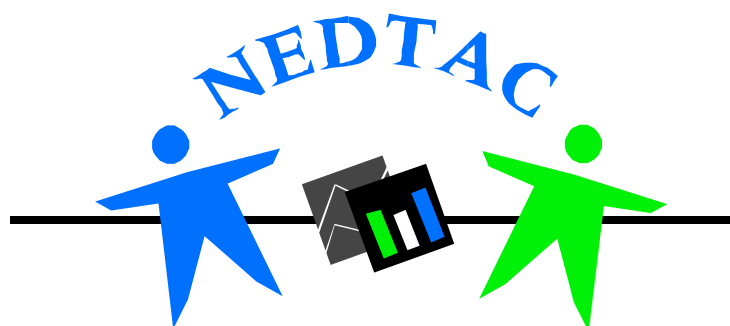


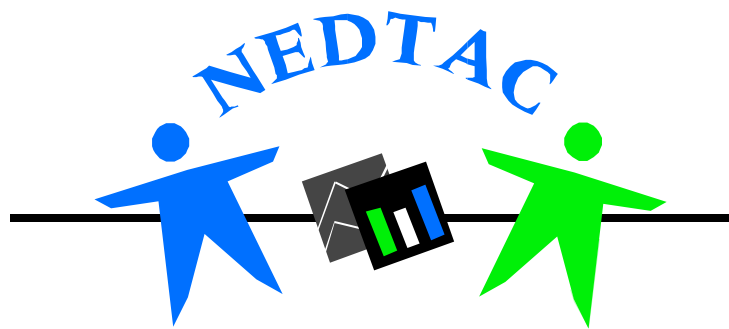
**NATIONAL EVALUATION DATA AND  
TECHNICAL ASSISTANCE CENTER**



**CROSS-SITE PROGRAM EVALUATION:  
SELECTED BIBLIOGRAPHIES, 1991-1998**

**April 1999**

# NATIONAL EVALUATION DATA AND TECHNICAL ASSISTANCE CENTER



## CROSS-SITE PROGRAM EVALUATION: SELECTED BIBLIOGRAPHIES, 1991-1998

April 1999

This document was supported by the Center for Substance Abuse Treatment, Department of Health and Human Services, Caliber/NEDTAC Contract No. 270-94-0001.

**CSAT**  
Center for Substance  
Abuse Treatment  
SAMHSA

## TABLE OF CONTENTS

	<u>Page</u>
<b>FOREWORD</b> .....	<b>i</b>
<b>I. OVERVIEW</b> .....	<b>1</b>
<b>II. ANNOTATED BIBLIOGRAPHY</b> .....	<b>2</b>
<b>III. SELECTED BIBLIOGRAPHY</b> .....	<b>8</b>
<b>IV. BENEFITS AND LIMITATIONS OF CROSS-SITE EVALUATIONS</b> .....	<b>9</b>

## FOREWORD

One of the missions of the Center for Substance Abuse Treatment (CSAT) is to enhance the nation's substance abuse treatment system by identifying, developing, and supporting appropriate policies, approaches, and programs. In short, for the field of substance abuse treatment, CSAT seeks to determine what works, for whom, how well, and at what cost.

Building knowledge through evaluation is the key to answering these questions. From CSAT's perspective, evaluation—including cost analysis and performance measurement—is an integral component of program management and part of an ongoing process of knowledge development, assessment, and improvement. Toward this end, CSAT's Program Evaluation Branch established the National Evaluation Data and Technical Assistance Center (NEDTAC) to advance state-of-the-art evaluation in the field of substance abuse treatment.

A primary NEDTAC activity was to provide evaluation technical assistance and support to substance abuse treatment providers and evaluators. One way in which NEDTAC pursued that objective was to produce a series of bibliographies in key topic areas. This document belongs to that series. Evaluations of multi-site programs present special challenges to evaluators concerned with the use of evaluation results. We hope this document will assist professionals within the substance abuse treatment community to meet those challenges and to think about effective and appropriate ways to serve diverse clients in substance abuse treatment programs.

