

Highlights

Most large-scale treatment studies adequately reflect the providers and clients upon whom they are based, but room for improvement exists.

Only high quality data can support treatment improvement and resource allocation.

Potential Sources of Bias in Substance Abuse Treatment Follow-up Studies



NEDS Analytic Summary Series

The Center for Substance Abuse Treatment (CSAT) works to improve the lives of those affected by alcohol and other substance abuse, and, through treatment, to reduce the ill effects of substance abuse on individuals, families, communities, and society at large. Thus, one important CSAT mission is to expand the knowledge about effective substance abuse treatment and recovery services. In support of these efforts, CSAT established the National Evaluation Data Services (NEDS) contract to provide a wide array of secondary data analysis products to the substance abuse treatment field.

Specifically, the NEDS project is focused upon providing CSAT with an analytic capability to use existing data to address policy- and practice-relevant topics as well as future research and evaluation activities. NEDS has developed several product lines designed to provide analytic findings to analytic substance abuse treatment policy makers, service providers, services researchers and evaluators in a format that is most useful to the end user.

The Analytic Summary is one of the NEDS product lines. The purpose of the Analytic Summary is to provide a brief summary of each technical report produced by NEDS written in non-technical language. Readers who find the Analytic Summary results of interest can contact the original NEDS technical report authors for more detailed information. Through this process, the NEDS Analytic Summaries provide information to the substance abuse treatment field and promote linkages among different areas in the field.

This Analytic Summary

This particular NEDS Analytic Summary is based on the NEDS Technical Report titled *Potential Sources of Bias in Substance Abuse Treatment Follow-up Studies* (Johnson, R. A. and Gerstein, D. R., July 1999). For a more thorough discussion of the analysis and findings, please obtain a copy of the complete Technical Report. Information for doing so is provided on the last page of this summary.

Contents

Analytic Importance	2
Analytic Approach	3
Findings	5
Implications	8
Future Steps	10



Potential Sources of Bias in Substance Abuse Treatment Follow-up Studies

Analytic Importance

One important function of CSAT is to expand the knowledge about and improve the availability of effective treatment and recovery services for those affected by alcohol and other substance abuse. In sponsoring these secondary analyses, CSAT is attempting to learn: *What actually constitutes effective (and cost-effective) treatment and for what populations?*

The value of research results to policy makers and to the general public is driven in part by the quality of the data upon which the results are based. Substance abuse treatment follow-up studies need to be of premium quality with well-constructed research designs, expert data collection and analysis, and clearly delineated results. Follow-up studies of substance abuse treatment must be of high quality so that they can be used to:

- Identify effective treatment approaches for continued support and funding
- Allocate resources to treatment approaches commensurate with their efficacy and efficiency
- Illuminate strategies to improve the effectiveness of treatment
- Add to the treatment knowledge base.

This analysis describes how future follow-up studies and analyses may be improved based upon the lessons learned in conducting four major studies during the 1990s.

Analytic Purpose

CSAT, other HHS agencies, and a State agency have sponsored four major large-scale studies of substance abuse treatment effectiveness in the past decade. Each study has produced useful insights for the treatment field, corroborating prior evaluations of the effectiveness and cost-benefits of substance abuse treatment provided to many types of client in several treatment settings. Each study also provides a valuable resource for the identification of best-practice treatment strategies.

The value of these treatment studies underscores how important it is to maintain a high standard of data quality and further improve upon design and implementation of treatment follow-up studies. Comparing research methods used in these studies can identify strategies for maximizing data quality. These strategies can be used in the analysis of existing data, the completion of current research, and the design of future studies. The questions assessed in this report are:

- How well do the data collected in larger-scale follow-up studies represent the populations of substance abusers and treatment clients in the U.S.?
- How can future studies be designed and analyzed to make findings more representative of these populations?
- What are the benefits of making large-scale follow-up studies more representative and accurate?

This analytic summary describes:

- **Attributes of high quality data**
- **Differences between treatment study groups and the providers and clients they represent**
- **Strategies to take these differences into consideration and minimize them in the future.**



Potential Sources of Bias in Substance Abuse Treatment Follow-up Studies (cont.)

High quality treatment studies select participants who truly represent groups of providers and clients, most of whom then supply requested information.



