NIH Public Access
Author Manuscript
Drug Alcohol Depend. Author manuscript; available in PMC 2013 May 11.

Published in final edited form as:
Drug Alcohol Depend. 2002 October 1; 68(2): 121–130.

'Intention-to-treat' meets 'missing data': implications of alternate strategies for analyzing clinical trials data

Charla Nich* and **Kathleen M. Carroll**
Department of Psychiatry, Division of Substance Abuse, VA CT healthcare Center (1510), Yale University, 950 Campbell Avenue, West Haven, CT 06515, USA

Abstract

True intention-to-treat analyses are rare in reports of randomized clinical trials. To highlight the complex issues that arise in conducting and interpreting data from intention-to-treat analyses in studies with substantial levels of protocol violation (e.g. attrition, noncompliance, or withdrawal of participants), data from a clinical trial of treatment for cocaine dependence were analyzed using three strategies to manage missing data: Strategy 1 addressed the effectiveness of treatments based on data collected from participants up to the point of dropout. Strategy 2 addressed the effectiveness of treatments based on data from the full intended duration of the protocol including data collected after participant dropout. The third strategy used a more novel approach, which used an intention-to-treat strategy for the full duration of the trial and the full sample, but also evaluated the effect of treatment retention outcomes by including an independent variable to reflect active treatment retention as a time-varying covariate. Conclusions about the relative efficacy of the study treatments varied to some extent depending on the analytic strategy used. These findings suggest that investigators should make every effort to conduct intent-to-treat analyses, but also to make use of multiple analytic strategies to fully understand the effects of the treatments studied. Moreover, regardless of the strategy used, investigators should clearly describe their handling of data from participants who violate the protocol.

Keywords

Research design; Randomized controlled trial; Attrition; Missing data

1. Introduction

1.1. Attrition and noncompliance in randomized clinical trials

Few clinical trials collect all outcomes on schedule for all study participants randomized to protocol treatments, and this is a particular problem in drug dependence, where attrition and noncompliance rates are comparatively high (Edwards and Rollnick, 1997; Howard et al., 1990; Mattson et al., 1998). A randomized clinical trial typically has as its final data set an incomplete block of data, with some participants not contributing any outcome data, some missing a few data points, and others contributing a full set of data (Wothke, 2000). The problem of missing data is common but is often ignored in analysis and presentation of trial findings due in part to lack of widely accepted strategies for handling the problem (Figueredo et al., 2000).

© 2002 Elsevier Science Ireland Ltd. All rights reserved.

*Corresponding author: Tel.: +1-203-932-3486x7405; fax: +1-203-932-3472, charla.nich@yale.edu (C. Nich).

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

1.2. Problems associated with missing data

An important, but under recognized, problem is the limitations that missing data place on how data from a clinical trial may be analyzed. For example, in a longitudinal study where multiple data points are collected on each participant over time, traditional repeated measures ANOVA-based models require all participants to have all data points collected on the same schedule. Thus, if a participant misses even a single data point, the investigator must then choose either to omit that participant from the dataset (listwise deletion), or to impute values for the missing data points in order to include all participants. In fact, any summary outcome measure, regardless of the analytic strategy, is affected by the missing data and thus subject to bias by being more closely linked to retention than the true outcome. For example, the outcomes 'days achieving target behavior' or 'percentage of days of symptom reduction' can be computed for any participant regardless of protocol compliance, but are nevertheless strongly influenced by missing data (and thus biased) if the participant drops out of treatment, does not fully comply with treatment, or otherwise violates the protocol.

1.3. Intention-to-treat analyses as the gold standard to address missing data

Forty years ago, the term 'intention-to-treat' was used to describe the principle of analyzing data from all participants randomized to treatment, regardless of their level of treatment received or protocol adherence (Hill, 1961). The intention-to-treat principle requires that all study participants be included in the analyses, regardless of whether the participant received any exposure to the assigned study treatment or complied with the treatment. In recent years, the importance of collecting full datasets from all participants randomized and analyzing all data during the time period of interest has been increasingly emphasized (Meyers, 1999; Feinstein, 1991; Lavori, 1992; Lee et al., 1991).

The intention-to-treat principle is critical because it results in unbiased and consistent interpretation of treatment effects, while analyses based on compliant subsamples are invariably biased. Participants who are randomized to treatment but who drop out (or otherwise violate the protocol) are likely to differ from participants who contribute complete data, thus distorting interpretation of the effectiveness of protocol treatments when data from protocol violators are omitted from the analysis. For example, one recent meta-analysis of medication trials for depression reported improvement rates of 50% when analyzing outcomes from all randomized participants, compared to the published improvement rates of 63–73% that did not follow the intention-to-treat principle (Bollini et al., 1999). This difference, which has considerable clinical implications, was attributed to missing data secondary to protocol attrition.

1.4. Rarity of true intention-to-treat analyses

Although intent-to-treat analysis is the undisputed 'gold standard' in reporting results of clinical trials, true intent-to-treat analyses are rare. Recent reports have found unsatisfactory rates of intention-to-treat reporting in dermatology (Adetugbo and Williams, 2000) (with 6% reporting outcomes based on the full intention-to-treat sample), obstetrics (Schultz et al., 1996) (8%), ophthalmology (Scherer and Crowley, 1998) (14%) and major medical journals (Hollis and Campbell, 1999; Ruiz-Canela et al., 2000) (48 and 47.7%, respectively). In psychiatry and substance abuse, a recent report by Ladoucuer et al. (2001) found that only half of all treatment studies described in five well-regarded journals even mentioned dropouts and only 18% included dropouts in statistical analyses.

