HHS Public Access

Author manuscript

Alcohol Clin Exp Res. Author manuscript; available in PMC 2016 September 01.

Published in final edited form as:

Alcohol Clin Exp Res. 2015 September ; 39(9): 1571–1581. doi:10.1111/acer.12797.

Guidelines for the Reporting of Treatment Trials for Alcohol Use Disorders

Katie Witkiewitz, PhD[1,2], **John W. Finney**[3], **Alex H.S Harris**[3,4], **Daniel R. Kivlahan**[5], and **Henry R. Kranzler**[6]

[1]Department of Psychology, University of New Mexico

[2]Center on Alcoholism, Substance Abuse, and Addictions, University of New Mexico

[3]Center for Innovation to Implementation, VA Palo Alto Health Care System, Menlo Park, CA

[4]VA Substance Use Disorder Quality Enhancement Research Initiative, VA Palo Alto Health Care System, Menlo Park, CA

[5]Veterans Health Administration, Washington, DC and Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle

[6]Center for Studies of Addiction, Department of Psychiatry, University of Pennsylvania Perelman School of Medicine and VISN4 MIRECC, Philadelphia VAMC, Philadelphia, PA

Abstract

Background—The primary goals in conducting clinical trials of treatments for alcohol use disorders (AUDs) is to identify efficacious treatments and determine which treatments are most efficacious for which patients. Accurate reporting of study design features and results is imperative to enable readers of research reports to evaluate to what extent a study has achieved these goals. Guidance on quality of clinical trial reporting has evolved substantially over the past two decades, primarily through the publication and widespread adoption of the Consolidated Standards of Reporting Trials (CONSORT) statement. However, there is room to improve the adoption of those standards in reporting the design and findings of treatment trials for AUD.

Methods—Narrative review of guidance on reporting quality in AUD treatment trials.

Results—Despite improvements in the reporting of results of treatment trials for AUD over the past two decades, many published reports provide insufficient information on design or methods.

Conclusions—The reporting of alcohol treatment trial design, analysis, and results requires improvement in four primary areas: (1) trial registration, (2) procedures for recruitment and retention, (3) procedures for randomization and intervention design considerations, and (4) statistical methods used to assess treatment efficacy. Improvements in these areas and the adoption

Correspondence should be addressed to Katie Witkiewitz, Department of Psychology, University of New Mexico, 1 University of New Mexico, MSC03 2220, Albuquerque, NM 87131-0001, Phone: 505-277-5953, Fax: 505-925-1394, katiew@unm.edu..

Disclosures: H. Kranzler has served as a consultant or advisory board member for Alkermes, Lilly, Lundbeck, Otsuka, Pfizer, and Roche. K. Witkiewitz has served as a consultant for Alkermes. H. Kranzler is a member of and K. Witkiewitz has received financial support from the American Society of Clinical Psychopharmacology's Alcohol Clinical Trials Initiative, which is supported by AbbVie, Ethypharm, Lilly, Lundbeck, and Pfizer.

of reporting standards by authors, reviewers, and editors are critical to an accurate assessment of the reliability and validity of treatment effects. Continued developments in this area are needed to move AUD treatment research forward via systematic reviews and meta-analyses that maximize the utility of completed studies.

Keywords

Alcohol clinical trials; CONSORT; eligibility criteria; multisite studies; alcohol use disorder

The ultimate goal of most randomized clinical trials (RCTs) is to identify efficacious treatments for a sample of individuals. The accurate reporting of trial design features, sample characteristics, statistical methods, and results is imperative for readers to determine whether a given study has achieved this goal. In addition, clearly reported information is needed to generate useful syntheses of research findings. Over the past few decades, several attempts have been made to improve the quality of reporting of RCTs, with the primary goal being increased transparency and confidence in the reliability and validity of trial findings. Importantly, many aspects of clinical trial design (e.g., eligibility criteria (Humphreys et al 2008), randomization allocation) and analysis (e.g., procedures for handling missing data), if not well chosen and well implemented, can introduce systematic biases into the estimates of treatment efficacy (Witkiewitz et al., under review). Reporting of design features and analysis methods that may impact estimates of treatment efficacy is particularly important in RCTs evaluating treatments for alcohol use disorders (AUDs), given that relatively few efficacious treatments exist (Jonas et al., 2014; National Institute for Health and Care Excellence, 2011; Zindel and Kranzler, 2014).

Witkiewitz and colleagues (under review) provide an overview of recommendations for the design and analysis of alcohol treatment studies, including (1) considerations for recruitment, randomization, and retention; (2) the selection of measures to assess outcomes, intervention fidelity, and adherence; (3) timing of assessments and the study of mechanisms of behavior change; and (4) statistical methods used to assess treatment efficacy.

Here, we provide a brief review of the Consolidated Standards of Reporting Trials (CONSORT; Altman et al., 2001; Begg et al., 1996; Schulz et al., 2010). We also discuss empirical findings that guide recommendations to improve the quality of reporting of findings from AUD treatment trials. To reduce overlap with prior reviews of reporting standards, we focus on the reporting of study features in four primary areas relevant to alcohol treatment studies: (1) trial registration, (2) procedures for recruitment and retention, (3) procedures for reporting on randomization and intervention design considerations, and (4) the statistical methods used to assess treatment efficacy.

Reporting Guidelines

The most influential approach to instituting reporting standards was developed by the Consolidated Standards of Reporting Trials (CONSORT) group in a statement first published in 1996 (Begg et al., 1996) and revised in 2001 (Altman et al., 2001; Moher et al., 2001) and 2010 (Schulz et al., 2010). The CONSORT statement is a description of critical reporting elements in clinical trials. It includes a diagram to summarize participant flow

through an RCT and a checklist of 25 reportable trial design considerations that have been shown to influence the reliability and validity of treatment efficacy estimates. The CONSORT statement has been endorsed by 585 journals, many of which require the inclusion of the CONSORT diagram with all RCTs. The CONSORT diagram (http://www.consort-statement.org/consort-statement/flow-diagram, last accessed March 22, 2015) summarizes participant flow, including enrollment, randomization, allocation, follow-up, and analysis in each of the intervention and control groups. The items in the CONSORT checklist (http://www.consort-statement.org/consort-statement/checklist#/checklists/, last accessed March 22, 2015) covers various aspects of trial reporting, including the specific hypotheses tested, description of primary and secondary outcomes (i.e., endpoints), details on the method of randomization, effect sizes and trial limitations. The widespread adoption and endorsement of the CONSORT guidelines have led to improvements in clinical trial reporting (To et al., 2013), but a systematic review of 16,604 RCTs concluded that many fail to follow the guidelines and suboptimal reporting of results is still a prevalent problem (Turner et al., 2012).

Notably, because the CONSORT statement was developed for pharmacological studies, many aspects of it are less relevant for non-pharmacological treatments (e.g., because blinding of therapists and participants is not always possible in behavioral intervention research) and trials that involve patient-reported outcomes, both of which are relevant to RCTs for AUDs. To address these limitations, the CONSORT group extended and amended the guidelines for non-pharmacological treatments (Boutron et al., 2008), including a modified participant flow diagram and the addition of an item that to detail the experimental and control treatments and how they are implemented. Other CONSORT extensions include guidelines for reporting RCTs in journal and conference abstracts (Hopewell et al., 2008a, 2008b) and recommendations for describing patient-reported outcomes (PRO) in RCTs (Calvert et al., 2013).