This bibliography, along with others in the series, was developed under the guidance and direction of Ron Smith, Ph.D., the NEDTAC Government Project Officer, Program Evaluation Branch, Office of Evaluation, Scientific Analysis, and Synthesis. We also wish to thank Beth Archibald Tang for compiling and reviewing this document.

Sharon Bishop  
Project Director  
National Evaluation Data and  
Technical Assistance Center (NEDTAC)

## **I. OVERVIEW**

## I. OVERVIEW

This list of annotated citations and other references concerns cross-site evaluation, with an emphasis on methodology and, where possible, homeless prevention programs. The selected citations list other, related resources. The chart at the end of this document highlights benefits and limitations of multisite and/or cross-site evaluations from selected articles cited in this bibliography. Included in this document are:

- An Annotated Bibliography
- A Selected Bibliography
- A Benefits and Limitations of Cross-Site Evaluations chart.

NEDTAC reviews are for informational purposes only and should not be interpreted as a CSAT endorsement of any specific resource or publication.

## **II. ANNOTATED BIBLIOGRAPHY**

## II. ANNOTATED BIBLIOGRAPHY

Compton, W.M., Lamb, R.J., & Fletcher, B.W. (1995). Result of the NIDA treatment demonstration grants' cocaine workgroup: Characteristics of cocaine users and HIV risk behaviors. Drug and Alcohol Dependence 37, 1-6.

This journal issue includes seven articles from the NIDA treatment demonstration grants cocaine workgroup. While this article discusses findings, the results indicated that it is possible to define variables precisely for analysis across sites and laid the groundwork for the next set of analyses in which the common theme of HIV risk behaviors among cocaine abusers was agreed upon. The paper concludes that cross-site collaborations can take different forms and this collection of papers represents one successful approach. (NCADI abstract modified)

Freedman, J.A. (1991). Multisite evaluations of health care policies and programs. In R.S. Turpin, & J.M. Sinacore (Eds.), Multisite evaluations (New Directions for Program Evaluation 50, pp. 97-107). San Francisco: Jossey-Bass.

The health care arena has a number of unique features that generally facilitate and influence the conduct of evaluations at multiple sites. This chapter discusses the technical, historical, and political features of the health system that promote multisite evaluations (MSEs), e.g., the availability of large, multisite health care databases, the growth of multi-institutional health care systems, and the history of multisite biomedical clinical trials. In addition, the authors discuss types of MSEs, with examples taken from studies that depend solely on secondary data, studies that depend on existing staff for data collection, studies that combine the two methods, and studies that depend on original, evaluation-specific data collection.

Greenberg, D. et al. (1994). Multisite employment and training program evaluations: A tale of three studies. Industrial and Labor Relations Review 47, 679-691.

Using a multilevel statistics framework, a study examined the role of multiple sites in the evaluation of three programs. Although multisite evaluations have great potential, they are more difficult and expensive, and have limited success in producing useful findings. (ERIC abstract modified)

Guastello, S.J., & Guastello, D.D. (1991). How organizations differ: Implications for multisite program evaluation. In R.S. Turpin, & J.M. Sinacore (Eds.), Multisite evaluations (New Directions for Program Evaluation 50, pp. 71-81). San Francisco: Jossey-Bass.

Assessment of variability in program outcomes across evaluation sites is facilitated by an understanding of the characteristics that differentiate organizations. (Author abstract)

Hansen, W.B., & Kaftarian, S.J. (1994). Strategies for comparing multiple-site evaluations under non-equivalent design conditions. Journal of Community Psychology (CSAP Special Issue), 170-187.

CSAP currently funds over 200 community prevention partnerships. Partnership evaluations consist of single-site evaluations without controls. In order to determine the utility for cross-site analysis, 144 partnership evaluations were analyzed for process, outcome, and impact measures. Results of this content analysis yielded about 20 categories each for process, outcome, and impact that were available for cross-site analysis. Pairing process measures (independent variables) with outcome and impact measures (dependent variables) yielded a number of possible analyses that, if pursued, hold the promise of significantly advancing plausible hypotheses about how partnerships operate to achieve changes in substance use and abuse and outcomes important to these ends. (NCADI abstract modified)

Lamberti, M.J., & Katzenmeyer, C. (1996, Winter). Transforming qualitative data from templates into quantitative assessment of multisite programs. New Directions for Program Evaluation, 72, 81-88.