However, implementation of the intention-to-treat principle in published reports is likely to improve in the near future. In particular, the CONSORT (Consolidated Standards of Reporting Trials) statement (Begg et al., 1996; Moher et al., 2001) represents a major step

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

forward in encouraging more consistent and thorough reporting of results from randomized clinical trials and thus in wider adoption of the intent-to-treat principle. The CONSORT checklist consists of 22 items that evaluate the internal and external validity of a trial as well as the nature and magnitude of selection bias. These items include number and reasons for participants not randomized, number of participants not receiving their assigned intervention, number of participants followed and not followed for assessment of outcome, number of participants withdrawn due to loss to follow-up or adverse reaction to study treatments, and numbers of participants completing the trial. Other reports (Meinert, 1998) have called for even more meticulous reporting of all key trial events, including numbers of participants in each condition receiving the full course of treatment, receiving a partial course of treatment, receiving no treatment, number of missed visits, dropouts, number of morbid events, number of deaths, and number of persons with whereabouts unknown. The FDA requires full intention-to-treat analyses for medication approval (Department of Health and Human Services, 1996).

Among the chief reasons for disparity between the intention-to-treat principle and its practice is the difficulty of tracking and collecting data from participants who violate the protocol, particularly among substance abusers. However, a number of studies have achieved high follow-up rates in clinical trials with drug users and a number of useful guides are available (Cottler et al., 1996; Twitchell et al., 1992; Zweben et al., 1998).

1.5. Methodological issues in analyzing data from protocol violators

What is much less clear, and the subject of this report, is how data that have been collected from protocol violators (e.g. post-dropout or noncompliance) should be analyzed (Lavori et al., 1999). For example, in a ‘classic’ intention-to-treat analysis, all data from all participants would be collected for the full duration of the treatment protocol and used to evaluate treatment effects. Thus, data collected after protocol violation (or data collected from randomized participants who never receive any protocol treatment) would be handled identically to those data collected from participants who were fully compliant with and completed treatment. While such an analysis would avoid the serious problems associated with attrition bias (Feinstein, 1979), it would not facilitate understanding of other key factors that may affect participant outcomes. For example, dropouts, or participants withdrawn from study treatments due to adverse events are likely to seek out and receive other types of non-protocol treatment that may have a profound effect on their symptoms, making it difficult to attribute outcomes to the study treatments themselves. In other words, classic intention-to-treat analyses typically do not allow for the complex methodological issues that arise when interpreting data from the individuals who violate study protocols.

Recent statistical advances have introduced likelihood-based models such as random effects regression models (Gibbons et al., 1988). These models, also known as hierarchical linear models (Bryk and Raudenbush, 1992), are iterative methods that utilize all existing data, both on the individual and group level, to estimate outcomes over time (analogous to a standard regression on each participant). The widespread availability of these methods of analysis in existing statistical packages (e.g. SAS PROC MIXED, BMDP 5V), as well as standalone software (e.g. MIXREG, HLM) has enhanced the opportunities for researchers to approach the intention-to-treat principle in clinical trials by allowing them to interpolate missing values, rather than delete participants with missing values from the analysis of treatment effects.

To address these issues, we have developed a novel analytic strategy that acknowledges the complexities of participants’ behavior during a treatment protocol and allows for data collected after the point of protocol violation to be included in the analysis. This strategy capitalizes on these statistical methods to augment classic intention-to-treat analyses by

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

including a 'dummy' variable to represent whether the data point was gathered prior to or after the point of protocol violation, and including that variable as a predictor (e.g. a time-varying covariate). Unlike using a fixed covariate that would provide a single overall effect estimate (such as gender), including a time-varying covariate that represents protocol compliance (e.g. whether the participant was in treatment or had dropped out at the time data were collected) allows for much fuller understanding of the effects of protocol compliance on treatment effectiveness. Thus, treatment effectiveness during the *full intended duration of the protocol* (classic intention-to-treat analyses) can be compared with the effectiveness of treatments *while participants are compliant with them*. This strategy is flexible enough to account for the complexity of a range of participant behaviors at the point of protocol violation. Thus, a major advantage of this approach is that it allows determination of whether participants' outcomes change at the point of protocol violation and if the nature of those events differs across study treatments, assuming the incomplete data are missing at random.

Data from a randomized clinical trial of treatments for cocaine dependence (Carroll et al., 1998) will be used to demonstrate the potential advantages of this approach, using a simplified case of considering the effects of protocol attrition (although the model could be used for medication noncompliance and other forms of protocol violation as well). Differences in findings that emerge from the following analytic strategies will be highlighted: (1) 'Standard' analysis or evaluation of protocol treatments based only on data collected prior to treatment dropout, with interpolation of missing values for treatment dropouts through random effect regression (e.g. no missing participants, interpolation of missing data). (2) A classic intention-to-treat analysis, sometimes called 'full analysis set', that included data collected from participants after they dropped out of treatment (e.g. no missing participants, no missing data) to evaluate how treatments differed during the full intended time span of treatment. (3) An intent-to-treat analysis using the novel strategy described above which used the full analysis set but also includes each participant's dropout status as a time-varying covariate, thus taking into account whether each outcome was collected before or after the point of attrition (e.g. no missing participants, no missing data, with inclusion of 'dropout status' as a time-varying covariate).

2. Methods

Because this report is intended as a practical demonstration of the implications of alternate analytic approaches to intent-to-treat analyses rather than as a report of a clinical trial, the protocol will be only briefly summarized below.