Compliance with Guidelines in Alcohol Clinical Trials

A systematic review and meta-analysis (Ladd et al., 2010) examined congruence with CONSORT reporting standards in 127 alcohol treatment RCTs. The authors found a significant improvement in the overall quality of reporting over time. However, Ladd and colleagues (2010) also found many CONSORT recommendations were followed in reports of less than 30% of studies, including the reporting of procedures used to identify the target sample size, method of randomization concealment, evidence for the success of the blinding procedure, stopping rules, details on protocol deviations, and identification of prognostic variables by treatment group. Many other critical aspects of trial design (e.g., the method used to generate the randomization schedule) and analyses (e.g., effect sizes for the intervention on primary and secondary outcomes) were reported by 60% or fewer of the studies reviewed.

Other recent studies have documented problems in the quality of reporting of alcohol treatment trial outcomes. A meta-analysis of pharmacotherapy for AUD delivered in outpatient settings (Jonas et al., 2014) included ratings of the potential risk of bias given the details provided in the original RCTs (including questions about randomization, blinding,

attrition, validity and reliability of measures, whether intent-to-treat analyses were used, fidelity, and missing data methods): only 7 of 122 RCTs met criteria for “low” risk of bias as defined by independent raters. Similarly, Del Re and colleagues (2013) evaluated the analytic approaches of 165 alcohol pharmacotherapy trials and found that only 44.8% (74 trials) reported use of an intent-to-treat (ITT) analytic strategy, often considered the “gold standard” method to analyze RCT data. Fewer than 40% of the studies that claimed to conduct ITT analyses actually conducted ITT analyses correctly, as called for by the CONSORT standards. In addition, the majority of trials did not report on which missing data strategy was used or relied on a demonstrably flawed method to handle missing data (mostly last observation carried forward; Hallgren and Witkiewitz, 2013).

Based on these reviews, it is clear that the majority of alcohol clinical trials do not meet the reporting standards set forth by the CONSORT group. With these prior findings as the foundation, the remainder of this paper will review selected reporting standards and empirical findings related to RCTs for alcohol and other substance use disorders. The primary goal of this review is to guide investigators, reviewers, and editors in improving the reporting of design features in alcohol treatment trials to facilitate the development of better treatments that can be made available to patients with AUDs.

Trial Pre-Registration as a Critical First Step

ClinicalTrials.gov is a repository for the registration and reporting of results from clinical trials in human participants. ClinicalTrials.gov was established to increase transparency in the design, conduct, analysis, and reporting of clinical trials, as required by Section 801 of the Food and Drug Administration Amendments Act (Public Law 110-95, 2007; Zarin and Tse, 2008). The European Medicines Agency (EMA), the European Union, and the U. S. National Institutes of Health (NIH) have also adopted the use of the registry.

As of December 2014, the ClinicalTrials.gov database included descriptions of 181,224 clinical trials being conducted in all 50 of the United States and in 187 countries. Of these registrations, 2,356 (1.3%) included individuals with AUDs or drinking level and/or AUD symptoms as the primary outcomes. This extensive database and pre-registration will only meaningfully increase transparency to the extent that it impacts peer-reviewed studies and their reports. Currently, most editors and reviewers are not referencing the registry record in evaluating submitted manuscripts detailing RCT results (Mathieu et al., 2013) and often considerable discrepancies exist between the proposed analyses that were initially reported in the registries and the final analyses presented in published reports (Dwan et al., 2011). Further, a review of the results from published trials that were registered in ClinicalTrials.gov suggests that most deviated from the pre-registered plan of analysis in metric, method of aggregation, or timing of primary outcome measures employed, and the majority of studies (69%) failed to include all participants in the final published analyses (Zarin et al., 2011). No study has explicitly examined discrepancies from trial registration in reporting alcohol treatment trials, but it is likely to be similarly problematic. It is imperative for trials to be pre-registered, to follow the proposed protocol, and to report any deviations from the registry protocol in the published trial results.

Reporting on Procedures for Recruitment and Retention

A discussion of all of the relevant procedures for recruitment and data collection is beyond the scope of the current review (see Witkiewitz et al., under review). A collection of articles in the July 2005 supplement to the *Journal of Studies on Alcohol* provides extensive coverage of issues related to recruitment and retention (Zweben et al., 2005), developing an assessment battery for data collection (Gastfriend et al., 2005), and coordinating and monitoring multisite studies (Youngblood et al., 2005). Given prior extensive coverage of these topics, we focus here on two areas that have been problematic in prior reports of alcohol clinical trial findings: eligibility criteria and sample description, and participant retention.

Eligibility criteria and sample description

Criteria for inclusion and exclusion must balance internal and external validity for a particular trial (Carroll and Rounsaville, 2003; Heather, 2014). Seminal research by Humphreys and colleagues (Humphreys and Weisner, 2000; Humphreys et al., 2008; Humphreys et al., 2005) has shown that most alcohol clinical trials apply extensive exclusion criteria that systematically oversample patients who are higher functioning, less ethnically and racially diverse, and more economically stable than those treated in clinical practice settings. For example, there is a tendency to exclude patients with psychiatric or other substance use comorbidity (Hoertel et al., 2014) and those with severe medical problems or unstable housing, all factors associated with a poorer prognosis in clinical settings (Humphreys et al., 2008). Importantly, specific exclusion criteria predict either underestimation or overestimation of treatment effect sizes, and the direction and degree of under- or overestimation can vary across studies (Humphreys et al., 2008). Given these findings, it is imperative that researchers carefully follow CONSORT guidelines in reporting all eligibility criteria, the rate at which potential participants are excluded from the final sample, and the basis for excluding potential participants.

Beyond describing the process by which participants were recruited, it is important that study reports clearly describe the sample's demographic and clinical characteristics. Investigators should report age, gender, race and ethnicity, weight and height or body mass index, education, employment status, marital status, the presence of comorbid psychiatric or drug use disorders, the severity and duration of AUD, and key baseline alcohol consumption measures. Other demographic and clinical characteristics relevant to the trial (e.g., concomitant medications, treatment history, housing status) may also need to be described. Reviews of alcohol treatment trials have shown that published reports are generally deficient in providing information on participants' demographic characteristics, and history of alcohol misuse or disorders (e.g., Swearingen et al., 2003). This information is critical in trying to evaluate the relative efficacy of different treatment approaches in research syntheses, as some participant populations may be more responsive to treatment ("easier to treat") than others.

Reporting on Study Retention

The CONSORT diagram provides clear instructions on how to present the number of individuals who prematurely discontinue the trial and those who complete it. It is particularly important to distinguish between patients who dropped out due to the treatment itself (e.g., those experiencing adverse events or no perceived treatment benefit), those who dropped out due to reasons potentially unrelated to the treatment (e.g., change in residence or employment status), and those who were lost to follow-up (i.e., did not respond to the researcher's prompts or reminders for assessment). Efforts made to increase retention should be described, as well as reasons for non-adherence (Zweben et al., 2009, 2005). There is also a CONSORT extension for reporting of the harms to patients in clinical trials (Ioannidis et al., 2004). Reporting on adverse events is commonly lacking in non-pharmacological trials (Jonas et al., 2013), but is essential in grading the evidence in terms of relative harms and benefits as part of systematic reviews. This information should be provided for all treatment conditions and at all assessment time points.