Exhibit 1 Summary of Four Substance Abuse Treatment Follow-Up Study Designs				
Study	Provider Sampling	Client Follow-up Sampling	Data Collection	Average Time to Follow-up
California Drug and Alcohol Treatment Assessment (CALDATA)	Random selection of California counties and providers within those counties	Random	1 retrospective interview	15 months
Services Research Outcomes Study (SROS)	Random selection of providers from a prior national study	Random	1 retrospective interview	5 1/2 years
National Treatment Improvement Evaluation Study (NTIES)	Intentional selection of specific CSAT demonstration projects, Undersampled methadone providers.	Random	Repeated interviews	11 months
Drug Abuse Treatment Outcome Study (DATOS)	Intentional selection of specific, well-established providers in 11 cities	Random sub-sample of intake interview completers	Repeated interviews	1 year

Analytic Approach

Four major large-scale studies of treatment effectiveness were reviewed in this analysis. They are described in Exhibit 1.

The goal underlying research studies such as these is to collect data that fully reflect the treatment and the results of treatment for an identified population of clients. To accomplish this, some treatment follow-up studies conduct multiple stages of sampling in one study. For example, a sample of treatment providers may first be chosen from whom clients then are sampled. Each stage is prone to challenges that can ultimately affect how well the data reflect the original population.

It is normal for there to be some differences between what we learn from a final data set and what would be true of the original study population. Such differences, known as “bias,” exist to some degree in all ambitious, large-scale studies. No affordable single study could eliminate all bias. Some studies have less bias than others. Many studies

adequately represent the original population, while some do not.

Understanding the types of bias can help researchers address them with existing data and minimize them in the future. Two types of measurement issues and two types of study design issues are most likely to occur in post-treatment follow-up studies.

Measurement Issues

Selecting study participants and obtaining their cooperation are key issues in obtaining rigorous measurements.

- **Selecting Participants.** The criteria used to select providers or clients for a treatment study may affect how well the data represent the original population. Generally, random selection assures the overall best representation.

- *Provider selection.* Researchers/evaluators often recruit providers to participate in studies based upon their treatment model or their participation in funding

Potential Sources of Bias in Substance Abuse Treatment Follow-up Studies (cont.)

A well-designed study will support desired comparisons and identify the groups of providers and clients to whom results apply.

It is particularly difficult to define an untreated group that is just like the treated one in every way except having come into treatment.



programs that are being studied. Additional criteria may include the provider's convenience or likelihood of cooperation.

- *Client selection.* Clients go through a selection process during which (1) they decide whether and where to apply for treatment, and (2) providers decide which clients to admit or refer elsewhere. Some clients seek treatment voluntarily, while others are required by authorities to enter treatment.

- **Adequate Response.** Most providers and clients agree to participate in studies and provide the requested data, but some do not. It is not always clear why non-response occurs. Non-response becomes a problem when respondents turn out to be very different from non-respondents. In that case, data from respondents may not reflect what non-respondents would have said. When there are non-respondents, no matter how large the remaining sample, results must be interpreted cautiously.

- *Provider non-response.* Some providers decline to participate in studies. Others do not follow through on their data collection responsibilities. Compared with cooperating providers, non-cooperating providers may be newer, less well organized, or understaffed. Records differ from provider to provider in both quality and content, making it difficult for some providers to supply requested data.

- *Client non-response.* Some clients are more difficult to locate

and interview than others in ways that could make respondents different from non-respondents. For example, homeless clients may be more difficult to locate, but their homelessness is important for understanding treatment effectiveness.

Study Design Issues

A well-designed study will support the types of comparisons analysts wish to make and define the population to which those results may be applied.

- **Comparison groups.** Treatment is considered effective when clients improve their behavior or status following treatment compared to before treatment. Knowing whether clients improved more with than without treatment would augment such conclusions. This requires comparable data on groups of substance abusers who did and who did not receive treatment.

It is very difficult, however, to find an untreated group, and it is unethical to withhold treatment. There is generally no master list of untreated substance abusers from which to develop a pure comparison group.

Comparisons of clients who receive different types of treatment can indicate whether one type is better than another. Some very large studies provide results from different types of treatment. However, such comparisons require very good data on the services provided to clients.

- **Defining the population.** Treatment studies provide information about what happens to people who seek treatment in certain treatment settings. Treatment studies

Potential Sources of Bias in Substance Abuse Treatment Follow-up Studies (cont.)

The four largest treatment studies over the past decade:

- Interviewed thousands of clients
- Selected providers to represent certain important types of providers
- Followed up with clients for at least a year.

often cannot provide information about what happens to individuals who do not seek help, sought but did not receive help, or received help in different settings such as self-help or individual counseling. Additional information may be needed to indicate how well treatment would work for all persons who are addicted to drugs or alcohol.

Robust studies will adequately address these measurement and study design issues in the planning stages. However, careful consideration of these issues is necessary at the analysis stage, as well.

Findings

There are many similarities and important differences between the four studies described in Exhibit 2.