2.1. Overview of trial and methods

After provision of informed consent, 122 cocaine- and alcohol-dependent individuals were randomly assigned to one of five treatment conditions: clinical management (CM) with disulfiram, 12 step facilitation (TSF) with disulfiram, cognitive-behavioral therapy (CBT) with disulfiram, TSF with no medication, and CBT with no medication (Carroll et al., 1998). To focus on the analysis methods rather than the treatments themselves, we will simplify the treatment designations for this report as follows: CBT = PSY1, TSF = PSY2, CM = PSYCONTROL, disulfiram = MED, and no medication = NOMED.

Participants assigned to MED received 250 mg of disulfiram per day, and compliance was monitored using a riboflavin monitoring procedure (DelBoca et al., 1996). All psychotherapies were manual guided (Carroll, 1998; Fawcett et al., 1987; Nowinski et al., 1992) and delivered in individual sessions offered over a 12-week course of treatment. Process analyses of videotapes of 741 treatment sessions indicated the study treatments were highly discriminable and delivered in accordance with manual guidelines (Carroll et al.,

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

2000). Participants also met weekly with an independent clinical evaluator who collected urine specimens for toxicology screens and self-reports of substance use.

2.2. Assessments

Assessments were administered before treatment, weekly during treatment, and at termination of treatment (12 weeks) by an independent clinical evaluator. As part of the informed consent procedure, prospective participants were informed that the treatment and data collection processes were independent and that assessments would take place regardless of their level of participation in study treatments. Thus, when a participant dropped out of treatment, whenever possible, they were tracked and interviewed to 12 weeks after they began treatment in order to collect information on their daily drug use and treatment utilization during the period between treatment dropout and the interview. For the analyses described here, we focus on a single outcome measure, frequency of cocaine use (the number of days in the preceding week the participant reported using cocaine, confirmed by urinalysis). Of the 122 participants randomized, 117 initiated treatment, and 39 completed treatment. Thirty-one of the treatment completers provided complete data for the 84 days of treatment. Of the 83 participants who did not complete treatment, 54 (65%) were successfully tracked and interviewed 12 weeks after they were randomized.

2.3. Data analysis

t-Tests and χ^2-tests were used to compare: (1) the sample of participants contacted after treatment dropout (early terminator sample) to participants who dropped out but were not assessed at 12 weeks, and (2) the sample contacted after treatment dropout to the sample of participants who completed treatment. Simple random effect regression analysis models, which use only the intercept as a random term, were run using MIXREG software (Hedeker, 1993) to evaluate outcomes using the three strategies (e.g. within-treatment data versus all available data). To account for the greater rate of change earlier in treatment, time was represented in logarithmic form (log time+1). For consistency with previous reports (Carroll et al., 1998, 2000), identical treatment contrasts were evaluated here using contrast coding for the medication effect and orthogonal coding for the psychotherapy effects: (1) effects for time, (2) ‘active psychotherapy’ versus ‘control psychotherapy’ (PSY1+PSY2 versus PSYCONTROL) by time, (2) CBT versus TSF (PSY1 versus PSY2) by time, and (3) disulfiram versus no medication (MED versus NOMED) by time. Also included in each model is the main effect of psychotherapy with contrasts 1 and 2 (PSYCH1, PSYCH2), the main effect of medication (MED), and the interaction of the psychotherapy contrast 2 and medication contrast.

As noted above, to account for the effect of treatment status, we computed a dummy variable to represent whether each data point was collected at a time when the participant was receiving their assigned study treatment or had dropped out. The dummy variable, representing treatment status (e.g. still in treatment versus dropped out) was then used as a predictor of outcome. For the final analysis, we computed interactions of the treatment contrasts with the treatment status variable.

Fig. 1 illustrates the differences among these analytic strategies. A circle represents each outcome (days of cocaine use). Closed circles in Fig. 1a represents data points gathered while the participant was in treatment (Strategy 1). Fig. 1b represents the data used in Strategy 2, which includes all data collected for the full intended duration of the protocol. Fig. 1c also includes all data gathered for the duration of the protocol, but, as in Strategy 3, the data gathered after the participant dropped out of treatment is handled differently, in this case with open circles.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

3. Results

3.1. Representativeness of the samples

Treatment dropouts who provided data at 12 weeks (n=54) did not differ significantly from the treatment dropouts who did not provide data (n=29) with respect to demographic characteristics, including age, education, gender, ethnicity, or socioeconomic status. Also, there were no significant differences in route of administration, age of first cocaine use, or frequency of cocaine or alcohol use at baseline. Treatment dropouts who did and did not provide 12-week data also did not differ significantly in the mean number of treatment sessions attended (e.g. compliance) or the percentage of treatment days abstinent from cocaine or other substances (e.g. outcome) while in treatment.

As expected, participants who provided data after dropping out (n=54) differed from those who completed treatment (n=39) in several ways. At baseline, treatment dropouts were more likely to have a DSM-IV Axis II disorder (68 versus 42%, χ^2=5.96, P<0.05). Treatment completers were also significantly more likely to have been assigned to MED rather than NOMED compared with the treatment dropouts who provided data (77 versus 52%, χ^2=6.06, P<0.01). During treatment, completers had a significantly higher rate of drug-free urine toxicology screens (62 versus 35%, χ^2= 10.05, P<0.05) and a higher rate of compliance with study medication (75.9 versus 58.7%, F=5.72, df=1, 44, P<0.05).

3.2. Comparison of analytic strategies

3.2.1. Strategy 1: standard analysis using random effect regression model on within-treatment data only (data collected prior to the point of treatment dropout or treatment completion)—Random effect regression analyses on all within-treatment data (122 participants, 903 data points) indicated the following: first, the effect for time was significant, indicating a general reduction in cocaine use across time for the sample as a whole (cocaine frequency by time: z=12.5, P<0.001). Second, participants assigned to either PSY1 or PSY2 reduced their cocaine use more over time with respect to those assigned to PSYCONTROL (z=2.01, P<0.05). Third, no significant differences were found between the medication conditions (MED versus NOMED) or the two types of psychotherapies (PSY1 versus PSY2), nor were there any interactions of medication and psychotherapy condition.

