Published in final edited form as:

Drug Alcohol Depend. 2016 January 1; 158: 1–7. doi:10.1016/j.drugalcdep.2015.11.004.

Measures of Outcome for Stimulant Trials: ACTTION Recommendations and Research Agenda

Brian D. Kiluk[a,*], **Kathleen M. Carroll**[a], **Amy Duhig**[b], **Daniel E. Falk**[c], **Kyle Kampman**[d], **Shengan Lai**[e], **Raye Z. Litten**[c], **David J. McCann**[f], **Ivan D. Montoya**[f], **Kenzie L. Preston**[f], **Phil Skolnick**[f], **Constance Weisner**[g], **George Woody**[h], **Redonna Chandler**[f], **Michael J. Detke**[i], **Kelly Dunn**[e], **Robert H. Dworkin**[j], **Joanne Fertig**[c], **Jennifer Gewandter**[j], **F. Gerard Moeller**[k], **Tatiana Ramey**[f], **Megan Ryan**[c], **Kenneth Silverman**[e], and **Eric C. Strain**[e]

[a]Yale University School of Medicine, West Haven, CT 06516, USA

[b]Xcenda, Palm Harbor, Florida, USA

[c]National Institute on Alcohol Abuse and Alcoholism (NIAAA), Bethesda, MD, USA

[d]Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

[e]Johns Hopkins School of Medicine, Baltimore, MD, USA

[f]National Institute on Drug Abuse (NIDA), Bethesda, MD, USA

[g]Division of Research, Kaiser Permanente, University of California – San Francisco, San Francisco, CA, USA

[h]University of Pennsylvania and Treatment Research Institute, Philadelphia, PA, USA

[i]Embera Neurotherapeutics, Indiana University School of Medicine, Carmel, IN, USA

[j]University of Rochester School of Medicine and Dentistry, Rochester, NY, USA

[*]Corresponding author at: Department of Psychiatry, Yale University School of Medicine, 950 Campbell Ave, 151D, West Haven, CT 06516, USA; Tel. (203) 937-3486; brian.kiluk@yale.edu.

Contributors

All authors participated in the ACTTION-sponsored meeting, which provided the content for this manuscript. Author Kiluk wrote the initial draft of the manuscript. All authors contributed to and have approved of the final manuscript.

Conflict of Interest

The views expressed in this article are those of the authors, none of whom have financial conflicts of interest specifically related to the issues discussed in this article. At the time of the meeting on which this article is based, several authors were employed by pharmaceutical companies and others had received consulting fees or honoraria from one or more pharmaceutical or device companies. Authors of this article who were not employed by industry or government at the time of the meeting received travel stipends, hotel accommodations, and meals during the meeting provided by the Analgesic, Anesthetic, and Addiction Clinical Trial Translations, Innovations, Opportunities, and Networks (ACTTION) public-private partnership with the US Food and Drug Administration (FDA), which has received research contracts, grants, or other revenue from the FDA, multiple pharmaceutical and device companies, and other sources. Preparation of background literature reviews and draft manuscripts was supported by ACTTION. No official endorsement by the FDA, US National Institutes of Health, or the pharmaceutical and device companies that have provided unrestricted grants to support the activities of ACTTION should be inferred.

Dr. Strain's participation in this activity was as a consultant to ACTTION. All opinions expressed and implied in this activity are solely those of Dr. Strain and do not represent or reflect the views of the Johns Hopkins University or the Johns Hopkins Health System.

[k]Virginia Commonwealth University, Richmond, VA, USA

Abstract

Background—The development and approval of an efficacious pharmacotherapy for stimulant use disorders has been limited by the lack of a meaningful indicator of treatment success, other than sustained abstinence.

Methods—In March, 2015, a meeting sponsored by Analgesic, Anesthetic, and Addiction Clinical Trial Translations, Innovations, Opportunities, and Networks (ACTTION) was convened to discuss the current state of the evidence regarding meaningful outcome measures in clinical trials for stimulant use disorders. Attendees included members of academia, funding and regulatory agencies, pharmaceutical companies, and healthcare organizations. The goal was to establish a research agenda for the development of a meaningful outcome measure that may be used as an endpoint in clinical trials for stimulant use disorders.

Results and Conclusions—Based on guidelines for the selection of clinical trial endpoints, the lessons learned from prior addiction clinical trials, and the process that led to identification of a meaningful indicator of treatment success for alcohol use disorders, several recommendations for future research were generated. These include a focus on the validation of patient reported outcome measures of functioning, the exploration of patterns of stimulant abstinence that may be associated with physical and/or psychosocial benefits, the role of urine testing for validating self-reported measures of stimulant abstinence, and the operational definitions for reduction-based measures in terms of frequency rather than quantity of stimulant use. These recommendations may be useful for secondary analyses of clinical trial data, and in the design of future clinical trials that may help establish a meaningful indicator of treatment success.

Keywords

Stimulant use disorders; Outcome measures; Clinical trials

1. INTRODUCTION

Sustained abstinence is considered the only outcome currently accepted by the US Food and Drug Administration (FDA) as a valid endpoint for clinical trials evaluating pharmacotherapies for drug use disorders (FDA: Psychopharmacologic Drugs Advisory Committee, 2013; Winchell et al., 2012). However, this endpoint is often considered unrealistic, and the lack of meaningful alternative indicators of treatment success (Carroll et al., 2014; Donovan et al., 2012) may be one factor that has hindered the development and approval of an efficacious pharmacotherapy for stimulant use disorders (see Acri and Skolnick, 2013). On March 24th and 25th, 2015, a meeting sponsored by the Analgesic, Anesthetic, and Addiction Clinical Trial Translations, Innovations, Opportunities, and Networks (ACTTION), a public-private partnership with the FDA, was convened to discuss ‘Measures of Outcome for Stimulant Trials’. ACTTION’s mission includes optimizing the design and execution of clinical trials to expedite the discovery and development of improved treatments. Participants were drawn from clinical investigators, representatives of the National Institute on Drug Abuse (NIDA), the National Institute on Alcohol Abuse and Alcoholism (NIAAA), the FDA, pharmaceutical companies, and healthcare organizations.