The Description of Random Assignment to Intervention or Comparator Conditions, Interventionists, and Assessment Procedures

Procedures for random assignment

The RCT is often considered the gold standard to evaluate intervention efficacy, although the procedures used in randomization can jeopardize the validity of findings (Berger and Weinstein, 2004; Berger, 2006; Grossman and Mackenzie, 2005; Viera and Bangdiwala, 2007). The CONSORT statement (Schulz et al., 2010) recommends that clinical trial reports include four aspects of randomization: the method used to generate the random allocation sequence (i.e., how the investigators ensured that patients have a known, not necessarily equal, probability of receiving each intervention before assignment and that the assignment is determined entirely by chance), the randomization type (e.g., blocked or stratified randomization), the allocation concealment method (i.e., the individuals enrolling patients have no knowledge of treatment assignment and are not allowed to accept or reject a patient based on treatment assignment), and how randomization was implemented, including who allocated, who enrolled, and who assigned participants to trial groups. In blinded treatment studies, the investigator should report whether the blind was broken and the circumstances that necessitated the unmasking of the blind (Haahr and Hróbjartsson, 2006).

Although numerous studies have examined biases resulting from inadequate randomization procedures across various types of clinical trials (Herbison et al., 2011; Savović et al., 2012; Schulz et al., 1995; Wood et al., 2008), we were unable to locate any studies of randomization biases that might occur specifically in alcohol clinical trials. Future research should examine the degree and frequency with which alcohol clinical trials violate recommended randomization procedures and the effects of randomization bias on estimates of intervention efficacy.

Describing the intervention and study procedures

The intervention is typically the primary focus of alcohol treatment trials, yet often little information is provided about the specific interventions tested. The CONSORT checklist clearly states that interventions, including control groups, should be described with "sufficient details to allow replication." For pharmacological interventions, this includes the drug name, dose, method of administration, titration procedures, and timing and duration of administration. The method for evaluating compliance with the medication, such as urinary riboflavin, medication event monitoring systems (MEMS), or blister packs, also should be reported.

For non-pharmacological interventions, the CONSORT checklist includes additional recommendations for reporting the specific details of the experimental and comparator conditions, including a description of the components of the interventions and any tailoring of the interventions to individual participants, details regarding standardization of the interventions and adherence to the intervention protocols, and how adherence to intervention protocols was assessed. All of these recommendations can be particularly important in alcohol treatment trials in which "treatment as usual" or "standard care" interventions are included as comparison interventions (Brigham et al., 2009), as these approaches may vary substantially across studies. This information is crucial to allow the relative effects of different interventions to be assessed in research syntheses.

It is also critical to report whether individual or group intervention delivery methods were utilized. If group interventions are used, the number of patients within each group and the potential impact of within-group effects on intervention outcomes should be reported. Although most trials have not shown a significant difference between group and individual intervention delivery methods (Graham et al., 1996; Marques and Formigoni, 2001; Sobell et al., 2009), the cohesion within any one group of a trial could impact the overall intervention efficacy, so that the clustering of individuals within groups should be considered in the analyses (Greenfield et al., 2013).

Describing the interventionists and treatment fidelity

In both pharmacological and non-pharmacological interventions for AUDs, there is often substantial patient interaction with a therapist, nurse, or other health care professional. The effects of such interventionists often explain a considerable amount of the variance in treatment outcomes (with estimates ranging from 3% to upwards of 12%: Imel et al., 2008; Moyers and Miller, 2013).

To help characterize the degree to which interactions with health care professionals might affect treatment outcomes, it is critical to report their training, the number of interactions they have with the patients, the average length of the meetings, and the content of those interactions including whether a prescribed intervention method was followed. This information also is critical for judging what would be required to implement the intervention in a real-world treatment setting, if it is shown to be efficacious.

In behavioral intervention trials, the interventionist can have a larger effect on the outcomes than the specific intervention being delivered (Miller and Moyers, 2014); thus in addition to

describing the interventionists' general training (e.g., degree, specialization), it is also important to report their training specific to the interventions being delivered in the trial for all conditions (Boutron et al., 2008). Moreover, the number of interventionists within an intervention condition and the number of patients seen by each interventionist are important to report. Interventions that are delivered by a single interventionist are entirely confounded, making it impossible to determine whether the intervention or the interventionist is the active ingredient in treatment efficacy. Conversely, in situations in which multiple interventionists treat numerous patients, clustering of patients within interventionists is important to consider. In such cases, authors should report the intraclass correlation (ICC) for interventionist effects. The ICC provides a statistical index of the percent of outcome variance explained by clustering of individuals within a particular interventionist, where an ICC of 0 would indicate that outcomes are not explained by clustering within interventionist and an ICC closer to 1.0 would indicate that outcomes are largely correlated across individuals within interventionist. If the ICC suggests a considerable effect of therapists and the sample size of the trial is sufficiently large (note that the "sufficiency" of the sample size is based on numerous factors, as discussed by Hoyle and Gottfredson, 2014), then ideally a mixed effects analysis would be reported, with interventionist included as a random effect, should be considered (see Lorentzen et al., 2013 for an example of this approach). For studies with smaller samples or multiple interventionists with only one patient, the number of interventionists and the number of patients seen by each should be reported.

When using a manualized intervention or protocol, it is also important to report on the fidelity of the treatment delivery and competence of the interventionist in delivering the intended interventions (Miller et al., 2005). Quantitative ratings of competence and adherence to the intervention (Carroll et al., 2000; Madson and Campbell, 2006) and whether there was ongoing interventionist monitoring and supervision should also be reported (Carroll and Rounsaville, 2007; Miller et al., 2005).

Describing the assessment procedures

Research protocols, in and of themselves, can also exert effects, with reactivity to assessment significantly affecting engagement in alcohol treatment and outcomes (Clifford et al., 2007; Kaminer et al., 2008; Maisto et al., 2007). Reactivity effects are likely to occur in the active intervention group and the control or placebo conditions, accounting for some improvement across groups and potentially obscuring between-group differences. Improvement across both groups may also be observed due to the phenomenon of regression to the mean (Finney, 2008), which may occur for several reasons. One potential explanation for regression to the mean is the pretreatment or "baseline" assessment occurring when the participant's AUD symptoms and alcohol consumption have increased to the point of prompting him or her to seek treatment. A longer baseline assessment window (reporting retrospectively on the 90 days prior to baseline) could help to provide a better picture of the pretreatment drinking behavior (Gueorguieva et al., 2012; Stasiewicz et al., 2013). Given prior findings on assessment reactivity and the effects of pretreatment drinking, it is critical to report such elements of the trial design as the length of time needed to complete assessments and the time periods covered by assessments.

Reporting of Statistical Methods to Assess Efficacy

Many statistical methods can be employed to assess intervention efficacy, with key determinants of validity being who was included in the analysis of outcomes (e.g., sample size and patient characteristics), which measures were analyzed, and how the analyses were conducted. Ideally, the statistical software syntax or programming code is included in supplementary materials, if allowed by the journal, and made available upon request by the corresponding author. Various programming tools can be used to enhance reproducibility without the need to share the trial data (Liu and Pounds, 2014). We review the relevant reporting aspects of these features in the context of the CONSORT reporting guidelines with particular consideration given to the analysis of drinking outcomes in alcohol clinical trials.