3.2.2. Strategy 2: intention-to-treat analysis (assesses differences in treatment effectiveness for the full intended duration of the protocol)—Analyses that included all available data points (122 participants, 1308 data points) concurred with those of the within-treatment data (122 participants, 903 data points), in indicating an overall reduction in frequency of cocaine use over time (z=7.6, P<0.05), as well as a significantly greater reduction in cocaine use for participants in the two experimental psychotherapies (PSY2 and PSY1) compared with participants in PSYCON-TROL (z=2.7, P<0.05). In addition, two significant effects not seen in the earlier model emerged. First, participants assigned to MED reduced their cocaine use significantly more over time than participants assigned to NOMED (z=4.2, P<0.05). Second, individuals assigned to PSY2 reduced their cocaine use significantly more over time than those assigned to PSY1 (z=2.78, P<0.05).

3.2.2.1. How well did the random effect regression model represent participants' actual cocaine use after treatment dropout?: Given that conclusions about study treatments differed based on whether data collected after treatment dropout were included in the two analytic strategies described above, we evaluated how well the random effect regression model used in Strategy 1 predicted the participants' cocaine use after treatment dropout by comparing their true values with those estimated by the model. As shown in Fig. 2, while all

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

true values after 5 weeks (the point by which most of the attrition occurred) fall within one standard error of the estimate, the model did tend to underestimate frequency of cocaine use for the treatment dropouts.

3.2.3. Strategy 3: intention-to-treat analysis with 'treatment status' included as a time-varying covariate (assess treatment differences during intended duration of the protocol as well as the effect of treatment dropout)—To evaluate whether the change in participants' cocaine use after treatment dropout would effect the analysis of treatment effects, a time-varying covariate, 'treatment status' (a dichotomous indicator of whether the outcome came from a participant while in treatment or after treatment drop out, coded −1 and 1, respectively) was added to the statistical model used in the intention-to-treat analysis above (Strategy 2). In other words, each data point indicating the frequency of the participant's cocaine use in the past week had its own covariate, treatment status. This analysis suggested that participants' cocaine use after treatment dropout was significantly greater than their cocaine use while they were in treatment (z=8.6, P <0.05). As shown in Table 1, results of this analysis were consistent with those of Strategy 2 in suggesting that that the sample as a whole reduced cocaine use over time (z=9.63, P <0.05), and that those assigned to PSY2 reduced their cocaine use more than individuals assigned to PSY1 (z=3.02, P < 0.05). However, these findings were inconsistent with those of the classic intention-to-treat analysis (Strategy 2), in that neither the effect of MED over NOMED nor that of PSY1 or PSY2 compared with PSYCONTROL were statistically significant.

3.2.3.1. How did Strategies 2 and 3 differ?: To evaluate possible reasons for the differences in findings between Strategies 2 and 3, which differed only in the inclusion of the time-varying covariate of 'treatment status', we explored how participants' behavior changed before and after they left treatment, and specifically whether there were differences in the intercept or slope by treatment assignment.

A significant MED by Time effect seen in Strategy 2 was not seen in Strategy 3. Strategy 3 has three significant interactions that were not modeled in Strategy 2: Medication by PSY2, Treatment Status by Medication by PSY2, and Treatment Status by Medication by PSY2 by time. Inspection of the data reveals that PSY2 participants on MED have lower scores than PSY2 not on meds, and this is not apparent for PSY1 (MED by PSY2 effect). Also, while in treatment PSY2 participants on MED have lower scores than out of treatment PSY2. For participants not in treatment, those in PSY2 MED have lower scores than those in PSY2 NOMED (Treatment Status by MED by PSY2 effect). Finally, at the point of termination, unlike other treatment groups, participants in PSY2/MED have a *lower* frequency of cocaine use than those in PSY2/MED who do not drop out. Participants who drop out of PSY2/MED increase cocaine use over time (Treatment Status by MED by PSY2 by Time effect).

There was also a significant effect of the active psychotherapies compared with the control by time in Strategy 2 is not seen in Strategy 3. This appears to be associated with an effect of PSYCH1 by Treatment Status, and PSYCH1 by Treatment Status by Time in Strategy 3. The first effect, PSYCH1 by time, is produced by the PSYCONTROL participants at the point of dropout, having higher scores than PSY2/PSY1 participants. At the termination point, however, the slope for dropouts in PSYCONTROL changes to an upward slope, indicating an increase in drug use, while the rate of change for the PSY2/PSY1 participants becomes flat, indicating no change in drug use from the point of dropout. Conversely, for participants who stay in treatment, the rate of change for participants in PSY1/PSY2 is greater than that of those who stay in PSYCONTROL.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Two additional effects are seen in Strategy 3 that were not assessed in Strategy 2. There is a Treatment Status by PSYCH2 effect, which suggests that PSY2 participants were using less cocaine at the drop out point than PSY1 dropouts. In addition, the Treatment Status by PSYCH2 by Time effect suggests that participants in PSY1 who do not drop out of treatment have the greatest rate of change.