- In CALDATA and SROS, treatment providers were selected randomly to represent certain types of providers; these studies include one interview with clients completed after treatment.
- In NTIES and DATOS, treatment providers were hand-picked to represent certain groups of providers.

These studies include several interviews with clients over time.

- Each study included clients in similar types of treatment settings, but the proportions treated in each setting differed across each study.
- CALDATA is the only one of the four studies to draw a random sample of both providers and clients. Unlike other studies, this assures that the providers include newly-established as well as long-lived and stable providers. CALDATA only collected data in one state (California).

Evaluating Bias from Non-Participation

National follow-up studies generally set the goal of obtaining cooperation from 80 percent or more of treatment providers and 70 percent or more of clients following treatment. The four studies met or approached these quality standards in some but not all instances (see Exhibit 3). This is often not true for smaller studies (with fewer than 200 clients) reported in peer reviewed literature. The respective study teams used many innovative strategies to work with providers and reach clients. Nonethe-

Exhibit 2 Similarities and Differences Across Studies	
Similarities	Differences
<ul style="list-style-type: none"> ■ Large samples at follow-up: ranging from 1,799 clients (SROS) to 5,388 clients (NTIES) interviewed. ■ Early 1990's: Patients were discharged from treatment between 1989 and 1994. ■ Data collection method: All four studies rely on client interviews ■ Geographical area: 1 State (CA) sample, 3 National ■ Monetary incentives: Clients received about \$15. ■ Data collectors: Mostly the same organizations 	<ul style="list-style-type: none"> ■ Follow-up response rates: Vary from 63 to 82 percent ■ Provider sampling strategy: Two studies employed random samples of providers while two studies selected providers purposively ■ Time to follow-up interview: Ranged from about a year to 5 years ■ Other data sources: Additional data were collected from program, clinical records, and urine samples (except CALDATA). One study collected official arrest records (NTIES).



Potential Sources of Bias in Substance Abuse Treatment Follow-up Studies (Cont.)

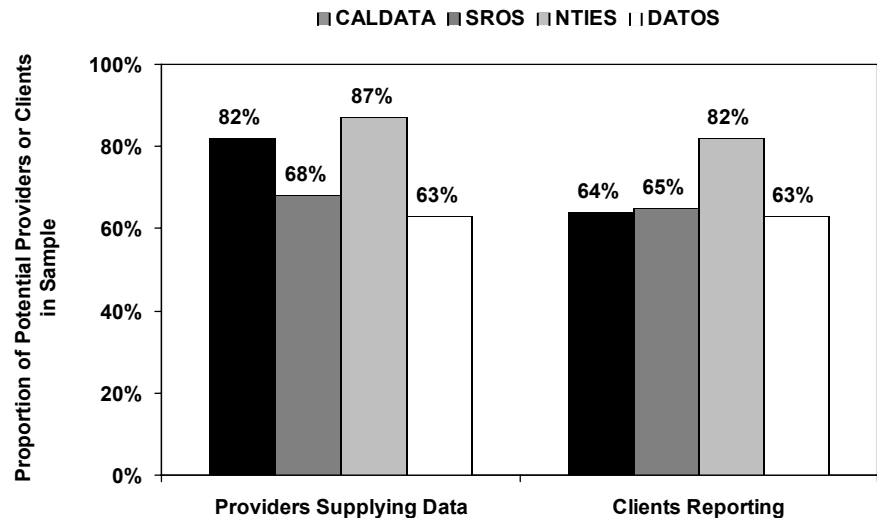
Most large-scale treatment studies approached the goal of obtaining data from 80 percent of selected treatment providers and 70 percent of clients.

One study concluded:

- Specific Non-participating treatment providers are generally similar to participating providers
- Specific A few demographic differences existed between participating and non-participating clients at treatment entry.



Exhibit 3
Provider and Client Response Rates



less, each study had some providers and some clients who elected not to participate in the study.

Many challenges hamper the achievement of high follow-up response rates. For example, of the original CALDATA client sample:

- 15 percent were treated by providers who declined to participate in the study
- 18 percent could not be located
- 10 percent refused to participate
- 6 percent had died or were otherwise inaccessible.

Reasons for non-response were documented as follows:

- Provider sample response rates were lower for methadone programs, due to non-cooperation of two large chains of private, for-profit providers.
- Clients were difficult to locate either because provider records were inadequate to help researchers locate

clients, or because former clients tended to move a lot, and family contacts were reluctant to disclose locating information. Moreover, some clients did not wish to reopen a “closed chapter” in their lives.

External data sources can determine how different the final respondent sample was from the original intended study population. Using existing survey data and program records, CALDATA researchers concluded that final data were not substantially biased.