Sample size and baseline differences in patient characteristics

The CONSORT statement recommends that authors report the method for determining sample size, whether *a priori* power analyses were conducted to determine the sample size, and whether/why the actual sample size differs from the intended sample size. The CONSORT diagram shows the number of individuals screened, the number of participants allocated to treatment conditions, and the number of participants who completed the trials. The sample size at each step of participant flow through the trial should be well documented, with sample sizes and reasons for exclusion, and participant dropout from the intervention and/or research follow-up carefully tracked. To allow an evaluation of the efficacy of the intervention, it is critically important to report whether there are differences in baseline characteristics or attrition rates between intervention groups.

Reporting in multisite studies

Many alcohol clinical trials are conducted across multiple sites (often in different states) or multiple treatment programs within a particular research site (Anton et al., 2006; Project MATCH Research Group, 1997). It is important to report the number of sites and the size of the sample within each site because in multisite trials both can influence the power to detect treatment effects (Feaster et al., 2011). The COMBINE study, the largest multisite alcohol pharmacotherapy trial conducted in the United States to date, included 1383 alcohol-dependent individuals recruited from alcohol treatment programs by 11 research teams across the United States. Because nearly all analyses indicated significant site effects, the investigators controlled for site in the primary outcome analyses (Anton et al., 2006).

A discussion of statistical methods that can be used effectively to control for site differences is beyond the scope of this review (see Witkiewitz et al., under review). However, at the very least, authors should provide the reader with an indication of whether site effects were detected and how site was included in the analyses. To quantify site effects on the primary outcome measures, investigators can calculate the ICC (Desai and Begg, 2008). Importantly, as the ICC gets larger, so does the number of participants needed to detect a given effect size. Reporting of ICC (whether at the level of the site, group, or therapist, as described above) is valuable to inform power estimates in subsequent studies. Graphical depictions of site effects, including simple forest plots of the confidence intervals of the treatment effects for each site are also helpful.

Outcome measure description

CONSORT standards require clearly defined primary and secondary outcome measures and, when applicable, a description of any methods used to enhance the quality of measurements (e.g., multiple observations, training of assessors). The standards also require specification of the primary time point for evaluating efficacy (e.g., end of active treatment intervention vs. a more remote follow-up assessment). The reporting of primary and secondary outcomes should be consistent with the analysis plan, as stated in the trial registration.

In most alcohol treatment studies, the primary and/or secondary outcome measures are likely to be based on alcohol consumption, often assessed via self-report. The current "gold standard" method to obtain self-reported alcohol consumption data is via a calendar method such as the Timeline Follow-back method (TLFB; Sobell and Sobell, 1992) or Form-90 (Miller, 1996). These methods provide estimates of daily alcohol consumption in a pre-selected standard drink unit and can yield reliable and valid estimates of alcohol consumption via a variety of assessment modalities (e.g., online, mobile, phone) in numerous studies (e.g., Bernhardt et al., 2009; Pedersen et al., 2012; Sobell et al., 1996).

Estimated alcohol consumption at the daily level is often aggregated to provide various estimates of drinking outcomes, such as percentage of days abstinent, drinks per drinking day, heavy drinking days (where "heavy" drinking is defined as four or more drinks for women and five or more drinks for men), or drinks per day (Anton et al., 2006; Project MATCH Research Group, 1997).

Investigators should report the standard drink unit size and the method of aggregation. The National Institute on Alcohol Abuse and Alcoholism (NIAAA) has recommended that the primary measure be heavy drinking, with the percentage of individuals with no heavy drinking days as the primary endpoint in alcohol clinical trials (Falk et al., 2010; Sobell et al., 2003). A binary measure of no heavy drinking has been proposed as a primary endpoint in the Food and Drug Administration's draft guidance for AUD medications development (Food and Drug Administration, 2015) and the European Medicines Agency (2010) has approved reductions in heavy drinking as a primary endpoint for trials with harm reduction goals. Yet, there is still a lack of consensus among researchers in the field regarding the optimal outcome for treatment studies and the selection of primary and secondary outcomes can depend on the goal of the trial (Donovan et al., 2012). We recommend that, when alcohol consumption is used as an outcome, information be provided on the use of biomarkers of alcohol use, which are commonly used to validate self-report data, along with the correspondence between the biomarker(s) and self-reported consumption measures (Litten et al., 2010). Evidence of efficacy for an intervention that is consistent across self-report measures and biomarker findings adds substantially to the validity of the findings.

When reporting on any measures used in a trial, an assessment of the reliability and validity of the measures is also warranted. Importantly, measures are not inherently reliable or valid (Vacha-Haase, 1998); rather, reliability and validity pertain to a particular administration of a measure. Thus, investigators should report any potential threats to the reliability or assessment validity of the measures in a given study (e.g., whether results were exclusively

for research or became part of the medical record), as well as the internal consistency reliability coefficients of the measures from the study sample.

Presentation of data and statistical analyses

Various statistical analyses can be appropriate to identify intervention effects in alcohol clinical trials. However, instead of discussing specific approaches (see Witkiewitz et al., under review), we provide general recommendations for reporting of results and encourage authors to follow the CONSORT recommendation to describe all analyses in enough detail to permit the results to be replicated by a knowledgeable reader who obtains the original datasets.

1. Means and standard deviations or frequencies of primary outcomes, and sample sizes, for each intervention group at each assessment point should be provided—Reporting these features allows for the estimation of effect sizes that can guide future research and indicates whether key variables are normally distributed. The distribution of alcohol treatment outcome measures is often skewed, with a preponderance of zero values (e.g., drinking days among abstainers), which can bias statistical tests that assume a normal distribution. Methods to transform outcome data to achieve a more normal distribution should be reported.

2. Primary analyses and any exploratory analyses should be clearly described and as pre-specified in the clinical trial registry—In the Abstract and the Results section, authors should report the estimate of the intervention effect (defined as the difference on the primary outcome measure between intervention and comparison groups), as well as the effect size and its 95% confidence interval (CI; Kelley and Preacher, 2012). Any deviations from the primary outcome analyses specified in the clinical trial registry should be clearly indicated (Hartung et al., 2014). Importantly, the pre-defined *alpha* should be reported and results that do not meet that level of significance (e.g., $p = 0.07$) should not be interpreted as “trend” level effects. The number needed to treat (NNT), defined as the number of patients that must be treated to prevent one additional negative outcome (e.g., heavy drinking day) or achieve one additional positive outcome (e.g., an abstinent day), and success rate (defined as the probability that an intervention patient compared to a control patient has the intended outcome) are recommended measures of effect size (Kraemer and Kupfer, 2006) and should also be reported with their 95% CI.

