

**CRITICAL ELEMENTS TO CONSIDER FOR
METHODOLOGICALLY SOUND
IMPACT EVALUATIONS OF
DRUG COURT PROGRAMS**

**By Charles Michael Johnson and
Shana Wallace
U.S. General Accounting Office**

To demonstrate their effectiveness, drug courts must build methodologically sound impact evaluations. To be methodologically sound, impact evaluations should include certain critical elements, including: a comparison group similar to that of the participants; the collection and analysis of critical data at several points during and post program; and the involvement of an experienced evaluator.

The best method for building a similarly situated comparison group is to randomly assign qualified drug court participants to this group. If that is not possible, the individuals in the comparison group should match the participants in the drug court as closely as possible.

Data should be collected from participants at intake, during program participation, upon graduation, and after program completion or termination. Data should be collected from all participants and comparison group members, and should include, among other information, data on relapse and recidivism. Data should be maintained in an automated data management system.

The involvement of a qualified evaluator is critical to the evaluation process, especially during the design phase. Evaluators will assist the team in all aspects of evaluation design, and will ensure that, among other things, the comparison group can withstand scrutiny.

Charles Michael Johnson is an Assistant Director with the U.S. General Accounting Office.¹ He was also the Senior Analyst-in-Charge of the two prior congressionally requested GAO reviews of federally-funded drug court programs and has helped to tailor numerous recommendations for improving data collection and evaluation efforts of federally-funded drug court programs.

Shana Wallace is a Senior Social Science Analyst also with the U.S. General Accounting Office. She was involved in drug treatment evaluation research throughout most of her twelve years working as a research analyst for the Federal Bureau of Prisons, and has specialized training in evaluation research methods.

Direct all correspondence to Charles Michael Johnson, Assistant Director, U.S. General Accounting Office, 441 G Street, NW, Washington, DC 20548; (202) 512-8777.

¹ The ideas in this paper represent the opinions of the authors and not those of the U.S. General Accounting Office.

ARTICLE SUMMARIES

**METHODOLOGICALLY
SOUND IMPACT
EVALUATIONS**

[9] Building a sound evaluation involves a carefully constructed comparison group, data collection, and the input of a qualified evaluator.

COMPARISON GROUP

[10] Drug courts should take great care in constructing a comparison group, focusing on certain baseline characteristics.

**DATA COLLECTION &
ANALYSIS**

[11] Individual data should be collected at several points in time from participants and non-participants and stored in an automated data management system.

EVALUATOR

INVOLVEMENT CRITICAL

[12] It is critical that drug courts identify an evaluator with relevant experience and publications, if possible.

INTRODUCTION

Through support from federal, state, and local governments, drug court programs have grown significantly over the last decade. To sustain this growth, drug courts will find it necessary to demonstrate program impact on the communities they serve. Although many drug court programs have undergone some type of evaluation, many of these evaluations have not included critical elements essential to conducting methodologically sound impact evaluations, including the use of individuals who have formal training or experience in conducting impact evaluations. As drug courts become institutionalized in many areas of the country, drug courts need to adopt best practices in all areas of management, including evaluation. Past evaluations have lacked information, which may be attributed to inadequacies in the (1) collection and utilization of data and (2) design strategies for completing impact evaluations. This article will assist the drug court practitioner in identifying and building methodologically-sound impact evaluations.

WHAT IS A METHODOLOGICALLY SOUND IMPACT EVALUATION?

[9] An impact evaluation is an attempt to answer whether drug court program participants fare better, usually in terms of criminal recidivism and substance abuse relapse, than if they had not gone through a drug court program. This usually involves comparing outcomes for drug program participants to those of similarly situated offenders who are eligible for, but not participating in, a drug court program. Completing a methodologically sound impact evaluation may be complicated and resource-intensive, often more so than implementing a process evaluation or reporting on program outputs or statistics. To maximize the opportunity for

success, data collection and program impact evaluation plans should begin during the design stage of a drug court program.

Three critical elements help ensure that an impact evaluation is methodologically sound:

- Comparing program participants with non-participants and being careful, when constructing a comparison group, to ensure that it is composed of individuals similar to those in the participant group;
- Collecting and analyzing various types of data at several points in time for both participants and the comparison group, and doing so for post-program criminal recidivism and substance abuse relapse as well; and
- Involving a qualified, experienced evaluator, social science analyst, or statistician throughout the evaluation (especially during the design phase).