3.2.3.2. Why did Strategies 2 and 3 differ?: Regarding the finding of a significant effect for MED in Strategy 2 but not Strategy 3, as noted above earlier, participants assigned to MED were significantly more likely to complete treatment than those not assigned to MED (38.5 versus 20.5%, $\chi^2=4.2$, $P<0.05$). *Thus, because participants' cocaine use increased significantly at the point of treatment dropout, when the variance associated with treatment participation is accounted for in Strategy 3 through the time-varying covariate, the medication effect suggested in Strategy 2 is no longer significant. In other words, while the overall effect of MED is fairly robust as suggested by the results of Strategy 2, it is likely that the medication itself, represented by the time-varying covariate in the model, is effective only while the participant is still involved in treatment and taking it.*

While both Strategy 1 and 2 found an effect for PSYCH1 by Time, Strategy 3 did not. While in treatment (Strategy 1), participants in PSY1 and PSY2 reduce their cocaine use more than participants assigned to PSYCONTROL. This effect remained significant in the analysis with all data (Strategy 2). However, when the effect of treatment dropout is included, the slopes for both groups change, with the PSY1/PSY2 rate of change lessening and the PSYCONTROL slope changing direction. Thus, the difference between the groups indicated by Strategy 1 and 2 is reflected in the Treatment Status by PSYCH1 by Time effect in Strategy 3.

4. Discussion

This report was intended to highlight the implications of different methods of analyzing clinical trials data when there are substantial or differential levels of protocol violation. First, we found that results based on different strategies for handling data from protocol violators (in this case, treatment dropouts) affected conclusions about the effectiveness of the study treatments. That is, the results based on the intention-to-treat analysis that included all data from the intended duration of protocol treatments (Strategy 2) are unlikely to mirror the outcomes observed in actual clinical practice, where full compliance is rare. This strategy suggested several significant treatment effects that were not seen when only those data collected before participants' dropout or treatment completion were used (Strategy 1). Second, the data supported our assumption that participants' behavior would change at the point of treatment dropout. In this case, cocaine use increased significantly after participants dropped out of treatment. This highlights the drawbacks of carrying values forward or imputing values for protocol violators (Lavori, 1992). Third, the approach that took into consideration the point of treatment dropout and included active treatment status (Strategy 3) addressed questions about the efficacy of the treatments in a clinical application, allocating variance due to the actual participation in treatment. Findings based on this strategy highlight potential weaknesses of the classic intention-to-treat approach, which as a rule does not take into account the point at which protocol deviation occurs. In the case of this trial, use of this analytic strategy indicated that two treatment effects suggested by the classic intention-to-treat analysis appeared to be artifacts of differential attrition and a greater increase in cocaine use frequency for the control group after dropout compared to the active groups.

This demonstration was intended to highlight the issues and dilemmas commonly faced by drug abuse and other investigators analyzing clinical trials data sets in which significant

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

numbers of participants violate the protocol (and thus tend not to contribute data after the point of protocol violation). In such cases, there have been until recently very limited options for handling the issue of missing data, most of which were highly problematic (e.g. imputing missing values, dropping participants with missing data). Moreover, as seen here, many participants seek alternative, non-protocol treatments after leaving a trial, making interpretation of the 'true' effects of study treatments less straightforward.

This demonstration also highlights several potential pitfalls associated with employing a single strategy in analyzing clinical trials data. First, when there is differential attrition or compliance by treatment, a single intent-to-treat analysis may not present a full picture of treatment effects, particularly if participants' symptoms change at the point of attrition. In such cases, supplemental analyses that take into account the point at which protocol violation (dropout, noncompliance, withdrawal) occurred might provide a more complete understanding of treatments' effects and their robustness. For example, in this study, the answer to the question, 'Is MED effective in reducing cocaine use?' might be 'Yes, participants assigned to MED had better outcomes during the full 12 weeks of the protocol (Strategy 2) and stayed in treatment longer. Because participants in MED were in treatment longer than those in NOMED, and because participants who dropped out of treatment subsequently used more cocaine than those who remained in treatment, if the effect of remaining in treatment is considered, retention is a significant predictor possibly overshadowing the independent effect of the medication (Strategy 3)'.

Another pitfall associated with relying solely on a single analytic strategy when using the intention-to-treat principle occurs when the active treatment and post-termination slopes shift differently by treatment group. In this case, the classic intent to treat analysis suggested a significant effect for PSY2 and PSY1 over PSYCON-TROL. However, our exploratory analysis suggested significantly higher levels of post-attrition cocaine use by dropouts assigned to PSYCONTROL versus those assigned to PSY1 and PSY2. This suggests that although PSY1/PSY2 were associated with improved outcomes compared to PSYCONTROL, this effect may have been downgraded by the significantly greater increase in cocaine use for those in PSYCONTROL seen post-termination. Thus, the apparent benefit of PSY2/PSY1 over PSYCONTROL may be in the active treatment, but also reflect some component of PSY1/PSY2 that is associated with continued treatment benefits post-termination that are not seen in PSYCONTROL.

This demonstration suggests that conclusions about the efficacy of study treatments may vary depending on the analytic strategy used, underscoring the importance of reporting not only the nature of the protocol deviation, as suggested by the CONSORT guidelines (Begg et al., 1996; Moher et al., 2001), but also describing the specific analytic strategy used. At a minimum, investigators should report the number of participants in each treatment group who dropped out or for whom data were unavailable for other reasons (withdrawal, noncompliance, loss to follow-up), as well as how their data were handled (method of imputation, how many values imputed, whether analysis accounted for differential retention or exposure to non-study treatments). The interpretations of the findings, of course, should be anchored in the methodological approach.