Use of template methodology in an evaluation of several teacher enhancement and development initiatives enabled evaluators to examine an extensive amount of qualitative data in a qualitative manner and to identify crosscutting themes and patterns. This chapter illustrates uses of templates for cross-site assessment and comparative analyses. These uses extend the utility of templates beyond identifying elements of best practice within a particular program to examining key elements across several program sites. (Author abstract modified)

Minnick, A., Klernpell, R.M., Micek, W., & Dudley, D. (1996). The management of a multisite study. Journal of Professional Nursing 12, 7-15.

This article presents a framework for multisite clinical studies based on research in 17 hospitals. Five areas of key research tasks are addressed: general systems design, public relations, human resource issues, data quality, and data management. (ERIC abstract modified)

Mowbray, C.T., & Herman, S.E. (1991). Using multiple sites in mental health evaluations: Focus on program theory and implementation issues. In R.S. Turpin, & J.M. Sinacore (Eds.), Multisite evaluations (New Directions for Program Evaluation 50, pp. 45-58). San Francisco: Jossey-Bass.

Human services programs at multiple sites are able to adapt to local needs and conditions, but this site-specific adaptation can foster differences in program implementation. (Author abstract)

Orwin, R.G., Goldman, H.H., Sonnefeld, L.J., Ridgely, M.S., Smith, N.G., Garrison-Mogren, R., O'Neill, E., & Sherman, A. (1994). Alcohol and drug abuse treatment of homeless persons: Results from the NIAAA Community Demonstration Program. Journal of Health Care for the Poor and Underserved, *5*, 326-352.

In a national evaluation, we assessed the implementation and outcomes of a multisite demonstration program for homeless persons with alcohol and other drug problems. We developed comprehensive case studies from data on client characteristics, utilization of services, implementation of interventions, and community systems of care at nine project sites. Client-level outcome data were analyzed to estimate the effectiveness of the interventions in a subset of projects with experimental or quasi-experimental evaluation designs. After controlling for baseline predictors, treatment clients in the majority of sites were significantly more likely than comparison clients to report improvement on one or more outcome dimensions. On alcohol use, for example, under conservative assumptions the average treatment client was drinking less at follow-up than were 57 percent of comparison clients. Analyses of predictor-by-treatment interactions suggested that clients with fewer problems benefited most from the interventions. The implementation analysis yielded a number of lessons for policy makers and program planners.

Orwin, R.G., Sonnefeld, L.J., Cordray, D.S., Pion, G.M., & Perl, H.I. (1998). Constructing quantitative implementation scales from categorical services data. Examples from a multisite evaluation. Evaluation Review, *22*, 245-288.

Like measures of outcome, measures of implementation are most useful and analytically powerful when measured at client-level and are quantitative. However, high-quality, individual-level, quantitative service utilization data can be expensive, intrusive, or otherwise impractical to obtain. Cruder data—for example, presence versus absence of a given service—are often more feasible to collect, as well as more likely to be available. Consequently, evaluators can benefit by finding ways to better exploit such data at the analysis phase to compensate for shortcomings at the collection phase. This article documents one such instance in which this was done. Specifically, it describes how quantitative, client-level implementation scales were derived from qualitative (categorical) data and used to support a cross-site synthesis of implementation and outcome analyses in a multisite evaluation. It also suggests additional scenarios in which quantitative implementation scales might be derived from qualitative services data.

Orwin, R.G., Sonnefeld, L.J., Garrison-Mogren, R., & Smith, N.G. (1994). Pitfalls in evaluating the effectiveness of case management programs for homeless persons: Lessons from the NIAAA community demonstration program. Evaluation Review *18*, 153-207.