For most RCTs, the primary analysis should be based on the original allocation of participants to intervention conditions, such that all individuals initially allocated to an intervention group are included in the analyses as members of that group (i.e., an “intention-to-treat” (ITT) analysis). As noted earlier, this principle has been compromised in some alcohol clinical trials by the use of a “modified” intention-to-treat approach or an incorrectly implemented ITT approach (Del Re et al., 2013). It is critical for authors to report the type of analyses that were conducted with regard to randomization and how allocation was preserved (or not) in the primary analysis of intervention effects. Exploratory analyses, particularly subgroup and adjusted analyses, should be clearly specified. Ideally, exploratory analyses are planned *a priori* based on prior research or theoretical considerations and

reflected in the trial registration. *Post-hoc* analyses (i.e. those proposed after examining the collected data) may lead to spurious findings that are less likely to be replicated (Freemantle, 2001; Pocock et al., 1987). Such findings, when reported, should be identified as having been proposed after data collection.

3. Attrition and methods for handling missing data should be reported— Missing data are, unfortunately, quite common in alcohol treatment trials. Although much has been learned about the factors that contribute to patient dropout from drug and alcohol treatment (Ball et al., 2006; Brorson et al., 2013; Coulson et al., 2009; Palmer et al., 2009), oftentimes the reasons for missing data in both clinical care and clinical trials are unknown. Attrition by treatment group, time-point, and site (if the study is a multi-site trial) should be reported. Any systematic differences in attrition rates (e.g., more attrition among females, or among older patients) should also be reported and variables associated with attrition should be considered as covariates in subsequent analyses (National Research Council, 2010). In addition, the level of adherence to the medication and the methods used to evaluate medication adherence, as well as attendance at behavioral treatment sessions should be reported, with consideration given to their use as covariates (Baros et al., 2007; Stout et al., 2014; Swift et al., 2011).

Regardless of the reasons for attrition, it is critically important to report the method used to address missing data in the analysis of the primary and secondary outcomes. In their review, Del Re and colleagues (2013) found that 23% of alcohol clinical trials did not provide sufficient information to verify the missing data procedure that was used in the analysis of the primary outcome. Of studies that reported the missing data procedure used, complete case analysis, assuming failure (i.e., heavy drinking), and last observation carried forward were the most commonly used approaches (Del Re et al., 2013). Two recent studies of different approaches to handling missing data in alcohol clinical trials, one of which use simulation based on the COMBINE study data (Hallgren and Witkiewitz, 2014) and the other of which used the actual COMBINE study data (Witkiewitz et al., 2014), showed that these approaches produced the most biased estimates of treatment effects, while maximum likelihood estimation and multiple imputation approaches provided the least biased estimates of treatment effects.

Researchers should specify a plan for handling missing data at the time of clinical trial registration. They should then report the method used to handle missing data, acknowledging any variation from the pre-specified plan, which should be based on consultation with the most recent statistical literature to identify the least biased methods to estimate treatment effects in the presence of missing data.

Conclusions and Future Directions

Despite the high prevalence and substantial public health impact of AUDs both in the United States and globally there are relatively few efficacious treatments for AUDs (Jonas et al., 2014; Magill and Ray, 2009; National Institute for Health and Care Excellence, 2011; Vasilaki et al., 2006; Zindel and Kranzler, 2014). Given these findings, it is critical that researchers continue to develop, evaluate, and disseminate efficacious treatments for AUDs.

A major barrier to the evaluation and dissemination of effective treatments is the lack of appropriate reporting of study features from alcohol treatment trials. Notably, the potential risks of bias have been shown to artificially inflate treatment effect sizes in AUD research (Jonas et al., 2014; Ladd et al., 2010), which ultimately could impede dissemination efforts when medications or behavioral interventions prove less effective in real-world settings (Humphreys and Weisner, 2000). Only with more complete and transparent reporting can potential biases in alcohol treatment trials be gauged.

Although the CONSORT statement and its extensions (Boutron et al., 2008; Calvert et al., 2013; Schulz et al., 2011) have generally improved reporting practices for RCTs, much improvement is still needed in trial reporting for all diseases (Turner et al., 2012), including AUDs (Ladd et al., 2010). We have highlighted several areas of particular relevance to AUD RCTs, including eligibility criteria, retention of participants, procedures for random assignment, descriptions of the interventions and interventionists, and several considerations in the reporting of statistical analyses consistent with trial registration. These recommendations for reporting after trial completion are included as a checklist in Table 1. Trial registration and careful documentation of any discrepancies between the analyses planned *a priori* and the completed primary or secondary outcome analyses are critical to evaluate the validity of trial findings and whether *post-hoc* explanations or “fishing expeditions” account for the apparent efficacy of the treatment, although sometimes these explorations lead to a new insight that requires replication. Journal editors and reviewers are often aware of deficits in clinical trial methodology (Harris et al., 2009), but are often unaware of discrepancies that exist between reported study methods and those in the planned protocol in the trial registry (Dwan et al., 2011; Mathieu et al., 2013). Authors could be required to provide a copy of the trial registration with submitted manuscripts, which would provide journal editors and reviewers with the opportunity to evaluate departures from the *a priori* analyses that reflect “cherry-picking” results for publication purposes (Zarin et al., 2011).

We conclude that poor reporting practices and deviations from planned pre-registered protocols for evaluating treatments for AUDs (especially unreported deviations) impede the application of the scientific method to the development of efficacious AUD treatments and degrade the evidence base needed to optimize systematic reviews. We recognize that journal page limits often restrict the amount of information that can be reported. Many journals now allow supplementary materials to be published online and using this outlet to incorporate all recommended reporting elements (both CONSORT and the recommendations in Table 1) should be considered. We also encourage all researchers, journal editors, and reviewers to work to ensure careful reporting practices so that published trial reports contribute meaningfully to the cumulative science of evidence-based treatments for AUDs.

Acknowledgements

This research was supported by grants from the National Institute on Alcohol Abuse and Alcoholism (AA022328 to K. Witkiewitz and AA021164 and AA023192 to H. Kranzler) and by the U.S. Department of Veteran Affairs. The views expressed do not reflect those of the Department of Veteran Affairs or other institutions.

References

Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, Gøtzsche PC, Lang T. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. Ann. Intern. Med. 2001; 134:663–694. [PubMed: 11304107]

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 R, Weiss RD, Williams LD, Zweben A. Combined pharmacotherapies and behavioral interventions for alcohol dependence: the COMBINE study: a randomized controlled trial. JAMA. 2006; 295:2003–17. [PubMed: 16670409]

Ball SA, Carroll KM, Canning-Ball M, Rounsaville BJ. Reasons for dropout from drug abuse treatment: symptoms, personality, and motivation. Addict. Behav. 2006; 31:320–30. [PubMed: 15964152]

Baros AM, Latham PK, Moak DH, Voronin K, Anton RF. What role does measuring medication compliance play in evaluating the efficacy of naltrexone? Alcohol. Clin. Exp. Res. 2007; 31:596–603. [PubMed: 17374038]

Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, Pitkin R, Rennie D, Schulz KF, Simel D, Stroup DF. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. JAMA. 1996; 276:637–639. [PubMed: 8773637]

Berger VW. A review of methods for ensuring the comparability of comparison groups in randomized clinical trials. Rev. Recent Clin. Trials. 2006; 1:81–6. [PubMed: 18393784]

Berger VW, Weinstein S. Ensuring the comparability of comparison groups: is randomization enough? Control. Clin. Trials. 2004; 25:515–24. [PubMed: 15465620]