The overall goal was to identify a research agenda for the development of outcome measures other than sustained abstinence that would be clinically meaningful and could be used as endpoints for stimulant use disorder clinical trials. The purpose of this review is to provide a summary of the state of knowledge regarding this topic area, as addressed at this meeting, and to make recommendations for the field moving forward.

2. CHARACTERISTICS OF A MEANINGFUL OUTCOME MEASURE

The FDA's Center for Drug Evaluation and Research (CDER) provides specific guidance to the research and pharmaceutical communities regarding the selection of endpoints for use in clinical trials. CDER has a formal for identifying specific measures that will aid in drug development, which include biomarkers and clinical outcome assessments (for more detailed information, see Qualification Process for Drug Development Tools, Center for Drug Evaluation and Research, 2014). In addition to the need for any assessment tool to have strong psychometric properties (e.g., reliability, validity), several aspects of the outcome measure should be considered for selection as an endpoint in stimulant trials and are discussed below.

Clinical outcome assessments are those that measure a patient's symptoms or level of functioning, and can provide both direct and indirect evidence of treatment response (depending on who is reporting the outcome: patient vs. clinician vs. observer). Of the various potential clinical outcome assessments possible, the FDA views patient-reported outcomes as the nearest to direct evidence for some conditions, as they come directly from the patient without interpretation from others. These are formally recommended "when measuring a concept best known by the patient or best measured from the patient perspective" (from FDA Guidance for Industry document: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims). Therefore, a patient-reported outcome may provide meaningful evidence of benefit from treatment for stimulant use, as the disorder is characterized by a wide variety of problems potentially better measured from the patient's perspective in some cases (more than mere frequency of drug consumption). However, in the treatment of stimulant use disorders, there is some disagreement regarding the validity of patient-reported drug use, drug-related symptoms and problems (Hjorthøj et al., 2012; Magura and Kang, 1996).

While most treatments (pharmacotherapy or behavioral) are designed to affect the target behavior of stimulant use, measuring rates of stimulant use may not be the sole indicator of treatment success. Treatment benefit is demonstrated by evidence of a positive impact on how an individual feels or functions in daily life; a meaningful outcome measure should be capable of indicating change in one of these areas. Although changes in biomarkers such as urine test results may be useful as an objective indicator of response to a therapeutic intervention, they are considered a surrogate (i.e., substitute) for how an individual feels or functions in their daily life, and may not be a particularly meaningful outcome of treatment for drug use disorders that are characterized by multiple physical and psychosocial problems/consequences (e.g., Winchell et al., 2012). Due to the chronic nature of stimulant use disorders, demonstrating significant change in physical and psychosocial domains is limited by the relatively short duration of most clinical trials. Therefore, a meaningful

outcome measure would be a level of reduced drug use that is predictive of long-term improvement in an individual's functioning in these areas. Several clinical trials have documented a statistically significant reduction in urine measures of stimulant use; however, identification of the specific level of reduced stimulant use (in terms of duration of abstinence and/or reduction in frequency of use) that is associated with clinically meaningful indices of long-term improvement has not been established.

3. CHALLENGES IN MEASURING REDUCTIONS IN STIMULANT USE

The existence of a valid, biological indicator for detecting stimulant use (i.e., urine testing) is a major advantage compared to other psychiatric disorders, yet also has important limitations as an outcome measure. In general, detection times for stimulant metabolites in urine are up to 2–3 days after the occurrence of drug use (Cone et al., 2003; Oyler et al., 2002; Preston et al., 2002), yet many additional factors result in substantial variability in the ability to detect urine metabolites (e.g., route of administration, dose/purity of drug, individual differences in drug metabolism, urine concentration, level of drug use chronicity). These factors often create wide variations in metabolite concentrations in urine (e.g., concentrations of benzoylecgonine, a cocaine metabolite, can be detected at 150ng/mL but concentrations greater than 900,000 ng/mL have been reported in some trials; Preston et al., 1998, 1997). While using urine drug screen results (or other biological indicators) as a primary efficacy endpoint in clinical trials have multiple advantages, strict reliance upon urinalysis results for evidence of treatment outcome is not as clear-cut as it may seem. Differences in the frequency of sample collection, the type of biological analysis (quantitative or qualitative), the threshold for determining abstinence (e.g., 150 vs. 300 ng/ml of benzoylecgonine), the resolution of discrepancy between biological and self-report data, and the handling of missing data all have a profound impact on the validity of the outcome measure.