CRITICAL ELEMENT 1: Comparing Program Participants with Non-participants

[10] It is important to have some means of comparing participants in drug court programs with non-participants. A common method is to use comparison groups. The comparison group consists of eligible offenders who are not program participants. This group should be composed of individuals who, taken together, have similar characteristics to the individuals who go through the drug court program. Ideally, the only difference between drug court participants and comparison group subjects would be participation in the drug court program. This is often difficult to achieve. The optimal way to construct drug court participant and comparison groups is through random assignment. If a program happens to have more volunteers for the drug court than could be accommodated, a program should strongly consider (1) randomly assigning individuals to the program

and (2) using those that do not get randomly assigned as the comparison group. However, even this may be difficult to achieve as selection pressures may come from judges, lawyers, drug court administrators, and the defendants themselves.

If a program cannot randomly assign individuals, then the program should match drug court participants with subjects in a comparison group, focusing on and controlling for certain baseline characteristics, such as demographics, criminal justice and substance abuse history, and drug treatment motivation. When comparing outcomes, more complex statistical analyses will be required to control for the inherent differences between the two groups. Remember that at the end of the evaluation, a program should be able to argue that if program participants fare better, it is because of the drug court program and not because of baseline characteristics or other factors, such as participants, when compared with the comparison group subjects, were already more motivated to change, were older (younger people are more likely to be recidivist or relapse), or had less extensive criminal careers.

CRITICAL ELEMENT 2: Collecting and Analyzing Data at Several Points in Time

[11] A program should be vigilant in collecting various types of data over several points in time, including the collection of post-program recidivism and relapse data. In collecting data, some essentials should be considered:

- Collect data at several points in time for both the participant group and the comparison group. This includes collecting data when program participants enter the program, during treatment, and as they leave. In addition, it is vital to collect post-program recidivism and relapse data to be able to report on the

impact that drug court programs have after participants leave the program. Stakeholders will be interested in knowing if there are lasting effects.

- Collect individual-level data, not aggregates or averages.
- Collect data on all individuals, including those who leave the program without completing it. Even if a participant does not make it through the program, he or she should still be included. Results will be suspect if only graduates are compared with the comparison group.
- Maintain data in an automated data management system.

Data should be collected throughout the program: during the program time period, at the time of program completion or graduation, and for some period after participants graduate from, or leave, the program. The same data elements also should be collected for the comparison group for an equivalent time frame (that is, a logically equivalent time; for example, when a subject leaves jail or probation). Post-program data should be collected for at least six months; preferably, data should be collected for one or two years for both the participants and comparison group. It is important to track individuals and not just aggregate data or averages. Knowing the average relapse rate for a group of individuals at one point in time is not sufficient for ensuring a methodologically sound impact evaluation. Individual-level data are needed to establish that participants and the comparison group remain equivalent for any characteristics associated with outcomes. If a program only has aggregate data and then loses some participants or non-participants during follow-up, the validity of the matched comparison could become questionable because the two groups can no longer be shown to be equivalent.

For both program participants and non-participants, data should be collected for the following characteristics:

- Demographics (age, sex, marital status, race, income, and education);
- Criminal history (past arrests, convictions, and incarcerations);
- Substance abuse history;
- Level of use of controlled or addictive substances at the point of arrest;
- Measure of drug treatment motivation (if possible);
- Level and type of substance abuse treatment;
- Substance abuse relapse while in the program;²
- Rearrest or conviction for a crime while in the program;
- Completion or non-completion of the drug court program;
- Whereabouts and contact information at the time of program completion;
- Substance abuse relapse after program completion; and
- Rearrest or conviction for a crime after program completion.

We recognize that obtaining post-program data for the comparison group can be difficult. Often, data collection strategies must be prearranged with officials from other jurisdictions or other parts of the criminal justice system to ensure the availability of data throughout the evaluation. Finally, to improve the efficiency of data collection and data analysis, as well as the methodological soundness of future

² “While in the program” refers to the drug court program for program participants. However, for comparison group subjects, this may refer to the time during which comparison group subjects are under supervision, in jail, or another comparable point in time.

impact evaluations, a program should use a computerized data management system if at all possible.

Feasibility of Collecting Post-Program Data

Information collected from the U.S. Department of Justice's Drug Courts Program Office (DCPO) grantees continues to support the feasibility of collecting post-program outcome data. An estimated two-thirds of the DCPO-funded drug court programs maintained criminal recidivism data on participants after they left the programs, according to the results of follow-up structured interviews the U.S. General Accounting Office (GAO) conducted during 2001 with a representative sample of DCPO-funded drug court programs.³ Of the remaining one-third that did not maintain post-program recidivism data, about 63 percent indicated that it would be feasible for their program to provide such data.⁴ These estimates suggest that about 86 percent of DCPO-funded drug court programs would be able to provide post-program recidivism data if requested.

Through these follow-up interviews, GAO also found that about one-third of the DCPO-funded drug court programs maintained substance abuse relapse data on participants after they left the program.⁵ Of the estimated two-thirds that did not maintain post-program substance abuse relapse data, about 30 percent indicated that it would be feasible for their program to provide such data. These

³ About 84 percent of these programs collected post-program outcome data for six months or more.