Bernhardt JM, Usdan S, Mays D, Martin R, Cremeens J, Arriola KJ. Alcohol assessment among college students using wireless mobile technology. J. Stud. Alcohol Drugs. 2009; 70:771–5. [PubMed: 19737502]

Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. Ann. Intern. Med. 2008; 148:295–309. [PubMed: 18283207]

Brigham GS, Feaster DJ, Wakim PG, Dempsey CL. Choosing a control group in effectiveness trials of behavioral drug abuse treatments. J. Subst. Abuse Treat. 2009; 37:388–97. [PubMed: 19553062]

Brorson HH, Ajo Arnevik E, Rand-Hendriksen K, Duckert F. Drop-out from addiction treatment: a systematic review of risk factors. Clin. Psychol. Rev. 2013; 33:1010–24. [PubMed: 24029221]

Calvert M, Blazeby J, Altman DG, Revicki DA, Moher D, Brundage MD. Reporting of patient-reported outcomes in randomized trials: the CONSORT PRO extension. JAMA. 2013; 309:814–22. [PubMed: 23443445]

Carroll KM, Nich C, Sifry RL, Nuro KF, Frankforter TL, Ball SA, Fenton L, Rounsaville BJ. A general system for evaluating therapist adherence and competence in psychotherapy research in the addictions. Drug Alcohol Depend. 2000; 57:225–38. [PubMed: 10661673]

Carroll KM, Rounsaville BJ. Bridging the gap: a hybrid model to link efficacy and effectiveness research in substance abuse treatment. Psychiatr. Serv. 2003; 54:333–9. [PubMed: 12610240]

Carroll KM, Rounsaville BJ. A vision of the next generation of behavioral therapies research in the addictions. Addiction. 2007; 102:850–62. discussion 863–9. [PubMed: 17523974]

Clifford PR, Maisto SA, Davis CM. Alcohol treatment research assessment exposure subject reactivity effects: part I. Alcohol use and related consequences. J. Stud. Alcohol Drugs. 2007; 68:519–28. [PubMed: 17568955]

Coulson C, Ng F, Geertsema M, Dodd S, Berk M. Client-reported reasons for non-engagement in drug and alcohol treatment. Drug Alcohol Rev. 2009; 28:372–8. [PubMed: 19594791]

Del Re AC, Maisel NC, Blodgett JC, Finney JW. Intention-to-treat analyses and missing data approaches in pharmacotherapy trials for alcohol use disorders. BMJ Open. 2013; 3:e003464.

Desai M, Begg MD. A comparison of regression approaches for analyzing clustered data. Am. J. Public Health. 2008; 98:1425–9. [PubMed: 18556621]

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]

Dwan K, Altman DG, Cresswell L, Blundell M, Gamble CL, Williamson PR. Comparison of protocols and registry entries to published reports for randomised controlled trials. Cochrane Database Syst. Rev. 2011 MR000031.

European Medicines Agency. Guideline on the development of medicinal products for the treatment of alcohol dependence. European Medicines Agency; United Kingdom: 2010.

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; 34:2022–34. [PubMed: 20659066]

Feaster DJ, Mikulich-Gilbertson S, Brincks AM. Modeling site effects in the design and analysis of multi-site trials. Am. J. Drug Alcohol Abuse. 2011; 37:383–91. [PubMed: 21854281]

Finney JW. Regression to the mean in substance use disorder treatment research. Addiction. 2008; 103:42–52. [PubMed: 17999707]

Food and Drug Administration. Alcoholism: Developing drugs for treatment (No. FDA D-0152-001). Silver Spring; MD: 2015.

Freemantle N. Interpreting the results of secondary end points and subgroup analyses in clinical trials: should we lock the crazy aunt in the attic? BMJ. 2001; 322:989–91. [PubMed: 11312237]

Gastfriend DR, Donovan D, Lefebvre R, Murray KT. Developing a baseline assessment battery: balancing patient time burden with essential clinical and research monitoring. J. Stud. Alcohol Suppl. 2005:94–103. discussion 92–3. [PubMed: 16223061]

Graham K, Annis HM, Brett PJ, Venesoen P. A controlled field trial of group versus individual cognitive-behavioural training for relapse prevention. Addiction. 1996; 91:1127–39. [PubMed: 8828241]

Greenfield SF, Cummings AM, Kuper LE, Wigderson SB, Koro-Ljungberg M. A qualitative analysis of women's experiences in single-gender versus mixed-gender substance abuse group therapy. Subst. Use Misuse. 2013; 48:750–60. [PubMed: 23607675]

Grossman J, Mackenzie FJ. The randomized controlled trial: gold standard, or merely standard? Perspect. Biol. Med. 2005; 48:516–34. [PubMed: 16227664]

Gueorguieva R, Wu R, Donovan D, Rounsaville BJ, Couper D, Krystal JH, O'Malley SS. Baseline trajectories of heavy drinking and their effects on postrandomization drinking in the COMBINE Study: empirically derived predictors of drinking outcomes during treatment. Alcohol. 2012; 46:121–31. [PubMed: 21925828]

Haahr MT, Hróbjartsson A. Who is blinded in randomized clinical trials? A study of 200 trials and a survey of authors. Clin. Trials. 2006; 3:360–5. [PubMed: 17060210]

Hallgren KA, Witkiewitz K. Missing data in alcohol clinical trials: A comparison of methods. Alcohol. Clin. Exp. Res. 2014; 38

Harris AHS, Reeder R, Hyun JK. Common statistical and research design problems in manuscripts submitted to high-impact psychiatry journals: what editors and reviewers want authors to know. J. Psychiatr. Res. 2009; 43:1231–4. [PubMed: 19435635]

Hartung DM, Zarin DA, Guise J-M, McDonagh M, Paynter R, Helfand M. Reporting discrepancies between the ClinicalTrials.gov results database and peer-reviewed publications. Ann. Intern. Med. 2014; 160:477–83.ClinicalTrials.gov [PubMed: 24687070]

Heather N. The efficacy-effectiveness distinction in trials of alcohol brief intervention. Addict. Sci. Clin. Pract. 2014; 9:13. [PubMed: 25127717]

Herbison P, Hay-Smith J, Gillespie WJ. Different methods of allocation to groups in randomized trials are associated with different levels of bias. A meta-epidemiological study. J. Clin. Epidemiol. 2011; 64:1070–5. [PubMed: 21474279]

Hoertel N, Falissard B, Humphreys K, Gorwood P, Seigneurie A-S, Limosin F. Do clinical trials of treatment of alcohol dependence adequately enroll participants with co-occurring independent mood and anxiety disorders? An analysis of data from the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). J. Clin. Psychiatry. 2014; 75:231–7. [PubMed: 24569017]

Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG, Schulz KF. CONSORT for reporting randomised trials in journal and conference abstracts. Lancet. 2008a; 371:281–3. [PubMed: 18221781]

Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG, Schulz KF. CONSORT for reporting randomized controlled trials in journal and conference abstracts: explanation and elaboration. PLoS Med. 2008b; 5:e20. [PubMed: 18215107]

Hoyle RH, Gottfredson NC. Sample Size Considerations in Prevention Research Applications of Multilevel Modeling and Structural Equation Modeling. Prev. Sci. 2014

Humphreys K, Harris AHS, Weingardt KR. Subject eligibility criteria can substantially influence the results of alcohol-treatment outcome research. J. Stud. Alcohol Drugs. 2008; 69:757–64. [PubMed: 18781251]