The lack of standard methods for the field regarding these and other issues necessitate multiple decisions that influence the interpretation of drug use outcome measures. First, the most suitable schedule of urine sample collection to detect stimulant use in clinical trials is unresolved. Due to the relatively brief window of detection, urine samples collected only once per week may result in an underestimation of stimulant use, yet collection of urines too frequently (e.g., 3 times per week) can result in overestimation of stimulant use due to carryover (Preston et al., 1997). The level of participant burden should also be considered when selecting a urine collection schedule. Second, there is no consensus regarding appropriate measurement of self-reported stimulant use. Many researchers create dichotomous variables to reflect daily stimulant use (yes/no), yet the lack of standardization across drugs, routes of administration, and the methodology used to obtain self-reports limits the ability to detect reductions in the quantity of stimulant use. Third, while there is general agreement that both self-report and biological results (e.g., urine drug screens) should be included in clinical trials with stimulant users (Donovan et al., 2012), combining these sources of data is not at all straightforward. The challenge arises when these two sources are incongruent or when one source is missing, particularly in instances when the outcome measure is a continuous variable (e.g., percentage of days abstinent). Several methods have been proposed for resolving the discrepancy through use of algorithms (Oden et al., 2011;

Preston et al., 1997; Somoza et al., 2008), but these can be complicated, often requiring costly quantitative urinalysis testing, and entail some level of assumptions. This issue is less complicated when using a dichotomous outcome measure (e.g., abstinent for at least 4 weeks), as the presence of a positive urine during the time period would indicate a failure to reach that cutoff (Carroll et al., 2014). Finally, regardless of the type of outcome measure considered (dichotomous or continuous; self-report or biological), missing data is arguably the biggest challenge to defining drug use/abstinence outcomes.

Although missing data are problematic in any clinical trial (Lavori et al., 2008; Siddique et al., 2008), the assumptions for handling missing data in drug use trials can lead to very different estimates of use and often determine whether an outcome measure detects a treatment effect or not (McPherson et al., 2015; Witkiewitz et al., 2014). For instance, rates of positive/negative urine drug screen results can vary widely depending on whether missing urine results are treated as 'positive' or 'missing' (e.g., many trials impute missing as indicative of 'positive' urine result). This is also true for calculation of continuous indicators of self-reported cocaine use, such as the percentage of days abstinent, wherein the selected denominator is a crucial decision for interpreting results (i.e., whether the denominator is the total number of urine samples expected or provided during the course of the trial; Carroll et al., 2014). A 2010 report produced by a National Academy of Sciences Panel on the Handling of Missing Data in Clinical Trials provides 18 recommendations for addressing missing data, noting that the preferred approach to the problem of missing data in clinical trials is to avoid missing data in the first place, as all strategies for statistical imputation or correction have important limitations including unverifiable assumptions and some level of subjectivity (National Research Council, 2010).

An additional challenge in measuring stimulant use as an outcome is in defining meaningful change in quantity/frequency of use (i.e., reduction). Because stimulants, such as cocaine and methamphetamine, are illicit drugs, there is no normative or established 'safe' level of use. Given the potential for serious harm due to the acute ingestion of cocaine or methamphetamine, it has been suggested that any reduction in the frequency an individual engages in this behavior would appear to benefit the individual (McCann et al., 2015). Others have argued that reduced use *per se* provides insufficient evidence of clinical benefit (Winchell et al., 2012). Furthermore, there are no standard units for quantifying illicit drugs (as there are with alcohol or tobacco), so calculating an outcome measure based on a reduction in self-reported quantity becomes virtually impossible. Thus, any definition of meaningful change would have to be defined by changes in the frequency (e.g., in days, weeks) rather than the amount, which limits the sensitivity of the measure (Carroll et al., 2014).

4. USING THE ALCOHOL FIELD AS A GUIDE

As there are similarities in the behavioral patterns of drug-taking for stimulants and alcohol (i.e., variable patterns that include periods of abstinence and relapse), the process that led to the identification of a standard, accepted and valid outcome measure of treatment success for alcohol use, 'percent of subjects with no heavy drinking days' (PSNHDDs), may serve as a useful model for the stimulant outcomes development process. Alcohol

pharmacotherapy clinical trials had a similar history as stimulant trials in that numerous outcomes were evaluated as efficacy endpoints, largely based on self-reported quantity and frequency of drinking rather than psychosocial or physical consequences of alcohol use (Falk et al., 2010). Continuous variables, such as the percentage of days abstinent, the percentage of heavy drinking days, or total alcohol consumption were the preferred measures, with outcomes typically presented as a comparison of group means to determine statistical significance. Continuous outcomes like these are desirable since they capture changes in the amount and pattern of drinking, and are generally known to have greater statistical power to detect a treatment effect than dichotomous outcomes. Yet despite the advantages of such continuous outcome measures, they are difficult to interpret in terms of clinical benefit or quantifying how many individuals achieved a 'good outcome'. On the other hand, dichotomous outcomes are more readily interpretable, have been promoted to increase acceptance of medication effects, and allow for the establishment of guidelines for medications development (Carroll et al., 2014; Falk et al., 2014). The most common dichotomous outcome in alcohol clinical trials had been the percentage of participants abstinent from alcohol. This outcome has traditionally been the primary goal of most treatment programs and, until recently, was the only outcome accepted by the FDA for Phase 3 registration trials. This indicator has limitations as well, as those who experienced minor slips or drank at low-risk levels are deemed treatment failures (Falk et al., 2010).

Notably, the non-abstinence based outcome PSNHDD has now been accepted as a primary efficacy endpoint by the FDA based on analyses commissioned by NIAAA on longitudinal and observational datasets (Food and Drug Administration Draft Guidance, 2015). Work by Falk and colleagues, using data from two sets of alcohol clinical trials, demonstrated that PSNHDD at the end of treatment (given various grace period lengths, as discussed below) was associated with fewer alcohol-related consequences and lower levels of drinking during a 1-year follow-up period (Falk et al., 2010). Although the 'low-risk' drinking group (i.e., those with no HDDs but continued alcohol use) had worse follow-up drinking outcomes than those who were completely abstinent, they fared significantly better than those who had HDDs, thereby validating this non-abstinence based outcome as associated with fewer subsequent psychosocial and physical consequences. This outcome is further supported through: prospective study data indicating the frequency of heavy drinking increased the relative risks of mortality from cardiovascular disease and cancer (Breslow and Graubard, 2008), epidemiologic data indicating individuals with no heavy drinking days had a lower risk for developing alcohol dependence and alcohol use disorders than those who experienced heavy drinking days (Dawson et al., 2007), as well as through data from treatment centers indicating low-risk drinkers were similar to abstinent individuals with respect to long-term psychosocial outcomes (Kline-Simon et al., 2013) and treatment utilization and costs (Kline-Simon et al., 2014).