4.1. Limitations

A major limitation of this demonstration is that data from all treatment dropouts were not available. Although 54 treatment dropouts (some of whom had no exposure to study treatment), were interviewed at the 12-week termination point, we were unable to contact the remaining 29 treatment dropouts, despite extensive efforts to track these individuals, which included multiple means of contact (phone, certified letters) as well as multiple locators and informants (each participant was asked to provide the name, address, and phone

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

number of at least three individuals who would always know their whereabouts). This highlights the difficulties faced by investigators, especially those working with substance users and other clinical samples characterized by poor treatment retention and compliance, in their efforts to conduct intention-to-treat analyses. Moreover, although these 29 individuals did not differ significantly from the 54 treatment dropouts we successfully reached on a range of baseline and treatment variables, it is possible that they differed in other ways. Thus, results based on inclusion of their data may have differed from those presented here.

4.2. Recommendations and conclusions

First, the intention-to-treat principle should be incorporated into the design of all randomized clinical trial protocols, by including the methods in the standard operating procedures, rather than at protocol conclusion, when options become restricted by the availability of the data. Although it is difficult in practice to separate data collection from treatment exposure, investigators must make every effort to collect data from participants who violate study protocols through dropout and noncompliance. In cases where participants cannot be reached after protocol violation, random effect regression analyses may provide reasonable estimates of treatment outcome, compared with approaches such as carrying values forward; imputing mean values, assuming all treatment dropouts had poor outcome, or deleting participants with missing values. However, as Lavori (1992) has noted, ‘there is no royal road to imputation’. Neither random effect regression approaches nor their alternatives are without bias and there remains no substitute for complete datasets in clinical trials.

Second, greater clarity and consistency should be used in reporting of ‘missingness’ in clinical trials, particularly around the distinction between missing participants and missing data. Investigators should report the number of participants included in analyses of treatment effects as well as how missing data were handled (Federal Register, 1998). Protocol adherence should be reported as an outcome as well as a descriptor of treatment efficacy.

Finally, while the gold standard of RCT reporting is intention-to-treat analysis, many trials continue to fall short of this standard and instead report on within-treatment outcomes only. We recommend greater efforts to anchor interpretation of study findings in the exact sample analyzed, as suggested by the CONSORT guidelines. In cases where results based on within-treatment data differ from those based on intention-to-treat data, investigators should report both, as the two types of analyses address different questions and together may provide an improved understanding of treatment effects (Meyers, 1999). Furthermore, conducting supplemental analyses that account for protocol violation in an intention-to-treat analysis may provide a more comprehensive understanding of treatment effects.

Acknowledgments

Support was provided by the National Institute of Drug Abuse grants P50 DA 09241, R01 DA10679, and K02 DA00248. We gratefully acknowledge Tami Frankforter and Theresa Babuscio who assisted in data management and manuscript preparation, and Seth Powsner, MD, who assisted in creating graphics.

References

Adetugbo K, Williams H. How well are randomized controlled trials reported in the dermatology literature? Arch Dermatol. 2000; 136 (3):381–385. [PubMed: 10724201]

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

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Bollini P, Pampallona S, Tibaldi G, Kupelnick B, Munizza C. Effectiveness of antidepressants. Meta-analysis of dose– effect relationships in randomized clinical trials. Br J Psych. 1999; 174:297–303.

Bryk AS, Raudenbush SW. Application of hierarchical linear models to assessing change. Psychol Bull. 1992; 101:147–158.

Carroll, KM. A Cognitive-Behavioral Approach: Treating Cocaine Addiction (NIH Publication 98-4308). NIDA; Rockville, MD: 1998.

Carroll KM, Nich C, Ball SA, McCance E, Rounsaville BJ. Treatment of cocaine and alcohol dependence with psychotherapy and disulfiram. Addiction. 1998; 93:713–728. [PubMed: 9692270]

Carroll KM, Nich C, Sifrey R, Nuro KF, Frankforter T, 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; 57:225–238. [PubMed: 10661673]

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

DelBoca FK, Kranzler HR, Brown J, Korner P. The assessment of compliance with pharmacotherapy through UV light detection of a riboflavin tracer. Alcohol Clin Exp Res. 1996; 20:1412– 1417. [PubMed: 8947318]

Department of Health and Human Services. Food and Drug Administration. Statistical Guidance for Clinical Trials of Non-Diagnostic Medical Devices. Government Printing Office; Washington: 1996.

Edwards AG, Rollnick S. Outcome studies of brief alcohol intervention in general practice: the problem of lost subjects. Addiction. 1997; 12:1699–1704. [PubMed: 9581002]

Fawcett J, Epstein P, Feister SJ, Elkin I, Autry JH. Clinical management—imipramine/placebo administration manual. NIMH. 1987; 23:309–324.

Federal Register. International conference on harmonization of technical requirements for registration of pharmaceutical for human use: Statistical Principles for Clinical Trials 97D:0174. 1998. p. 49583-49598.

Feinstein, A. Intent-to-treat policy for analyzing randomized trials: statistical distortions and neglected clinical challenges. In: Cramer, J.; Spilker, B., editors. Patient Compliance in Medical Practice and Clinical Trials. Raven Press; New York: 1991. p. 359-370.

Feinstein, AR. ‘Compliance bias’ and the interpretation of the therapeutic trials. In: Haynes, RB.; Taylor, DW.; Sackett, DL., editors. Compliance in Healthcare. Johns Hopkins Press; Baltimore, MD: 1979. p. 309-322.

Figueredo AJ, McKnight PE, McKnight KM, Sidani S. Multivariate modeling of missing data within and across assessment waves. Addiction. 2000; 95 (Suppl):S361–S380. [PubMed: 11132363]

Gibbons RD, Hedeker D, Waternaux C, Davis JM. Random regression models: a comprehensive approach to the analysis of longitudinal psychiatric data. Arch Gen Psychiatry. 1988; 50:739–750. [PubMed: 8357299]

Hedeker, DH. Computer Program. NIMH Division of Services Research; Rockville, MD: 1993. MIXREG: A Fortran Program for Mixed-Effects Linear Regression Models.