⁴ GAO noted a 95 percent confidence interval, ranging from 45 to 78 percent, for this estimate.

⁵ About 84 percent of these programs collected substance abuse relapse data for six months or more after participants left the program.

estimates suggest that about 50 percent of DCPO-funded drug court programs would be able to provide post-program substance abuse data if requested.

According to survey results from data collected by the Drug Court Clearinghouse and Technical Assistance Project at American University (2000 and 2001), a significant number of the drug court programs were able to provide post-program outcome data. For example, about 47 percent of the DCPO-funded adult drug court programs that responded to the Drug Court Clearinghouse's 2000 operational survey reported maintaining some type of follow-up data on program participants after they left the program.⁶ Of these drug court programs, about 92 percent said that they maintained follow-up data on recidivism and about 45 percent said that they maintained follow-up data on relapse. Of the DCPO-funded adult and juvenile drug court programs that were in operation for at least one year and that responded to the Drug Court Clearinghouse's annual survey published in 2001,⁷ about 56 percent were able to provide follow-up data on program graduates' recidivism and about 55 percent were able to provide follow-up data on program graduates' relapse.

⁶ The Drug Court Clearinghouse's operational survey was administered to various adult drug court program stakeholders, including the judge and court officials, treatment providers, defense counsel, and participants. The response rate for the year 2000 survey was estimated at 88 percent.

⁷ The Drug Court Clearinghouse, under a grant from the U.S. Department of Justice, had been administering an annual data collection survey to collect operational and program participant data from operating adult, juvenile, family, and tribal drug court programs.

CRITICAL ELEMENT 3: Involving a Qualified Evaluator

[12] Because of the complexities involved in developing and executing a methodologically sound impact evaluation, it is important to obtain the services of a qualified, experienced evaluator. This evaluator could help create or review a program's evaluation design, advise on the evaluation effort as problems arise, and facilitate or perform the appropriate impact analyses. It is important to identify an evaluator who has previous experience with program evaluation and preferably one who has published the results of such work in a peer-reviewed journal. The evaluator should have specialized training in implementing studies that withstand some common threats to validity. This will ensure that if the program evaluation yields positive results, they can be attributed to the program itself and not to a myriad of other possible explanations. Competing explanations may occur when there are doubts about whether the comparison group is really equivalent to the participant group. For example, individuals who volunteer to go into a drug court program could be more motivated to change—both lifestyle and substance abuse patterns—than the subjects of the comparison group, who did not volunteer for a drug court program. Another problem may occur when one is able to obtain good follow-up data on drug court graduates, but unable to obtain follow-up data from a portion of the comparison group. This problem may compromise the validity of the study because it is possible that this missing portion of the comparison group differed in some important way from the non-missing portion.

As aspects of a program's evaluation change over time—and they generally do—the evaluator may advise the program on how to avoid any later threats to validity. The evaluator may also advise the program about how to analyze collected data, provide guidance on statistical analyses, or

better employ analyses that are appropriate for the types of data collected. Most evaluations suffer from some data loss and attrition (that is, the loss of some of the evaluation participants). For the comparison group, some sophisticated analyses may be required to control for characteristics that are not identical to those of the program participants. Finally, if random assignment is not possible, rely heavily on an evaluator to decide how to structure the comparison group.

CONCLUSION

Based on GAO's 2002 follow-up report, it would appear that drug court programs are closer than in previous years to having the data to (1) improve the quality of future impact evaluations and (2) better enable drug court programs, evaluators, and researchers to address program impact. However, if meaningful impact evaluations on the growing number of drug court programs are to be done, oversight agencies—such as the U.S. Department of Justice, the U.S. Department of Health and Human Services, and state and local governments—must encourage the continued collection and utilization of data on factors affecting program operations and outcomes. In particular, data should be collected on participants after they leave the program. Further, the effectiveness of programs could be demonstrated more definitively in future impact evaluations by including recidivism and, to the extent feasible, relapse data for both program participants and non-participants.

Without the inclusion of such data from a broad range of drug court programs, it will not be possible for drug court programs, researchers, or evaluators to adequately respond to issues raised on the overall impact of drug court programs.

REFERENCES

Drug Court Clearinghouse and Technical Assistance Project.
2000 Operational survey. Washington, DC:
American University Office of Justice Programs.

_____. *2001 Annual adult survey.* Washington, DC:
American University Office of Justice Programs.

_____. *2001 Annual juvenile survey.* Washington,
DC: American University Office of Justice Programs.

U.S. General Accounting Office. (1997). *Drug courts:
Overview of growth, characteristics, and results.*
GGD-97-106. Washington, DC: Author.

_____. (2002). *Drug courts: Better DOJ data
collection and evaluation efforts needed to measure
impact of drug court programs.* GAO-02-434.
Washington, DC: Author.