Two important aspects of the Falk et al. study validating PSNHDD that may be useful for identifying an outcome measure for stimulant trials are: (1) the evaluation of various grace periods that impacted the treatment's effect size, and (2) the use of the Drinker Inventory of Consequences (DRINC; Miller et al., 1995) as a measure of alcohol-related consequences. First, in the alcohol studies it was found that the detected differences between an active drug and placebo increased with each additional month of a grace period, such that the largest

effects were found in the final month of treatment (i.e., the longer the grace period, the larger the end of treatment effect). By discounting slips during an initial period of pharmacotherapy, a grace period may offer greater sensitivity for detecting success/failure outcomes in stimulant trials (McCann et al., 2015; McCann and Li, 2012). Second, a tool for assessing physical and psychosocial consequences of alcohol use (e.g., the DRINC) was sensitive to changes in alcohol use, and was able to differentiate low-risk from high-risk drinking. Although a parallel instrument has been developed to assess the consequences of drug use, the Inventory of Drug Use Consequences (InDUC; Tonigan and Miller, 2002), it is not specific for stimulant use (items refer to consequences of drinking or drug use) and is rarely included within stimulant treatment clinical trials.

There are three advantages that the alcohol field has relative to clinical trials for stimulant use disorders. The first is an accepted measure of a standard drink, given the clear and known size and alcohol content of different marketed beverages. Given that they are illicit, stimulants such as cocaine and methamphetamine have no similar standard approach to sizing and purity. Second, the alcohol field developed the concept of a ‘heavy drinking day’ that has been linked to alcohol-related consequences (Breslow and Graubard, 2008; Jackson, 2008). No such analogue has been developed for stimulant use. Third, medications (such as naltrexone) with established efficacy in the treatment of alcohol use disorders are available for use in validating clinical trial endpoints. Such medications are lacking for the treatment of stimulant use disorders.

5. WHERE ARE WE NOW?

NIDA has been particularly interested in the potential validation of an outcome measure other than sustained abstinence (e.g., ‘intermittent abstinence’ or ‘reduction in use’) that is clinically meaningful (i.e., associated with improvements in physical or psychosocial functioning) for use in clinical trials of pharmacotherapies for cocaine and methamphetamine use disorders. Several funding opportunities have been issued in recent years to address this area, and some resulting studies have produced promising findings (e.g., Carroll et al., 2014; Crits-Christoph et al., 2013; Garner et al., 2014; Kiluk et al., 2014; Lai et al., 2015). For instance, using pooled data from 5 randomized controlled trials of treatment for cocaine dependence, Carroll and colleagues (2014) identified several continuous (e.g., percentage of days abstinent, percentage of negative urine samples, maximum days of continuous abstinence), and dichotomous (e.g., achieving at least 3 weeks of continuous abstinence) measures that were significantly associated with cocaine use and a measure of ‘good functioning’ (defined as 0 days cocaine use, and 0 days of legal, employment, and psychological problems as reported on the Addiction Severity Index – ASI; McLellan et al., 1992) during a 12-month follow-up period. In a complementary analysis with the same dataset, Kiluk and colleagues (2014) used longitudinal growth curve modeling to demonstrate greater rates of abstinence during treatment (indicated by continuous or dichotomous measures) were associated with fewer reported problems across all ASI domains (‘global problems’) during follow-up. Although the magnitude of these relationships in both analyses were relatively modest, the findings hold promise for the use of a dichotomous outcome measure [≥ 3 weeks of continuous abstinence, which is an outcome measure used in some of the earliest trials for cocaine dependence (Gawin et al.,

1989)] as an indicator of treatment success, as it has been associated with less cocaine use and fewer physical and psychosocial problems in the long term.

There is also promising evidence that a reduction in cocaine use can have health benefits. For example, Lai and colleagues (2015) in a pilot study reported both cocaine abstinence and reduction in cocaine use (indicated by self-reported days of use) were associated with a decrease in endothelin-1 (ET-1), which is a marker of endothelial dysfunction and damage (and can contribute to hypertension and cardiac disease). Furthermore, the study found that the number of days of cocaine use was positively associated with ET-1 levels. This association held after controlling for family history of heart attack, baseline ET-1, and cardiovascular risk profile, suggesting a reduction in the number of days of cocaine use is independently associated with less endothelial damage. Although preliminary due to small sample size (n=57), these results offer promise for ET-1 as a potentially valid health outcome measure for medication trials that target cocaine use (McCann et al., 2015). For a thorough review of the evidence regarding biomarkers for cocaine use disorders, see Bough et al. (2014).

In terms of potential clinical outcome assessments for stimulant trials, the Cocaine Selective Severity Assessment (CSSA; Kampman et al., 1998) has been evaluated as an indicator of cocaine treatment benefit. The CSSA has been found to be a valid and reliable indicator of cocaine withdrawal syndrome (Gawin and Kleber, 1986), and several trials have found it predictive of poor treatment response (e.g., Ahmadi et al., 2006; Kampman et al., 2001, 2002). However, despite the CSSA's strong sensitivity at identifying poor treatment responders, its specificity at identifying patients likely to do well is fairly low, thereby limiting its utility as an outcome measure for clinical trials. The CSSA may be more useful as a stratification variable in study design of the clinical trial, or as a potential moderator in a Phase III medication trial that may inform labelling of medication effects. The ASI (McLellan et al., 1992) is another clinical outcome assessment that has been widely used in clinical trials with stimulant users, and has sound psychometric characteristics as a reliable and valid measure of severity (Alterman et al., 2007; McLellan et al., 2006). However, despite the clinical utility of the ASI, it has several limitations as an indicator of treatment success in clinical trials (Makela, 2004); most notably, the restricted functional range in scores found for most non-drug use problem areas limits the sensitivity to detect improvements due to treatment. Whereas treatments may demonstrate an effect on the frequency of stimulant use, improvements have not generally extended to the ASI composite scores representing psychosocial and other addiction-related problem areas (e.g., Crits-Christoph et al., 2001).