- Non-cooperating providers were similar to participating providers. This is based upon comparisons using the National Drug and Alcoholism Treatment Unit Survey (NDATUS, now the Uniform Facility Data System). Residential and methadone treatment providers were broadly similar, while other treatment modalities were very similar.
- Responding clients were similar to non-responding clients, based on analyses of program records at treatment entry. Respondents were more likely than non-respondents to be female, Hispanic, to use public

Statistical techniques can improve upon the interpretation of data by making study populations appear to better reflect broader populations.



Potential Sources of Bias in Substance Abuse Treatment Follow-up Studies (cont.)

funds to pay for treatment, and to have been unemployed at admission.

Since the differences between participating and non-participating clients are largely due to poor information on where they could be located, future improvements are feasible.

Evaluating Differences Between Treatment and Other Population Samples

These four studies show that treatment is effective for those who sought and received services in selected settings. These clients may, however, be different from non-treated substance abusers. For example, prior research has demonstrated that:

- Between half and two-thirds of individuals entering substance abuse treatment do so under pressure from the criminal justice system.
- While treated and untreated individuals may have had similar levels and duration of substance use, those who seek treatment may have had more psychological distress and have more support systems in place. One strategy for determining how well treatment might work for a larger population would be to compare substance abusers in treatment studies with those in large-scale population studies (e.g., the National Household Survey on Drug Abuse along with data on criminal justice populations). Analysts could assess major differences in the two populations and determine whether the different characteristics of the general population would predict better or worse outcomes based on how well similar groups of clients fared in treatment studies.

Statistical Methods can Adjust for Bias

Statistical methods can be used to handle some of the challenges presented by biased data. Three such methods are suggested:

- **Weighting Adjustments.** Sample weights are used to determine how many people in a study population are represented by one person's data. Weights can be adjusted to make a study population resemble larger populations. External data such as the NHSDA, in concert with data from criminal justice populations, could provide the standard for adjusting treatment study weights to reflect the household population in need of treatment.
- **Likelihood-Based Methods.** Analysts can assess what factors made some clients less likely to participate in the study than others, and whether those factors also predict poorer outcomes. For example, clients who dropped out of treatment early may be less likely to participate in follow-up studies and have poorer outcomes. Analysts could then adjust downward their conclusions about treatment effects to better reflect what would be expected from non-participants.
- **Imputation of missing data.** Missing data can be imputed, or derived, for individuals using established, straightforward techniques. One technique applies an average of data calculated for similar types of clients to replace missing values. Imputing data is often necessary because many statistical analyses will leave out all data for an individual if one variable is missing.

Researchers can strive to improve upon the rigor of follow-up studies by reading technical literature, improving the execution of follow-up studies, and trying different methods.



Potential Sources of Bias in Substance Abuse Treatment Follow-up Studies (cont.)

Imputing helps when all data are missing for some clients, e.g. clients not interviewed at follow-up.

Results using any of these methods need to be compared to results that did not use the methods to account for the effect of the procedure.

Implications

Non-random entry into treatment, non-random sampling of providers, and provider and client non-response challenge the validity and generalizability of follow-up study findings.

Implications for Research/Analysis

The importance of collecting accurate, complete data should be underscored. Methods for dealing with missing data would not be needed if data were not missing in the first place. There are a number of opportunities to improve future follow-up studies.

- **Access available guides and resources on study design and implementation.** In recent years, CSAT, NIDA, and other agencies and organizations have prepared guidebooks for conducting substance abuse treatment follow-up studies and analyzing the data. For example, CSAT prepared a comprehensive resource package known as the Integrated Evaluation Methods package. Among the many components of the IEM package are:
 - A manual for locating and interviewing clients following treatment (*Staying in Touch*)
 - Guidelines for selecting appropriate research designs

- Core data sets that have already been reviewed by researchers/evaluators and that constitute generally agreed upon and feasible items to collect in follow-up studies

- Guides for analyzing data.

- **Evaluate the relative costs and benefits of alternative data collection methods.** Several procedures exist for locating clients, securing their participation and cooperation, and completing data collection. Exploratory studies could help researchers determine whether alternative data-collection methods improve data quality.

- **Improve execution of the follow-up tasks.** Locating clients may require extra creativity and time. Locating some clients may take longer than originally desired. The costs of extending data collection need to be assessed in light of the value of each additional completed interview.

- Obtaining response rates of 90 percent may cost several times more per completed interview than attaining rates of 50-70 percent.

- One study (not reported here) found that the conclusions drawn about treatment clients continued to evolve as response rates grew higher. Results for some types of clients (e.g., pregnant heroin users) could not be derived with lower response rates.

