | 57               | 54%     | 29.00          | 78%     | 16.82                | .001 | 4 v other            |
| 9 Number, percent, attaining 1+ weeks of continuous abstinence  | D    | 139                                  | 71%     | 66                           | 68%     | 73               | 70%     | 32.00          | 71%     | 4.85                 | .18  | ns                   |
| 10 Number, percent of participants completely abstinent from cocaine during treatment                                   | D    | 25                                   | 13%     | 11                           | 11%     | 16               | 15%     | 8.00           | 24%     | 3.55                 | .31  | ns                   |
| 11 Number, percent completing treatment and abstinent in last week  | D    | 56                                   | 31%     | 26                           | 28%     | 35               | 37%     | 11.00          | 30%     | 2.15                 | .54  | ns                   |
| 12 Percent reduction in frequency of cocaine use (28 days prior/days last 4 weeks)                                      | C    | 0.53                                 | 0.38    | 0.58                         | .39     | .65              | .37     | .62            | .32     | 2.15                 | .09  | ns                   |
| 13 Number, percent attaining 50% reduction in frequency   | D    | 60.0                                 | 31%     | 26                           | 27%     | 38               | 36%     | 15.00          | 41%     | 3.43                 | .33  | ns                   |
| 14 Number, percent attaining 75% reduction in cocaine use   | D    | 23                                   | 12%     | 9                            | 9%      | 25               | 24%     | 12.00          | 32%     | 18.12                | .00  | 3 v other, 4 v other |
| 15 Number, percent reporting no cocaine use/no days of legal, employment, or psychological problems on ASI last 28 days | D    | 20                                   | 11%     | 8                            | 12%     | 5                | 6%      | 9.00           | 24%     | 8.9                  | .03  | 4 v other            |

Note.

<sup>1</sup> Indicates comparison condition for behavioral therapy for each study. Study A, Clinical Management; Study B, Interpersonal Therapy (IPT); Study C, Treatment as usual (TAU), Study D, TAU, Study E, CBT.



<sup>2</sup> TST=Twelve Step Facilitation, CBT=Cognitive Behavioral Therapy, CM=Prize Contingency Management