5.1 Perspectives to consider

As the field continues to evaluate potential outcome measures for use in stimulant clinical trials, it is important to consider/integrate the perspectives from various stakeholders including regulatory agencies, third party payers, healthcare organizations, and patients as to what might be considered meaningful. Notably, payers and the public generally see the meaningful outcome as abstinence; reductions in drug use (as opposed to reductions in alcohol use) are not as accepted as a good outcome. For instance, data presented at the

meeting regarding responses to a double-blinded online survey conducted with a sample of 34 payers (pharmacy and medical directors from US managed care plans) indicated the highest rated concepts in terms of value to their formulary decision-making process for new stimulant addiction pharmacotherapies were: (1) abstinence, (2) an effect on healthcare resource use (e.g., emergency room visits and hospitalizations), and (3) an impact on comorbidities (Duhig, 2015). Certainly, abstinence is not questioned as a meaningful endpoint for clinical trials, as it is the goal of most behavioral treatments and pharmacotherapies. However, the value of the impact on comorbidities and resource use is noteworthy, as these outcomes are not often measured or reported in clinical trials of pharmacotherapies for stimulant use. Most early stage pharmacotherapy trials exclude individuals with medical and/or psychiatric comorbidities in order to limit the impact on the potential effect of the medication under study. Also, while studies have indicated drug use treatment can reduce emergency department utilization and hospital admissions (Laine et al., 2001, 2005), determining the relative cost-benefit of pharmacotherapies in clinical trials is complicated by the relatively small numbers of subjects enrolled, the short duration of the trials, and the lag between treatment and health care benefits.

A meaningful outcome from a regulatory perspective should also be considered. In the Code of Federal Regulations, Title 21 Part 314, regarding the application for FDA approval to market a new drug, there is the potential to receive accelerated approval of new drugs for serious or life-threatening illnesses based on a surrogate endpoint or an effect on a clinical endpoint other than survival or irreversible morbidity (SECTION 314.510). This regulation states “FDA may grant marketing approval for a new drug product on the basis of adequate and well-controlled clinical trials establishing that the drug product has an effect on a surrogate endpoint that is reasonably likely, based on epidemiologic, therapeutic, pathophysiologic, or other evidence, to predict clinical benefit or on the basis of an effect on a clinical endpoint other than survival or irreversible morbidity. (http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/cfrsearch.cfm?cfrpart=314&showfr=1&subpartnode=21:5.0.1.1.4.8). Again, as drug use during the brief window of a clinical trial is considered a surrogate endpoint (Winchell et al., 2012), the goal should be to demonstrate that some measure of drug abstinence (or reduced use) is ‘reasonably likely’ to be predictive of clinical benefit, such as improved physical or psychosocial functioning. However, the question remains: what specific duration of abstinence and what type of physical or psychosocial outcome assessment would be meaningful and/or persuasive to accept a pharmacotherapy as efficacious?

6. CONCLUSIONS AND RECOMMENDATIONS

Considering the guidance of the FDA’s CDER (Center for Drug Evaluation and Research, 2014), the lessons learned from other addiction trials, and the process of identifying a valid endpoint for measuring treatment success in alcohol use clinical trials, the following conclusions and recommendations are proposed (see Table 1):

7. SUMMARY

This ACTTION-sponsored meeting brought together participants from academia, FDA, NIDA, NIAAA, and healthcare and pharmaceutical organizations with the targeted goal of generating a research agenda for the field of stimulant use. The goal was to move toward identifying a clinically meaningful outcome measure, other than long-term abstinence, for use in stimulant treatment clinical trials. The participants agreed that the alcohol treatment outcome measure (PSNHDD) development was a helpful guide in the quest to validate a similar endpoint for the treatment of stimulant use disorders. Patterns of stimulant use are not dissimilar from patterns of alcohol use, and there are some parallels in the handling of missing data in clinical trials. However, there are also challenges unique to measuring stimulant use that require careful consideration of the various methods to identify a valid outcome measure. Through discussion of these challenges, review of the current state of the evidence, and attention to the perspectives of multiple stakeholders, a list of research recommendations for generating a meaningful endpoint to be used in clinical trials for stimulant use disorders was developed. Although this is not the first time a list of research recommendations have been proposed from an expert panel or task force regarding meaningful outcome measures in drug use clinical trials (e.g., Clinical Trials Network, 2010; Donovan et al., 2012; Tiffany et al., 2012), the current list represents both an expansion and refinement of prior recommendations based on the knowledge gained in recent years and the feedback generated from various stakeholders. The above recommendations should be considered a blueprint, for both the design of future clinical trials, as well as for potential secondary analyses of stimulant use clinical trial data intended to reveal evidence of meaningful treatment benefit to advance the development/approval of new drugs to treat stimulant use disorders.

Acknowledgments

Role of Funding Source

The work of authors who are NIH Intramural staff was supported in part by the NIDA Intramural Research Program. The work of authors Kiluk and Carroll was supported in part by NIDA grant P50-DA09241 (Carroll, PI). Author Strain was supported in part by NIDA grant K24-DA023186. The funding sources had no further role in the writing of this report.