Hill, AB. Principles of Medical Statistics. Oxford University Press; New York: 1961.

Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. BMJ. 1999; 319:670–674. [PubMed: 10480822]

Howard KI, Cox WM, Saunders SM. Attrition in substance abuse comparative treatment research: the illusion of randomization. NIDA Res Monogr. 1990; 104:66–79. [PubMed: 2092226]

Ladoucuer R, Cosselin P, Laberge M, Blaszezynski A. Dropouts in clinical research: do results reported reflect clinical reality? Behav Therapist. 2001; 24:44–46.

Lavori PW. Clinical trials in psychiatry: should protocol deviation censor patient data? Neuropsychopharmacology. 1992; 6:39–48. [PubMed: 1571068]

Lavori PW, Bloch DA, Bridge PT, Leiderman DB, LoCastro JS, Somoza E. Plans, designs, and analyses for clinical trials of anti-cocaine medications: where we are today. J Clin Psychopharmacol. 1999; 19:246–256. [PubMed: 10350031]

Lee YJ, Ellenberg JH, Hirtz DG, Nelson KB. Analysis of clinical trials by treatment actually received: is it really an option? Stat Med. 1991; 10:1595–1605. [PubMed: 1947515]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Mattson ME, DelBoca FK, Carroll KM, Cooney NL, DiClemente CC, Donovan D, Kadden RM, McRee B, Rice C, Rytarik RG, Zweben A. Compliance with treatment and follow-up protocols Project MATCH: predictors and relationship to outcome. Alcoholism: Clin Exp Res. 1998; 22:1328–1339.

Meinert CL. Beyond CONSORT: need for improved reporting standards for clinical trials. JAMA. 1998; 279:1487–1489. [PubMed: 9600487]

Meyers, W. Dealing with dropouts in clinical studies. In: Meibohm, A.; Arseven, E.; Alemavehu, D., editors. Biopharmaceutical Report. Vol. 7. Biopharmaceutical Section of American Statistical Association; Alexandria, VA: 1999 Spring.

Moher D, Schultz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. JAMA. 2001; 285:1987–1991. [PubMed: 11308435]

Nowinski, J.; Baker, S.; Carroll, KM. NIAAA Project MATCH Monograph Series. Vol. 1. National Institute on Alcohol Abuse and Alcoholism; Rockville, MD: 1992. Twelve-step facilitation therapy manual: a clinical research guide for therapists treating individuals with alcohol abuse and dependence. (DHHS Publication No. (ADM) 92-1893)

Ruiz-Canela M, Martinez-Gonzalez MA, de Irala-Estevez J. Intention to treat analysis is related to methodological quality. BMJ. 2000; 320:1007. [PubMed: 10753165]

Scherer RW, Crowley B. Reporting of randomized clinical trial descriptors and use of structured abstracts. JAMA. 1998; 280 (3):269–272. [PubMed: 9676678]

Schultz KF, Grimes DA, Altman DG, Hayes RJ. Blinding and exclusions after allocation in randomised controlled trials: survey of published parallel group trials in obstetrics and gynaecology. BMJ. 1996; 312:742–744. [PubMed: 8605459]

Twitchell GR, Hertzog CA, Klein JL, Schuckit MA. The anatomy of a follow-up. Br J Addict. 1992; 87:1327–1333. [PubMed: 1392554]

Wothke, W. Longitudinal and multi-group modeling with missing data. In: Little, TD.; Schnabel, KU.; Baumert, J., editors. Modeling Longitudinal and Multiple Group Data: Practical Issues, Applied Approaches and Specific Examples. Lawrence Erlbaum Associates; Mahwah, NJ: 2000.

Zweben, A.; Barret, D.; Carty, K.; McRee, B.; Morse, P.; Rice, C. Strategies for Facilitating Protocol Compliance in Alcoholism Treatment Research. Vol. 7. NIAAA; Bethesda, Maryland: 1998.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