It is generally believed that homeless individuals—particularly those with substance abuse problems or mental illness—are unable to access the full range of available benefits and community services on their own. In recent years, community service providers have increasingly looked toward case management as the intervention of choice for solving this problem. Yet the

evaluation findings of the NIAAA community demonstration program, which included three case management projects, showed few clear indications of case management effectiveness. The results of these studies—which focused on homeless clients with alcohol and other drug problems—are consistent with studies of case management services for the homeless mentally ill. This article identifies several phenomena that potentially explain the apparent lack of positive effects and suggests that the reasons for negative findings may lie as much with the evaluations as with the interventions. These phenomena include bias due to differential attrition from measurement, lack of intervention intensity, lack of distinction between groups, contextual factors, and additional design and measurement issues. Suggestions for improving future evaluations of case management effectiveness are offered. (Author abstract)

Ross, J., Chipungu, S., Hermann, J., Peterson, J., et al. (1993). CSAP PPWI demonstration program findings. Final Report (Volume 1). Calverton, MD: Macro International.

Perinatal substance use is a major health concern throughout this country. Since 1989, the Center for Substance Abuse Prevention (CSAP) has funded numerous Pregnant and Postpartum Women and their Infants (PPWI) demonstration grant projects. As part of the PPWI demonstration initiative, CSAP conducted a cross-site evaluation of these programs. This volume is one of three final reports prepared under the cross-site evaluations of the PPWI Demonstration Program. This study reports findings from the implementation experience, achievement of client-level and system-level goals, as well as presenting a background of the CSAP PPWI demonstration program, evaluation design and methods, and project and participant characteristics. The report concludes with a discussion of conclusions and recommendations.

Seltzer, M.H. (1994). Studying variation in program success: A multilevel modeling approach. Evaluation Review 18, 342-361.

Ways in which multilevel modeling techniques can broaden the kinds of questions asked in multisite evaluations are illustrated through the example of the “Transmath” program for secondary schools. These techniques provide a means of moving beyond estimating effects to investigating how differences across sites relate to differences in success. (ERIC abstract modified)

Sinacore, J.M., & Turpin, R.S. (1991). Multiple sites in evaluation research: A survey of organizational and methodological issues. In R.S. Turpin, & J.M. Sinacore (Eds.), Multisite evaluations (New Directions for Program Evaluation 50, pp. 5-18). San Francisco: Jossey-Bass.

In order to conduct an evaluation at multiple sites, a balance is needed between the benefits and the challenges of structural, technical, and personal issues. (Author abstract)

Substance Abuse and Mental Health Services Administration. (1994). Cost-benefit issues in programs for prevention of alcohol and other drug use. Rockville, MD: NCADI.

Cost-benefit methodology has become a much higher priority area of investigation with the advent of health care system reform and with Federal support increasingly being directed toward programs that demonstrate cost-effectiveness. Cost-effectiveness studies of individual programs help build the case for prevention from successful efforts. Interventions among pregnant women, worksite EAPs, and others that have been subjected to cross-site evaluations are cited as proven successes. It is contended that the comparison of average per-person costs for primary prevention and substance abuse treatment is not an appropriate measure of cost-effectiveness, because relatively few drug and alcohol users receive treatment. The creation of a standardized taxonomy of prevention components and lists of direct community indicators of AOD abuse, identification of social psychology techniques that can be applied to measuring changes on the community level, and the imposition of rigorous evaluation criteria when the prevention program grant is awarded may be helpful in cost-effectiveness studies. (NCADI abstract modified)

Tushnet, N.C. (1995, April). Toward a general approach to multisite program evaluation. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA. (ERIC Document Reproduction Service No. ED 383 773)

Evaluations of multisite programs present special challenges to evaluators concerned with the use of evaluation results. This paper offers an approach that holds promise for mitigating the problems inherent in evaluating programs at multiple sites. A brief analysis is presented of the challenges that these programs present and the reasons these challenges exist. In sum, these challenges exist because each site implements a program differently. Two Federal education programs are described and used as examples throughout the text. A second section describes a general theoretical approach to evaluation that incorporates the works of Chen, Derrida, and others who emphasize the deconstruction of concepts and activities. The evaluation framework incorporates the elements of context, problem definition, program, funded projects, activities, and outcomes. (ERIC abstract modified)

U.S. Department of Health and Human Services. (n.d.). OSAP national evaluation of the community partnership demonstration program: Evaluation plan. Washington, DC: GPO.