We thank Valorie Thompson and Andrea Speckin for their assistance in organizing the meeting, and Sharon Hertz, MD, Allison H. Lin, PharmD, PhD, and Celia Winchell, MD, for their valuable participation.

References

Acri, JB.; Skolnick, P. Pharmacotherapy of substance use disorders. In: Charney, D.; Nestler, E.; Sklar, P.; Buxbaum, J., editors. Neurobiology Of Mental Illness. 4. Oxford University Press; London: 2013. p. 235-245.

Ahmadi J, Kampman K, Dackis C. Outcome predictors in cocaine dependence treatment trials. Am J Addict. 2006; 15:434–439. [PubMed: 17182445]

Alterman AI, Cacciola JS, Habing B, Lynch KG. Addiction Severity Index Recent and Lifetime summary indexes based on nonparametric item response theory methods. Psychol Assess. 2007; 19:119–132. [PubMed: 17371127]

Bough KJ, Amur S, Lao G, Hemby SE, Tannu NS, Kampman KM, Schmitz JM, Martinez D, Merchant KM, Green C, Sharma J, Dougherty AH, Moeller FG. Biomarkers for the development of

new medications for cocaine dependence. Neuropsychopharmacology. 2014; 39:202–219. [PubMed: 23979119]

Breslow RA, Graubard BI. Prospective study of alcohol consumption in the United States: quantity, frequency, and cause-specific mortality. Alcohol Clin Exp Res. 2008; 32:513–521. [PubMed: 18215212]

Carroll KM, Kiluk BD, Nich C, DeVito EE, Decker S, LaPaglia D, Duffey D, Babuscio TA, Ball SA. Towards empirical identification of a clinically meaningful indicator of treatment outcome for drug addiction: features of candidate indicators and evaluation of sensitivity to treatment effects and relationship to one year cocaine use follow-up outcomes. Drug Alcohol Depend. 2014; 137:3–19. [PubMed: 24556275]

Center for Drug Evaluation and Research. U.S. Department of Health and Human Services. Guidance for Industry and FDA Staff: Qualification Process for Drug Development Tools. Silver Spring, MD: 2014.

Clinical Trials Network. [accessed on October 23, 2015] Treatment Effect and Assessment Measures Task Force, Executive Summary. 2010. http://ctndisseminationlibrary.org/PDF/522.pdf

Cone EJ, Sampson-Cone AH, Darwin WD, Huestis MA, Oyler JM. Urine testing for cocaine abuse: metabolic and excretion patterns following different routes of administration and methods for detection of false-negative results. J Anal Toxicol. 2003; 27:386–401. [PubMed: 14606991]

Cottler LB, Compton WM, Ben-Abdallah A, Horne M. Achieving a 96.6 percent follow-up rate in a longitudinal study of drug abusers. Drug Alcohol Depend. 1996; 41:209–217. [PubMed: 8842633]

Crits-Christoph P, Gallop P, Gibbons MB, Sadicario JS, Woody G. Measuring outcome in the treatment of cocaine dependence. J Alcohol Drug Depend. 2013; 1:1–8.

Crits-Christoph P, Siqueland L, McCalmont E, Weiss RD, Gastfriend DR, Frank A, Moras K, Barber JG, Blaine JD, Thase ME. Impact of psychosocial treatments on associated problems of cocaine-dependent patients. J Consult Clin Psychol. 2001; 69:825–830. [PubMed: 11680559]

Dawson DA, Goldstein RB, Grant BF. Rates and correlates of relapse among individuals in remission from DSM-IV alcohol dependence: a 3-year follow-up. Alcohol Clin Exp Res. 2007; 31:2036–2045. [PubMed: 18034696]

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 endpoints in clinical trials. Addiction. 2012; 107:694–708. [PubMed: 21781202]

Duhig, AM. U.S. Payer Feedback on Stimulant Addiction Treatment. Proceedings of the Measures of Outcome for Stimulant Trials, an ACTTION sponsored conference; 2015.

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]

Falk DE, Litten RZ, Anton RF, Kranzler HR, Johnson BA. Cumulative proportion of responders analysis (CPRA) as a tool to assess treatment outcome in alcohol clinical trials. J Stud Alcohol Drugs. 2014; 75:335–346. [PubMed: 24650828]

Farabee D, Hawken A, Griffith P. Tracking and incentivizing substance abusers in longitudinal research: results of a survey of National Institute on Drug Abuse-funded investigators. J Addict Med. 2011; 5:87–91. [PubMed: 21769053]

Festinger DS, Marlowe DB, Dugosh KL, Croft JR, Arabia PL. Higher magnitude cash payments improve research follow-up rates without increasing drug use or perceived coercion. Drug Alcohol Depend. 2008; 96:128–135. [PubMed: 18395365]

Food and Drug Administration Draft Guidance. Alcoholism: Developing Drugs for Treatment, Guidance for Industry. Silver Spring, MD: 2015.

Food and Drug Administration: Psychopharmacologic Drugs Advisory Committee. Probuphine (Buprenorphine Hydrochloride Subdermal Implant) For Maintenance Treatment Of Opioid Dependence. Silver Spring, MD: 2013.