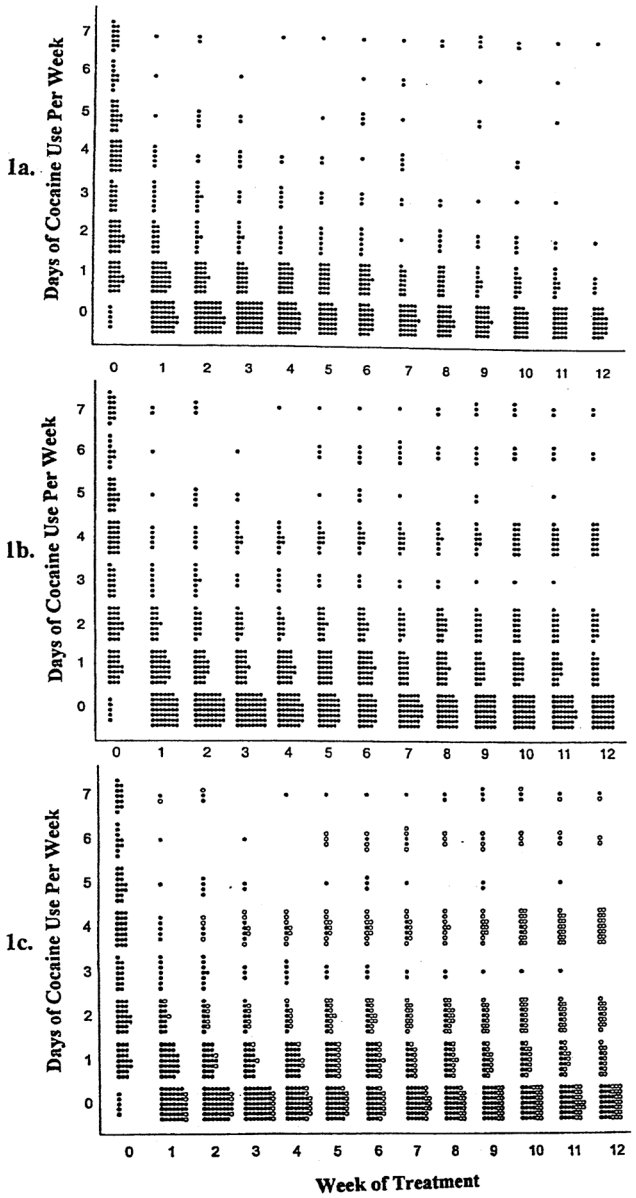

Fig. 1.

Individual data points used in three statistical models. (a) Strategy 1. Standard analysis using random effect regression model on within-treatment data only. (b) Strategy 2. Random effect regression model on all data for the full intended duration of the protocol, including data points collected after treatment termination. (c) Strategy 3. Random effect regression model on all data for full intended duration including both data points collected after treatment termination and a covariate for whether the outcome was collected during treatment. Actual data point (●); Actual data point gathered after treatment termination (○).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

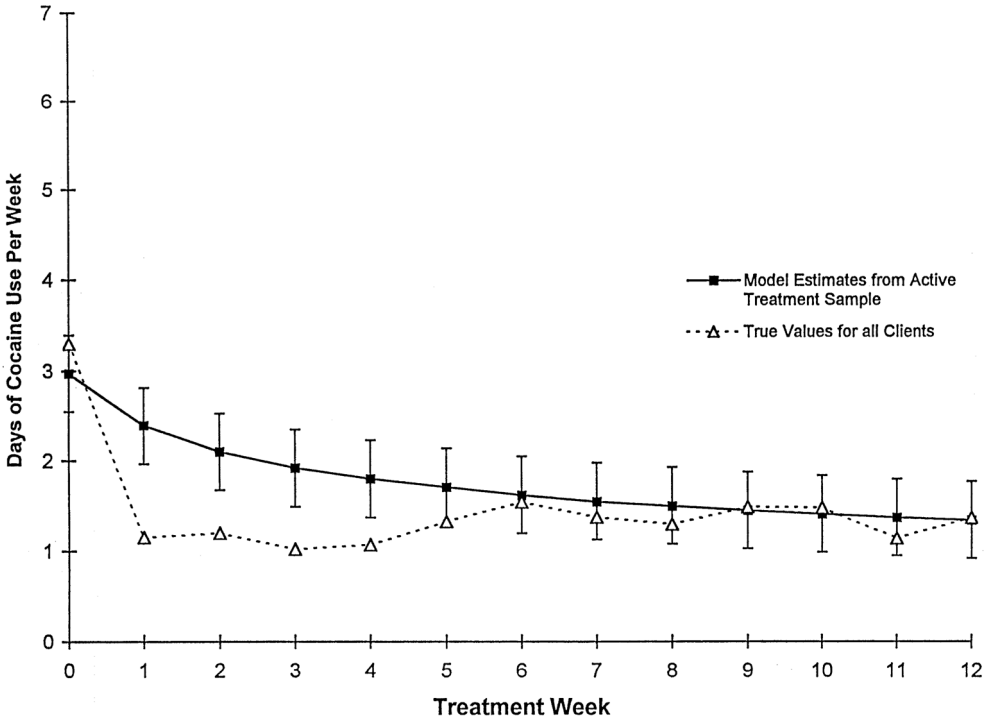

Fig. 2.

Comparison of model estimates from data collected within-treatment (Strategy 1, *n*=122/903) to mean observed values from data collected both within-treatment and after treatment termination (Strategy 2, *n*=122/1308).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 1

Comparison of treatment effects for different samples and data points

Sample size (subjects/observations)	Strategy 1		Strategy 2		Strategy 3	
	Standard analysis using random effect regression model on within-treatment data only		Random effect regression model on all data for the full intended duration of the protocol		Random effect regression model on all data for full intended duration including a covariate for whether the outcome was collected during treatment	
	122/903		122/1308		122/1308	
	z-score	*P*-value	*z*-score	*P*-value	z-score	*P*-value
Time (overall slope/change over treatment)	−12.50	<0.01	−7.61	<0.01	−9.63	0.00
Medication by Time	−0.48	0.62	−4.24	<0.01	0.19	0.85
Psych Contrast 1 by Time (PSY1+PSY2 versus PSYCONTROL)	−2.01	0.04	−2.66	<0.01	1.18	0.24
PSYCH2 by Time (PSY1 versus PSY2)	1.46	0.14	2.78	<0.01	3.02	0.00
PSYCH2 by Med by Time	−0.64	0.32	−1.44	0.15	1.63	0.10
Active Treatment Status	–	–	–	–	8.60	0.00
Treatment Status by Med	–	–	–	–	−1.29	0.20
Treatment Status by PSYCH1	–	–	–	–	−2.92	0.00
Treatment Status by PSYCH2	–	–	–	–	−2.28	0.02
Treatment Status by Med by PSYCH1	–	–	–	–	−2.60	0.01
Treatment Status by Med by Time	–	–	–	–	0.77	0.44
Treatment Status by PSYCH1 by Time	–	–	–	–	2.84	0.00
Treatment Status by PSYCH2 by Time	–	–	–	–	2.34	0.02
Treatment Status by PSYCH1 by Med by Time	–	–	–	–	2.14	0.03