This OSAP demonstration program is based on the tenet that key segments of the community affected by AOD abuse must be involved in the development and implementation of prevention programs. The central component of the 95 programs funded to date is a coalition of relevant organizations representing the public, health, criminal justice, education, voluntary, business, and private communities. Cross-site comparisons of the programs are designed to identify successful strategies as well as common inhibitors to the processes of forming effective partnerships and implementing prevention strategies. Assessing the effectiveness and impact of the prevention strategies also will be conducted across sites, with uniform data collection instruments leading to a cross-site analysis of prevention approaches. (NCADI abstract modified)

Weaver, F.M. (1994). Conducting simultaneous large multisite evaluations on parallel topics. Evaluation Practice 15, 37-44.

Simultaneous evaluations of the existing Department of Veterans' Affairs pharmaceutical distribution system and an alternative "prime vendor" distribution concept in the private sector were conducted. This review of the methodology illustrates the problems and advantages in conducting large multisite evaluations simultaneously. (ERIC abstract)

### **III. SELECTED BIBLIOGRAPHY**

### III. SELECTED BIBLIOGRAPHY

- Aktan, G.B. (1998). Evolution of a substance abuse prevention program with inner city African-American families. Drugs and Society, 12(1/2), 39-52.
- Conrad, K.J., Hultman, C.J., & Lyons, J.S. (Eds.). (1993). Treatment of the chemically dependent homeless: Theory and implementation in fourteen American projects. Binghamton, NY: Haworth Press.
- Cousineau, M.R., & Wittenberg, E. (1995). Study of the health care for the homeless program: Executive summary. Bethesda, MD: HRSA.
- Orwin, R.G. et al. (1993). Community demonstration grant projects for alcohol and drug abuse treatment of homeless individuals: Executive summary. Rockville, MD: NIAAA.
- Orwin, R.G., Goldman, H.H., Sonnefeld, C.J., Ridgely, M.S., et al. (1994). Alcohol and drug abuse treatment of homeless persons: Results from the community demonstration program. Journal of Health Care for the Poor and Underserved, 5(4), 326-352.
- Skogan, W.G., & Lurigio, A.J. (1991). Multisite evaluations in criminal justice settings: Structural obstacles to success. In R.S. Turpin, & J.M. Sinacore (Eds.), Multisite evaluations (New Directions for Program Evaluation 50, pp. 83-96). San Francisco: Jossey-Bass.

## **IV. BENEFITS AND LIMITATIONS OF CROSS-SITE EVALUATIONS**

#### IV. BENEFITS AND LIMITATIONS OF CROSS-SITE EVALUATIONS

CITATION	BENEFITS	LIMITATIONS	NOTES
Conrad, Hultman, & Lyons (1993)	<ul style="list-style-type: none"> <li>• The cross-site allows for answering questions about program efficacy in the most scientifically rigorous manner possible.</li> </ul>	<ul style="list-style-type: none"> <li>• While programs profiled were well-planned and conceptualized, the programs need to be allowed time for project start-up.</li> <li>• Required interventions need to be of sufficient duration to allow for meaningful data collection.</li> <li>• Plans for tracking and follow-up should be developed before the first client is assessed.</li> </ul>	<ul style="list-style-type: none"> <li>• This book concerns the 14 NIAAA-funded Cooperative Agreement program for homeless, AOD-using people.</li> <li>• Summary was obtained from the first chapter but concerns all the programs profiled.</li> </ul>
Freedman (1991)	<ul style="list-style-type: none"> <li>• With the availability of large, multisite databases, there is a:               <ul style="list-style-type: none"> <li>- Shift to national focus of health services</li> <li>- Wealth of data in health care sector and national data organizations to facilitate access to the data</li> <li>- Designated data-processing staff.</li> </ul> </li> <li>• Benefits of multisite evaluations (MSEs) include:               <ul style="list-style-type: none"> <li>- History of multisite clinical trials in biomedical research</li> <li>- Highly specialized/trained professionals in health care (including the medical profession)</li> <li>- Body of evidence that treatment varies across health care settings and regions</li> <li>- Demand for uniform health care standards</li> <li>- Growing number of multi-institutional health care systems.</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• Large-scale MSEs are expensive to conduct.</li> <li>• Researchers may experience difficulty in gaining the cooperation of (diverse) programs that are frequently geographically dispersed and under a variety of organizational leadership structures.</li> <li>• Reliability and validity of data from national organizations' databases are not guaranteed.</li> <li>• Researchers may encounter difficulty in obtaining a large enough sample size for some "rare" event.</li> <li>• Some professionals may resist "soft" research methods (e.g., surveys, focus groups, psychometric scales, etc.).</li> </ul>	<ul style="list-style-type: none"> <li>• What type of MSE?               <ul style="list-style-type: none"> <li>- Studies that depend solely on secondary data sources for evaluation of policies and programs</li> <li>- Studies that depend on the labor of existing program staff for data collection</li> <li>- Studies that combine secondary data sources with original data collection</li> <li>- Studies that primarily depend on original, evaluation-specific data collection.</li> </ul> </li> </ul>