Humphreys K, Weingardt KR, Horst D, Joshi AA, Finney JW. Prevalence and predictors of research participant eligibility criteria in alcohol treatment outcome studies, 1970-98. Addiction. 2005; 100:1249–57. [PubMed: 16128714]

Humphreys K, Weisner CM. Use of exclusion criteria in selecting research subjects and its effect on the generalizability of alcohol treatment outcome studies. Am. J. Psychiatry. 2000; 157:588–94. [PubMed: 10739418]

Imel ZE, Wampold BE, Miller SD, Fleming RR. Distinctions without a difference: direct comparisons of psychotherapies for alcohol use disorders. Psychol. Addict. Behav. 2008; 22:533–43. [PubMed: 19071978]

Ioannidis JPA, Evans SJW, Gøtzsche PC, O'Neill RT, Altman DG, Schulz K, Moher D. Better reporting of harms in randomized trials: an extension of the CONSORT statement. Ann. Intern. Med. 2004; 141:781–8. [PubMed: 15545678]

Jonas DE, Amick HR, Feltner C, Bobashev G, Thomas K, Wines R, Kim MM, Shanahan E, Gass CE, Rowe CJ, Garbutt JC. Pharmacotherapy for adults with alcohol use disorders in outpatient settings: a systematic review and meta-analysis. JAMA. 2014; 311:1889–900. [PubMed: 24825644]

Jonas, DE.; Cusack, K.; Forneris, CA.; Wilkins, TM.; Sonis, J.; Middleton, JC.; Feltner, C.; Meredith, D.; Cavanaugh, J.; Brownley, KA.; Olmsted, KR.; Greenblatt, A.; Weil, A.; Gaynes, BN. Comparative Effectiveness Review No. 92. Agency for Healthcare Research and Quality; Rockville, MD: 2013. Psychological and Pharmacological Treatments for Adults With Posttraumatic Stress Disorder (PTSD).

Kaminer Y, Burleson JA, Burke R. Can assessment reactivity predict treatment outcome among adolescents with alcohol and other substance use disorders? Subst. Abus. 2008; 29:63–9. [PubMed: 19042325]

Kelley K, Preacher KJ. On effect size. Psychol. Methods. 2012; 17:137–52. [PubMed: 22545595]

Kraemer HC, Kupfer DJ. Size of treatment effects and their importance to clinical research and practice. Biol. Psychiatry. 2006; 59:990–6. [PubMed: 16368078]

Ladd BO, McCrady BS, Manuel JK, Campbell W. Improving the quality of reporting alcohol outcome studies: effects of the CONSORT statement. Addict. Behav. 2010; 35:660–6. [PubMed: 20207490]

Litten RZ, Bradley AM, Moss HB. Alcohol biomarkers in applied settings: recent advances and future research opportunities. Alcohol. Clin. Exp. Res. 2010; 34:955–67. [PubMed: 20374219]

Liu Z, Pounds S. An R package that automatically collects and archives details for reproducible computing. BMC Bioinformatics. 2014; 15:138. [PubMed: 24886202]

Lorentzen S, Ruud T, Fjeldstad A, Høglend P. Comparison of short- and long-term dynamic group psychotherapy: randomised clinical trial. Br. J. Psychiatry. 2013; 203:280–7. [PubMed: 24029539]

Madson MB, Campbell TC. Measures of fidelity in motivational enhancement: a systematic review. J. Subst. Abuse Treat. 2006; 31:67–73. [PubMed: 16814012]

Magill M, Ray LA. Cognitive-behavioral treatment with adult alcohol and illicit drug users: a meta-analysis of randomized controlled trials. J. Stud. Alcohol Drugs. 2009; 70:516–27. [PubMed: 19515291]

Maisto SA, Clifford PR, Davis CM. Alcohol treatment research assessment exposure subject reactivity effects: part II. Treatment engagement and involvement. J. Stud. Alcohol Drugs. 2007; 68:529–33. [PubMed: 17568956]

Marques AC, Formigoni ML. Comparison of individual and group cognitive-behavioral therapy for alcohol and/or drug-dependent patients. Addiction. 2001; 96:835–46. [PubMed: 11399215]

Mathieu S, Chan A-W, Ravaud P. Use of trial register information during the peer review process. PLoS One. 2013; 8:e59910. [PubMed: 23593154]

Miller, WR. Form 90: A structured assessment interview for drinking and related behaviors. Project, MA., editor. National Institute on Alcohol Abuse and Alcoholism; Bethesda, MD: 1996.

Miller WR, Moyers TB. The forest and the trees: relational and specific factors in addiction treatment. Addiction. 2014

Miller WR, Moyers TB, Arciniega L, Ernst D, Forcehimes A. Training, supervision and quality monitoring of the COMBINE Study behavioral interventions. J. Stud. Alcohol. Suppl. 2005:188–195. [PubMed: 16223070]

Moher D, Schulz KF, Altman DG, the CONSORT Group. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. Ann. Intern. Med. 2001; 134:657–662. [PubMed: 11304106]

Moyers TB, Miller WR. Is low therapist empathy toxic? Psychol. Addict. Behav. 2013; 27:878–84. [PubMed: 23025709]

National Institute for Health and Care Excellence. Alcohol-use disorders: Diagnosis, assessment and management of harmful drinking and alcohol dependence. National Institute for Health and Care Excellence; Manchester, UK: 2011.

National Research Council. The prevention and treatment of missing data in clinical trials. The National Academies Press; Washington, DC, US: 2010.

Palmer RS, Murphy MK, Piselli A, Ball SA. Substance user treatment dropout from client and clinician perspectives: a pilot study. Subst. Use Misuse. 2009; 44:1021–38. [PubMed: 19938942]

Pedersen ER, Grow J, Duncan S, Neighbors C, Larimer ME. Concurrent validity of an online version of the Timeline Followback assessment. Psychol. Addict. Behav. 2012; 26:672–7. [PubMed: 22486334]

Pocock SJ, Hughes MD, Lee RJ. Statistical problems in the reporting of clinical trials. A survey of three medical journals. N. Engl. J. Med. 1987; 317:426–32. [PubMed: 3614286]

Project MATCH Research Group. Matching Alcoholism Treatments to Client Heterogeneity: Project MATCH posttreatment drinking outcomes. J. Stud. Alcohol. 1997; 58:7–29. [PubMed: 8979210]

Public Law 110-95. Food and Drug Administration Amendments Act of 2007. 2007.

Savović J, Jones HE, Altman DG, Harris RJ, Jüni P, Pildal J, Als-Nielsen B, Balk EM, Gluud C, Gluud LL, Ioannidis JPA, Schulz KF, Beynon R, Welton NJ, Wood L, Moher D, Deeks JJ, Sterne JAC. Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. Ann. Intern. Med. 2012; 157:429–38. [PubMed: 22945832]

Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. Int. J. Surg. 2011; 9:672–677. [PubMed: 22019563]

Schulz KF, Altman DG, Moher D, the CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. PLoS Med. 2010; 7:e1000251. [PubMed: 20352064]

Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA. 1995; 273:408–12. [PubMed: 7823387]

Sobell LC, Brown J, Leo GI, Sobell MB. The reliability of the Alcohol Timeline Followback when administered by telephone and by computer. Drug Alcohol Depend. 1996; 42:49–54. [PubMed: 8889403]

Sobell, LC.; Sobell, MB. Timeline Follow-Back: A technique for assessing self-reported alcohol consumption. Human Press; Totowa, NJ, US: 1992.