- **Improve the representative sampling of treatment providers.** Selecting providers at random can be effective and efficient. Moreover,

Potential Sources of Bias in Substance Abuse Treatment Follow-up Studies (cont.)

Policymakers should compare results from multiple studies wherever possible, demand rigorous designs and execution, and adequately resource follow-up efforts.

Regardless of how good statistical techniques are, they can never replace complete data nor repair poor study designs or execution.



effort needs to be made to secure cooperation of certain provider groups that often have not participated in follow-up studies, such as large multi-site, proprietary chains. Monetary incentives for treatment providers might help.

- **Assess how well treatment studies represent original populations.** Compare treatment studies with broader populations to assess differences in represented groups. The National Household Survey on Drug Abuse and surveys of criminal justice populations could be used to compare clients with other chronic drug users. This would help analysts to statistically adjust study results to help audiences understand how the results might apply to general populations.
- **Investigate the causes of the different response rates achieved in the four studies.** Why are response rates considerably higher in one study compared to another? Understanding of the causes of this differential would help to improve response rates in future studies, thereby strengthening the data.
- **Explore the use of advanced statistical methods** to assess and correct biases due to non-random entry into treatment and to non-response.

Increasing participation by treatment providers and clients is the most desirable of these approaches for handling bias.

Implications for Policy-Makers/Users

The major studies in the substance abuse treatment field each possess strong and yet distinct designs that con-

sistently demonstrate treatment effectiveness. Recent analyses of several databases have also demonstrated that the economic benefits of treatment outweigh costs.

Challenges remain inherent in such ambitious efforts. Policy makers can reinforce development and application of improved science practice to meet these challenges.

- **Compare multiple studies.** Where possible, compare several studies to assess the degree of agreement on major, similar topics.
- **Anchor or “calibrate” policy studies by the results of treatment follow-up studies.** The impact of potential policy initiatives can be assessed through continued analysis of these large data sets.
- **Demand high quality data to inform best practice and increase accountability.** It is important to continue to ensure that those who direct and conduct research possess the appropriate technical, oversight, and other resources to provide the best quality information.
- **In order to obtain high quality data, ensure that studies are appropriately funded.** Adequate resources can help ensure that high quality data are obtained and meaningful results are produced.

Implications for Treatment Providers/Clients

Treatment providers are at the heart of any treatment follow-up study. They provide access to clients, share data, and help interpret results.

Potential Sources of Bias in Substance Abuse Treatment Follow-up Studies (cont.)

- **Work collaboratively with experienced researchers/evaluators.** Specifically, providers should ensure that data collection efforts are designed to efficiently provide valid insights regarding services and clients.
- **Maintain better information to help locate clients.** Providers can collect, update, and retain client identifiers, case numbers from other government agencies, and addresses and phone numbers of not only the client but family, friends, lawyers, parole officers, or other case workers. This information could be gathered as part of a treatment provider's intake and treatment exit procedures. This will also help providers conduct their own follow-up studies.
- **Continue efforts to adopt standard and consistent data collection and reporting formats.** Efforts to document the effectiveness of treatment services in the future will benefit as the internal consistency of information collected in the field increases. Automated data processing efforts under development now may augment these efforts.

Treatment practice is unlikely to significantly improve if useful, readily applicable and accurate results remain elusive.

Future Steps

The experience encapsulated in the four studies reviewed in this report can inform ongoing design and implementation of treatment studies. By continually improving the quality of data on substance abuse treatment services, researchers/ evaluators, policy makers, and practitioners will all enjoy a much stronger foundation upon which to build effective and efficient treatment programs to control the nation's problems with substance abuse. Clients are the ultimate beneficiaries of more effective treatment services.

References

Johnson, R. A. and Gerstein, D. R., *Potential Sources of Bias in Substance Abuse Treatment Follow-up Studies*, July 1999. Prepared under the NEDS contract by the National Opinion Research Center (NORC).

For more information, please contact the National Evaluation Data Services analysis team at (703) 385-3200, or visit the NEDS Web site.

**Find more on the web
<http://neds.calib.com>**

The authors of this Analytic Summary are Elizabeth Thune, Ph.D. and Doug Fountain.

This analytic summary was produced by Caliber Associates under the NEDS contract (270-97-7016). We wish to thank Dr. Ron Smith, CSAT Program Evaluation Branch, GPO for the NEDS contract, for his overall guidance and direction. We also wish to recognize the many NORC and Caliber staff who contributed to production of this analytic summary and the technical report upon which it is based.



The perspective offered in this document is solely that of the author(s) and does not reflect the policies or views of the Federal government, or any of its Departments or Agencies.