CITATION	BENEFITS	LIMITATIONS	NOTES
Guastello & Guastello (1991)	<ul style="list-style-type: none"><li>• Multisite evaluations (MSEs) are highly recommended:<ul style="list-style-type: none"><li>- To generalize program outcomes over organizational differences</li><li>- To generalize conclusions over many possible examples of a hypothetical program</li><li>- To evaluate the equivalence of controls.</li></ul></li></ul>	<ul style="list-style-type: none"><li>• MSEs can be conducted to detect the limits to generalizability or the interaction between organizational characteristics and program features.</li></ul>	<ul style="list-style-type: none"><li>• Overall, the entire matter of the impact of organizational characteristics on program outcomes is under-researched.</li></ul>
Lamberti & Katzenmeyer (1996)	<ul style="list-style-type: none"><li>• The use of template methodology allows for identifying patterns and cross-cutting themes across programs.</li><li>• Templates allow for quantitative analysis of qualitative data.</li></ul>	<ul style="list-style-type: none"><li>• The structured nature of the templates may hamper richness of comments during analysis.</li></ul>	<ul style="list-style-type: none"><li>• Future efforts may involve training site visitors to comment as necessary and not to be hampered by the structured process.</li></ul>

CITATION	BENEFITS	LIMITATIONS	NOTES
<p>Mowbray &amp; Herman (1991)</p>	<ul style="list-style-type: none"> <li>• Multisite evaluations (MSEs) offer many advantages in the development of program initiatives; they:               <ul style="list-style-type: none"> <li>- Permit testing of contextual effects</li> <li>- Allow opportunities for positive, local adaptations</li> <li>- Provide applicable models for program development.</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• The advantages can also present major difficulties because they may complicate ensuring standards of scientific rigor and useable results and perhaps foster negative outcomes and/or suspicion that the (national) model was not followed.</li> <li>• Site characteristics, while understood to be contextually different, may have greater variety and impact than initially expected.</li> <li>• Much time must be spent on education, training, motivation, and data quality control with program implementors when utilizing program models in MSEs.</li> <li>• Current literature is concerned with measuring application at the individual subject level rather than at the program level.</li> </ul>	<ul style="list-style-type: none"> <li>• This article offers a modified research, development, and diffusion approach to measuring the application at the program level.</li> <li>• In order to overcome limitations, researchers should:               <ul style="list-style-type: none"> <li>- Adhere to program theory in order to accurately judge the effects of the implementing organization's characteristics on the process, amount, and type of services delivered (and other variables)</li> <li>- Conduct a process evaluation since it is a logical step from theory into practice; a process evaluation is a critical factor in assessing the validity of the intervention</li> <li>- Gather careful program- and client-level data since the methods for doing so have not been addressed in the literature</li> <li>- Modify program-level data collection techniques carefully since the literature does not emphasize program theory or implementation measurement</li> <li>- Utilize a concept map in order to maintain fidelity to the program model.</li> </ul> </li> </ul>