Garner BR, Scott CK, Dennis ML, Funk RR. The relationship between recovery and health-related quality of life. J Subst Abuse Treat. 2014; 47:293–298. [PubMed: 25012552]

Gawin FH, Kleber HD. Abstinence symptomatology and psychiatric diagnosis in cocaine abusers. Clinical observations. Arch Gen Psychiatry. 1986; 43:107–113. [PubMed: 3947206]

Gawin FH, Kleber HD, Byck R, Rounsaville BJ, Kosten TR, Jatlow PI, Morgan C. Desipramine facilitation of initial cocaine abstinence. Arch Gen Psychiatry. 1989; 46:117–121. [PubMed: 2492422]

Hjorthøj CR, Hjorthøj AR, Nordentoft M. Validity of Timeline Follow-Back for self-reported use of cannabis and other illicit substances - systematic review and meta-analysis. Addict Behav. 2012; 37:225–233. [PubMed: 22143002]

Jackson KM. Heavy episodic drinking: determining the predictive utility of five or more drinks. Psychol Addict Behav. 2008; 22:68–77. [PubMed: 18298232]

Kampman KM, Alterman AI, Volpicelli JR, Maany I, Muller ES, Luce DD, Mulholland EM, Jawad AF, Parikh GA, Mulvaney FD, Weinrieb RM, O'Brien CP. Cocaine withdrawal symptoms and initial urine toxicology results predict treatment attrition in outpatient cocaine dependence treatment. Psychol Addict Behav. 2001; 15:52–59. [PubMed: 11255939]

Kampman KM, Volpicelli JR, McGinnis DE, Alterman AI, Weinrieb RM, D'Angelo L, Epperson LE. Reliability and validity of the Cocaine Selective Severity Assessment. Addict Behav. 1998; 23:449–461. [PubMed: 9698974]

Kampman KM, Volpicelli JR, Mulvaney F, Rukstalis M, Alterman AI, Pettinati H, Weinrieb RM, O'Brien CP. Cocaine withdrawal severity and urine toxicology results from treatment entry predict outcome in medication trials for cocaine dependence. Addict Behav. 2002; 27:251–260. [PubMed: 11817766]

Kiluk BD, Nich C, Witkiewitz K, Babuscio TA, Carroll KM. What happens in treatment doesn't stay in treatment: Cocaine abstinence during treatment is associated with fewer problems at follow-up. J Consult Clin Psychol. 2014; 82:619–627. [PubMed: 24635550]

Kleschinsky JH, Bosworth LB, Nelson SE, Walsh EK, Shaffer HJ. Persistence pays off: follow-up methods for difficult-to-track longitudinal samples. J Stud Alcohol Drugs. 2009; 70:751–761. [PubMed: 19737500]

Kline-Simon AH, Falk DE, Litten RZ, Mertens JR, Fertig J, Ryan M, Weisner CM. Posttreatment low-risk drinking as a predictor of future drinking and problem outcomes among individuals with alcohol use disorders. Alcohol Clin Exp Res. 2013; 37(Suppl 1):E373–380. [PubMed: 22827502]

Kline-Simon AH, Weisner CM, Parthasarathy S, Falk DE, Litten RZ, Mertens JR. Five-year healthcare utilization and costs among lower-risk drinkers following alcohol treatment. Alcohol Clin Exp Res. 2014; 38:579–586. [PubMed: 24117604]

Lai H, Stitzer M, Treisman G, Moore R, Brinker J, Gerstenblith G, Kickler TS, Li J, Chen S, Fishman E, Lai S. Cocaine abstinence and reduced use associated with lowered marker of endothelial dysfunction in African Americans: a preliminary study. J Addict Med. 2015; 9:331–339. [PubMed: 26164164]

Laine C, Hauck WW, Gourevitch MN, Rothman J, Cohen A, Turner BJ. Regular outpatient medical and drug abuse care and subsequent hospitalization of persons who use illicit drugs. JAMA. 2001; 285:2355–2362. [PubMed: 11343483]

Laine C, Lin YT, Hauck WW, Turner BJ. Availability of medical care services in drug treatment clinics associated with lower repeated emergency department use. Med Care. 2005; 43:985–995. [PubMed: 16166868]

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]

Magura S, Kang SY. Validity of self-reported drug use in high risk populations: a meta-analytical review. Subst Use Misuse. 1996; 31:1131–1153. [PubMed: 8853234]

Makela K. Studies of the reliability and validity of the Addiction Severity Index. Addiction. 2004; 99:398–410. discussion 411-398. [PubMed: 15049734]

McCann D, Ramey T, Skolnick P. Outcome measures in medication trials for substance use disorders. Curr Treat Options Psychiatry. 2015; 2:113–121.

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 Ther. 2012; 18:414–418. [PubMed: 22070720]

McLellan AT, Cacciola JC, Alterman AI, Rikoon SH, Carise C. The Addiction Severity Index at 25: origins, contributions and transitions. Am J Addict. 2006; 15:113–124. [PubMed: 16595348]

McLellan AT, Kushner H, Metzger D, Peters R, Smith I, Grissom G, Pettinati H, Argerious M. The fifth edition of the Addiction Severity Index. J Subst Abuse Treat. 1992; 9:199–213. [PubMed: 1334156]

McPherson S, Barbosa-Leiker C, Mamey MR, McDonell M, Enders CK, Roll J. A 'missing not at random' (MNAR) and 'missing at random' (MAR) growth model comparison with a buprenorphine/naloxone clinical trial. Addiction. 2015; 110:51–58. [PubMed: 25170740]

Miller, WR.; Tonigan, JS.; Longabaugh, R. Test Manual. NIAAA; Rockville, Maryland: 1995. The Drinker Inventory of Consequences (DrInC): An Instrument For Assessing Adverse Consequences Of Alcohol Abuse.

National Research Council. Committee on National Statistics. Panel on Handling Missing Data in Clinical Trials. The National Academy Press; Washington, DC: 2010. The Prevention and Treatment of Missing Data in Clinical Trials.