**Table 5**  
Relationship of candidate indicators to cocaine use and functioning during follow-up, N=434

| Outcome indicator  | Days of cocaine Use Month 1 | Days of cocaine Use Month 3 | Days of cocaine Use Month 6 | Days of cocaine Use Month 12 | Totally abstinent throughout follow-up | Good functioning at Month 1 | Good functioning at Month 3 | Good functioning at Month 6 | Good functioning at Month 12 |
|--|-----------------------------|-----------------------------|-----------------------------|------------------------------|--|-----------------------------|-----------------------------|-----------------------------|------------------------------|
| 1 Days retained in treatment protocol  | r<br>-0.12                  | -0.06                       | -0.08                       | .02                          | .05                                    | .10                         | .05                         | .08                         | .04                          |
|  | p<br>.01                    | .23                         | .09                         | .76                          | .33                                    | .04                         | .35                         | .12                         | .47                          |
| 2 Percent cocaine negative urine specimens   | r<br>-0.31                  | -0.28                       | -0.30                       | -0.16                        | .33                                    | .33                         | .29                         | .25                         | .22                          |
|  | p<br>.00                    | .00                         | .00                         | .01                          | .00                                    | .00                         | .00                         | .00                         | .00                          |
| 3 Maximum consecutive days of abstinence   | r<br>-0.30                  | -0.24                       | -0.26                       | -0.12                        | .30                                    | .34                         | .26                         | .24                         | .17                          |
|  | p<br>.00                    | .00                         | .00                         | .02                          | .00                                    | .00                         | .00                         | .00                         | .00                          |
| 4 Percent days of abstinence   | r<br>-0.39                  | -0.37                       | -0.35                       | -0.24                        | .19                                    | .23                         | .21                         | .18                         | .14                          |
|  | p<br>.00                    | .00                         | .00                         | .00                          | .00                                    | .00                         | .00                         | .00                         | .01                          |
| 5 Maximum days of consecutive abstinence during participants last two weeks of treatment | r<br>-0.46                  | -0.35                       | -0.30                       | -0.21                        | .32                                    | .31                         | .33                         | .19                         | .24                          |
|  | p<br>.00                    | .00                         | .00                         | .00                          | .00                                    | .00                         | .00                         | .00                         | .00                          |
| 6 Number and percent completely abstinent last two weeks of treatment                    | r<br>-0.30                  | -0.25                       | -0.19                       | -0.07                        | .28                                    | .29                         | .31                         | .21                         | .19                          |
|  | p<br>.00                    | .00                         | .00                         | .25                          | .00                                    | .00                         | .00                         | .00                         | .00                          |
| 7 Percent attaining 3+ weeks of abstinence   | r<br>-0.33                  | -0.30                       | -0.28                       | -0.16                        | .25                                    | .26                         | .26                         | .24                         | .24                          |
|  | p<br>.00                    | .00                         | .00                         | .00                          | .00                                    | .00                         | .00                         | .00                         | .00                          |
| 8 Percent attaining 2+ weeks of abstinence   | r<br>-0.26                  | -0.26                       | -0.28                       | -0.14                        | .24                                    | .24                         | .26                         | .22                         | .20                          |
|  | p<br>.00                    | .00                         | .00                         | .01                          | .00                                    | .00                         | .00                         | .00                         | .00                          |
| 9 Percent attaining 1+ week of abstinence  | r<br>-0.27                  | -0.22                       | -0.24                       | -0.10                        | .11                                    | .21                         | .17                         | .15                         | .17                          |
|  | p<br>.00                    | .00                         | .00                         | .05                          | .02                                    | .00                         | .00                         | .00                         | .00                          |
| 10 Percent completely abstinent during treatment   | r<br>-0.14                  | -0.08                       | -0.11                       | -0.09                        | .23                                    | .28                         | .17                         | .14                         | .19                          |
|  | p<br>.00                    | .12                         | .03                         | .09                          | .00                                    | .00                         | .00                         | .00                         | .00                          |
| 11 Completed treatment and abstinent in the last week                                    | r<br>-0.30                  | -0.24                       | -0.22                       | -0.09                        | .23                                    | .22                         | .23                         | .19                         | .12                          |
|  | p<br>.00                    | .00                         | .00                         | .09                          | .00                                    | .00                         | .00                         | .00                         | .03                          |
| 12 Percent reduction in frequency of cocaine use   | r<br>-0.32                  | -0.26                       | -0.22                       | -0.11                        | .18                                    | .24                         | .22                         | .17                         | .14                          |
|  | p<br>.00                    | .00                         | .00                         | .07                          | .00                                    | .00                         | .00                         | .00                         | .01                          |
| 13 Percent attaining 50% reduction   | r<br>-0.02                  | -0.01                       | .02                         | .04                          | -0.16                                  | -0.04                       | -0.05                       | -0.09                       | .00                          |
|  | p<br>.65                    | .76                         | .67                         | .49                          | .00                                    | .40                         | .30                         | .06                         | .94                          |

| Outcome indicator  | Days of cocaine Use Month 1 | Days of cocaine Use Month 3 | Days of cocaine Use Month 6 | Days of cocaine Use Month 12 | Totally abstinent throughout follow-up | Good functioning at Month 1 | Good functioning at Month 3 | Good functioning at Month 6 | Good functioning at Month 12 |
|--|-----------------------------|-----------------------------|-----------------------------|------------------------------|--|-----------------------------|-----------------------------|-----------------------------|------------------------------|
| 14 Percent attaining 75% reduction   | r<br>-.08                   | -.07                        | -.04                        | -.01                         | -.11                                   | -.04                        | .00                         | -.01                        | .07                          |
|  | p<br>.08                    | .14                         | .42                         | .92                          | .02                                    | .39                         | .94                         | .88                         | .21                          |
| 15 Percent abstinent and reporting no days of problems in ASI legal, employment, and psychological domains | r<br>-.20                   | -.20                        | -.15                        | -.08                         | .29                                    | .37                         | .28                         | .24                         | .21                          |
|  | p<br>.00                    | .00                         | .00                         | .16                          | .00                                    | .00                         | .00                         | .00                         | .00                          |