Sobell LC, Sobell MB, Agrawal S. Randomized controlled trial of a cognitive-behavioral motivational intervention in a group versus individual format for substance use disorders. Psychol. Addict. Behav. 2009; 23:672–83. [PubMed: 20025373]

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]

Stasiewicz PR, Schlauch RC, Bradizza CM, Bole CW, Coffey SF. Pretreatment changes in drinking: relationship to treatment outcomes. Psychol. Addict. Behav. 2013; 27:1159–66. [PubMed: 23276313]

Stout RL, Braciszewski JM, Subbaraman MS, Kranzler HR, O'Malley SS, Falk D. What happens when people discontinue taking medications? Lessons from COMBINE. Addiction. 2014; 109:2044–52. [PubMed: 25098969]

Swearingen CE, Moyer A, Finney JW. Alcoholism treatment outcome studies, 1970-1998. An expanded look at the nature of the research. Addict. Behav. 2003; 28:415–36. [PubMed: 12628616]

Swift R, Oslin DW, Alexander M, Forman R. Adherence monitoring in naltrexone pharmacotherapy trials: a systematic review. J. Stud. Alcohol Drugs. 2011; 72:1012–8. [PubMed: 22051215]

To MJ, Jones J, Emara M, Jadad AR. Are reports of randomized controlled trials improving over time? A systematic review of 284 articles published in high-impact general and specialized medical journals. PLoS One. 2013; 8:e84779. [PubMed: 24391973]

Turner L, Shamseer L, Altman DG, Weeks L, Peters J, Kober T, Dias S, Schulz KF, Plint AC, Moher D. Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals. Cochrane Database Syst. Rev. 2012; 11 MR000030.

Vacha-Haase T. Reliability generalization: Exploring variance in measurement error affecting score reliability across studies. Educ. Psychol. Meas. 1998; 58:6–20.

Vasilaki EI, Hosier SG, Cox WM. The efficacy of motivational interviewing as a brief intervention for excessive drinking: a meta-analytic review. Alcohol Alcohol. 2006; 41:328–35. [PubMed: 16547122]

Viera AJ, Bangdiwala SI. Eliminating bias in randomized controlled trials: importance of allocation concealment and masking. Fam. Med. 2007; 39:132–7. [PubMed: 17273956]

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; 38:2826–34. [PubMed: 25421518]

Witkiewitz, K.; Finney, J.; Harris, A.; Kivlahan, D.; Kranzler, H. Recommendations for the design and analysis of treatment trials for alcohol use disorders. under review

Wood L, Egger M, Gluud LL, Schulz KF, Jüni P, Altman DG, Gluud C, Martin RM, Wood AJG, Sterne JAC. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. BMJ. 2008; 336:601–5. [PubMed: 18316340]

Youngblood ME, Murray KT, Devine E, Latham PK, Hubatch S. Coordinating and monitoring multisite clinical trials that combine pharmacological and behavioral treatments. J. Stud. Alcohol. Suppl. 2005:82–91. discussion 65. [PubMed: 16223060]

Zarin DA, Tse T. Medicine. Moving Toward Transparency of Clinical Trials. Science. 2008; 319:1340–1342. (80-.). [PubMed: 18323436]

Zarin DA, Tse T, Sheehan J. The Proposed Rule for U.S. Clinical Trial Registration and Results Submission. N. Engl. J. Med. 2015; 372:174–180. [PubMed: 25539444]

Zarin DA, Tse T, Williams RJ, Califf RM, Ide NC. The ClinicalTrials.gov results database--update and key issues. N. Engl. J. Med. 2011; 364:852–60.ClinicalTrials.gov [PubMed: 21366476]

Zindel LR, Kranzler HR. Pharmacotherapy of alcohol use disorders: seventy-five years of progress. J. Stud. Alcohol Drugs. 2014; 75(Suppl 1):79–88.

Zweben A, Barrett D, Berger L, Murray KT. Recruiting and retaining participants in a combined behavioral and pharmacological clinical trial. J. Stud. Alcohol Suppl. 2005:72–81. discussion 65. [PubMed: 16223059]

Zweben A, Fucito LM, O'Malley SS. Effective Strategies for Maintaining Research Participation in Clinical Trials. Drug Inf. J. 2009; 43

Table 1

Checklist for Reporting Results of Alcohol Treatment Trials

Abstract – Description of trial design, participants (including sample size and demographics), main findings (including effect sizes and/or 95% confidence intervals), and conclusions from the study
Introduction – Description of background for the study, why the current study is necessary in the context of existing research, and the specific objectives and hypotheses of the current study
Methods – Participants
- Participant selection procedures, how sample size was determined, and setting of recruitment reported
- Detailed description of inclusion and exclusion criteria, rationale for inclusion and exclusion criteria, and the numbers excluded and reasons for exclusion in the sample
- Participant demographic and clinical characteristics reported, for entire sample and by intervention groups. Differences in demographic or clinical characteristics across intervention groups reported.
- Provide CONSORT Flow Diagram to characterize retention in the study and reasons for dropping out
- Report numbers experiencing adverse events and type and severity of adverse events
Methods – Randomization
- Follow CONSORT guidance for reporting on method used to generate the randomization sequence, the type of randomization, the allocation concealment method, and randomization implementation methods
- If study is blinded then report of who was blinded and if and when the blind was broken
Methods – Descriptions of Interventions and Procedures
- For pharmacological interventions, report the drug name, dose, method of administration, titration procedures, timing and duration of administration, and adherence monitoring methods
- For non-pharmacological interventions, report specific details of the experimental and comparator conditions, including a description of type and format of intervention, and details regarding standardization of the interventions and adherence to the intervention protocols
- For all trial types describe the degree and type of interactions with any health professionals related to the trial, including number of interventionists, training of interventionists, and number of patients seen by each interventionist. Report ICCs primary and secondary outcomes by interventionist
- Describe the assessment procedures, including duration, timing and location of all research assessments
- In multi-site studies provide a description of the number of sites and characteristics of sites, indicate whether site effects were detected and how site was included in analyses. Report ICCs for primary and secondary outcomes by site
- Define primary and secondary outcomes, methods used to enhance quality of outcome assessment, and time points at which outcomes were assessed. Outcome selection is based on the trial registration. Reliability and validity of assessment of primary and secondary outcomes reported.
Methods and Results – Statistical Analyses
- Means and standard deviations or frequencies of primary and secondary outcomes, and sample sizes, for each intervention group at each assessment point reported
- Primary analyses and any exploratory analyses clearly described and as pre-specified in the clinical trial registry. Any deviations from the primary outcome analyses specified in the clinical trial registry clearly described. If randomized design report how allocation was preserved in the primary analysis

- Report the estimate of the intervention effect (defined as the difference on the primary outcome measure between intervention and comparison groups), as well as the effect size and its 95% confidence interval
- Clearly describe attrition and methods for handling missing data. Adherence to intervention reported and considered in the analyses of intervention effects.