Oden NL, VanVeldhuisen PC, Wakim PG, Trivedi MH, Somoza E, Lewis D. Power of automated algorithms for combining time-line follow-back and urine drug screening test results in stimulant-abuse clinical trials. Am J Drug Alcohol Abuse. 2011; 37:350–357. [PubMed: 21854277]

Oyler JM, Cone EJ, Joseph RE Jr, Moolchan ET, Huestis MA. Duration of detectable methamphetamine and amphetamine excretion in urine after controlled oral administration of methamphetamine to humans. Clin Chem. 2002; 48:1703–1714. [PubMed: 12324487]

Preston KL, Epstein DH, Cone EJ, Wtsadik AT, Huestis MA, Moolchan ET. Urinary elimination of cocaine metabolites in chronic cocaine users during cessation. J Anal Toxicol. 2002; 26:393–400. [PubMed: 12422991]

Preston KL, Goldberger BA, Cone EJ. Occurrence of cocaine in urine of substance-abuse treatment patients. J Anal Toxicol. 1998; 22:580–586. [PubMed: 9847008]

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]

Scott CK. A replicable model for achieving over 90% follow-up rates in longitudinal studies of substance abusers. Drug Alcohol Depend. 2004; 74:21–36. [PubMed: 15072804]

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]

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]

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. 2012; 107:709–718. [PubMed: 21981638]

Tonigan JS, Miller WR. The Inventory of Drug Use Consequences (InDUC): test-retest stability and sensitivity to detect change. Psychol Addict Behav. 2002; 16:165–168. [PubMed: 12079257]

Winchell C, Rappaport BA, Roca R, Rosebraugh CJ. Reanalysis of methamphetamine dependence treatment trial. CNS Neurosci Ther. 2012; 18:367–368. [PubMed: 22533722]

Witkiewitz K, Falk DE, Kranzler HR, Litten RZ, Hallgren KA, O'Malley SS, Anton RF. Methods to analyze treatment effects in the presence of missing data for a continuous heavy drinking outcome measure when participants drop out from treatment in alcohol clinical trials. Alcohol Clin Exp Res. 2014; 38:2826–2834. [PubMed: 25421518]

Highlights

- Meaningful outcome measures, other than sustained abstinence, would be beneficial.
- There are multiple challenges unique to measuring stimulant use.
- Meaningful alternative outcomes should be associated with functional benefits.
- A patient reported outcome may be useful for validating a stimulant use endpoint.

Table 1

Conclusions and Recommendations

Conclusions	Recommendations for future research
1. Qualified Patient Reported Outcomes are accepted by the FDA for measuring treatment benefit, and may be used to validate a measure of stimulant use as a surrogate endpoint that is reasonably likely to predict clinical benefit. This would parallel the alcohol field's use of the DRINC as a tool for validating PSNHDD as a meaningful endpoint.	Emphasis should be placed on the development or modification of a standardized Patient Reported Outcome instrument for measuring the physical and psychosocial problems/consequences that characterize stimulant use disorders. Such an instrument should demonstrate sensitivity to changes in stimulant use.
2. Long-term abstinence should remain the goal of treatments for stimulant use disorders, and the target endpoint for clinical trials. However, alternative outcomes (e.g., periods of 'intermittent abstinence') should be evaluated with respect to associated functional outcomes.	Secondary analysis of clinical trial data should compare various patterns of stimulant use (e.g., 1–4 days per month vs. >4 days per month) and the associated physical and psychosocial consequences. This may help define a pattern of 'intermittent abstinence' that is associated with more favorable physical or psychosocial outcomes as compared to 'regular use' (in the same manner that low-risk drinkers had more favorable outcomes compared to heavy drinkers).
3. Measures of 'reduction' in stimulant use that are based on the quantity of use per day (either through self-report or quantitative urinalysis) are too unreliable for consideration as a valid outcome measure. There does not appear to be a stimulant use equivalent of a 'heavy drinking day' in terms of amount consumed per day.	Reduction-based measures defined by a reduction in the quantity of use per day/episode (self-report or urinalysis) should be abandoned. Any measure of reduction should be based on the frequency of days of stimulant use (either per week or per month).
4. Urine drug screens should continue to be an essential component of clinical trials for stimulant use disorders. However, they should mainly be used to corroborate self- reported use/ abstinence, rather than as a primary outcome measure (i.e., moving away from a primary emphasis on comparing the percentage of positive/negative urine results across treatment conditions).	In clinical trials, urine collection might be limited to once per week during the treatment period to confirm abstinence and reduce participant burden. Future trials might also explore the use of contingency management procedures to increase the accuracy of self- reported stimulant use (e.g., provide $10 voucher if urine result matches self-report).
5. Missing data are one of the biggest threats to measurement of stimulant use/abstinence in clinical trials. All methods of statistical imputation to handle missing data include subjective assumptions. The best way to handle missing data is to limit it as much as possible.	All methods for collecting data from treatment drop-outs should be utilized. This includes: providing transportation to/from appointments, offering alternate locations for assessment interviews, obtaining permission to contact significant others who might know how to contact the individual if staff unable to do so, and incorporating multiple methods of communication with participants (e.g., phone, text message, email, social media, and postal mail). Persistence is key in this endeavor (Cottler et al., 1996; Farabee et al., 2011; Festinger et al., 2008; Kleschinsky et al., 2009; Scott, 2004)
6. The existence of an established efficacious medication for reducing alcohol use facilitated the process of validating a meaningful outcome measure for use in clinical trials for the treatment of alcohol use disorders. Although such a medication does not exist for stimulant use disorders, there are behavioral treatments with established efficacy at reducing stimulant use (e.g. Contingency Management) that may be useful for validating an outcome measure.	Data from contingency management trials for stimulant use disorders could be used for defining an outcome measure that is indicative of treatment success. Such a measure could then be used to evaluate success in pharmacotherapy trials.