CITATION	BENEFITS	LIMITATIONS	NOTES
<p>Orwin, Sonnefeld, Garrison-Mogren, &amp; Smith (1994)</p>	<ul style="list-style-type: none"> <li>• If programs utilize the following suggestions, multisite evaluations (MSEs) could improve understanding of the reasons behind the “negative findings.” The programs profiled should incorporate:                             <ul style="list-style-type: none"> <li>- Stronger designs</li> <li>- Better follow-up</li> <li>- Longer program and follow-up duration</li> <li>- Improved outcome measurement</li> <li>- Integrated process and outcome evaluation</li> <li>- Development and use of program theory.</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• As a result of the multisite evaluation, case management effectiveness was found to be lacking, based on the points that:                             <ul style="list-style-type: none"> <li>- Case management is not the key ingredient</li> <li>- CM model is based on faulty assumptions</li> <li>- Tasks are too complicated to be carried out by case managers on behalf of severe mentally ill (SMI) patients.</li> </ul> </li> <li>• “Poor” data may be due to:                             <ul style="list-style-type: none"> <li>- Bias from differential attrition (those who are doing well are harder to track)</li> <li>- Lack of sufficient intervention intensity</li> <li>- Lack of distinction between groups (all received some form of CM)</li> <li>- Contextual factors related to the network of services</li> <li>- Low statistical power and selection bias</li> <li>- Measurement issues.</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• While focusing on the phenomena behind the reasons for negative findings, the article asks, “Does case management (CM) work?”</li> <li>• Three NIAAA-funded CM projects are profiled: Boston, Louisville, and Minneapolis.</li> <li>• Data were obtained from the ASI, treatment-housing forms, products from national process evaluations, and projects’ final reports.</li> <li>• The results are restricted to reduced AOD use, increased employment/economic security, improved mental/physical health status, increased residential stability.</li> <li>• The results are measured in terms of the likelihood and magnitude of improvement for each of the above factors.</li> </ul>

CITATION	BENEFITS	LIMITATIONS	NOTES
Seltzer (1994)	<ul style="list-style-type: none"> <li>• When the program participant is viewed as the key unit of analysis, discussions about the units of analysis are no longer required to be single units, due to the recent development of multilevel analytic techniques.</li> <li>• Such a technique is termed hierarchical linear modeling (HM) and involves with-in site (level 1) and between-site (level 2) analyses.</li> <li>• Some variation in program effects seen in a multisite evaluation (MSE) may be due to unique contextual factors and events.</li> <li>• Differences in program successes are to some extent systematically related to differences in certain facets of implementation and certain site characteristics.</li> <li>• Program sites are likely to differ in terms of client compositional characteristics, various ways of implementing intervention techniques, program support from principals may vary, etc.</li> <li>• Models are needed in MSEs to reflect the location of program participants within different sites and to allow for the effects of a program's substantial differing across sites.</li> <li>• Analysis of field notes and transcribed interviews help researchers find differences in program philosophy, organizational structure, thus helping to locate "outliers," or those programs that deviate markedly from the patterns of relationships discerned via the between-site models.</li> </ul>	<ul style="list-style-type: none"> <li>• False impression that evaluators working in the qualitative tradition are insensitive to the program context.</li> <li>• Using the student, for example, as the unit of analysis conceals variation in program effectiveness.</li> <li>• Restricting analyses to context means may result in a loss of power and can place limitations on the kinds of questions able to be addressed.</li> </ul>	<ul style="list-style-type: none"> <li>• The article asks the general question, "For an outcome of interest, how much of a difference might we expect between those who participate in a particular program or intervention versus those who do not?"</li> </ul>

CITATION	BENEFITS	LIMITATIONS	NOTES
Sinacore & Turpin (1991)	<ul style="list-style-type: none"><li>• Advantages of multisite evaluations (MSEs) include increased statistical power and increased validity and reliability due to increased sample size.</li><li>• Evaluation of programs at multiple sites is advantageous when generalizing program effects across a diverse range of individuals.</li></ul>	<ul style="list-style-type: none"><li>• MSE analysis is limited to original data.</li></ul>	<ul style="list-style-type: none"><li>• This chapter discusses the issues that investigators face when conducting MSEs.</li></ul>
Skogan & Lurigio (1991)		<ul style="list-style-type: none"><li>• Some agencies may be inherently difficult to evaluate via MSEs.</li></ul>	<ul style="list-style-type: none"><li>• This chapter describes obstacles to MSEs and reviews two projects from the CJ system.</li></ul>

**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.**