

# NEDS

NATIONAL EVALUATION DATA SERVICES

## COMPARISONS OF CLIENT OUTCOMES ACROSS CITIES: DOES CASE MIX MATTER?

October 2002



**CSAT**  
Center for Substance  
Abuse Treatment  
SAMHSA

  
*The* LEWIN GROUP

# NEDS

NATIONAL EVALUATION DATA SERVICES

## **COMPARISONS OF CLIENT OUTCOMES ACROSS CITIES: DOES CASE MIX MATTER?**

**Prepared by**

**Ansari Z. Ameen  
Jennifer L. Kowalski  
Christel M. Villarivera  
Henrick J. Harwood**

**The Lewin Group, Inc.  
3130 Fairview Park Drive, Suite 800  
Falls Church, VA 22042**

**October 2002**

This document was supported by the Center for Substance Abuse Treatment, Department of Health and Human Services, Caliber/NEDS Contract No. 270-00-7078. 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.

**CSAT**  
Center for Substance  
Abuse Treatment  
SAMHSA

# TABLE OF CONTENTS

Page

**FOREWORD**

**ACKNOWLEDGMENTS**

**ABSTRACT**

**EXECUTIVE SUMMARY..... i**

**I. INTRODUCTION..... 1**

1. INTRODUCTION TO AND USE OF PERFORMANCE MEASUREMENT ANALYSIS ..... 1

2. CONTEXT/RATIONALE FOR THE ANALYSIS ..... 4

3. PRIOR RESEARCH/ONGOING ASSESSMENTS OF TREATMENT SYSTEM PERFORMANCE ..... 6

4. OVERVIEW OF THE ANALYSIS ..... 10

5. ORGANIZATION OF THIS REPORT ..... 11

**II. METHODS ..... 13**

1. THE TARGET CITIES DATA SET ..... 13

1.1 Data Source ..... 13

1.2 Analysis Database ..... 15

1.3 Exclusion Criteria..... 16

2. ANALYTIC METHODS/APPROACH..... 17

2.1 Analytic Approach ..... 17

2.2 Model Specification ..... 20

## TABLE OF CONTENTS (CONT.)

	<u>Page</u>
<b>III. FINDINGS .....</b>	<b>27</b>
1. SAMPLE DESCRIPTION .....	27
1.1 Client-level Demographics.....	29
1.2 Criminal Justice History.....	29
1.3 Primary Substance of Use .....	29
1.4 Type of Treatment.....	30
1.5 Substance Use Days/Substance Use Problem Days.....	31
1.6 Standardized Performance Measurements .....	32
2. REGRESSION RESULTS.....	33
2.1 Findings By Type of Treatment .....	34
2.2 Findings By Primary Substance of Use.....	35
2.3 Summary .....	36
<b>IV. SUMMARY AND IMPLICATIONS.....</b>	<b>39</b>
1. SUMMARY OF FINDINGS .....	39
2. IMPLICATIONS FOR TREATMENT PROVIDERS .....	42
3. IMPLICATIONS FOR POLICYMAKERS.....	43
4. IMPLICATIONS FOR RESEARCHERS/EVALUATORS.....	44
5. CONSTRAINTS OF FINDINGS.....	45
<b>REFERENCES.....</b>	<b>49</b>
<b>APPENDIX A REGRESSION MODEL SPECIFICATIONS</b>	
<b>APPENDIX B REGRESSION FINDINGS</b>	

## FOREWORD

The mission of the Center for Substance Abuse Treatment (CSAT), Substance Abuse and Mental Health Services Administration (SAMHSA), is to improve the lives of individuals and families affected by alcohol and drug abuse by ensuring access to clinically sound, cost-effective addiction treatment that reduces the health and social costs to our communities and the nation. As part of its mission, CSAT supports the development of innovative treatment approaches, based on sound data and state-of-the-art analyses, and disseminates information on treatment approaches shown to be effective for curbing addiction and related behaviors.

In 1997, CSAT established the National Evaluation Data Services (NEDS) contract to support the CSAT mission. In 2000, through a new contract (Contract No. 270-00-7078), CSAT continued and expanded the scope of NEDS. NEDS activities help to foster collaboration and partnering among the public and private sectors along the Federal-state-local community-based treatment continuum. The three major activities of NEDS, under the current contract, are to assist in developing data infrastructure vehicles and tools, to perform treatment services secondary analyses on existing data, and to support the Government Performance Results Act (GPRA) activities. NEDS, through its Secondary Analysis Technical Reports, provides evidence-based information on substance abuse treatment issues relevant to treatment needs, access, utilization, efficacy, effectiveness, and efficiency. NEDS analyses focus on treatment needs, services received, and populations of interest to the substance abuse treatment field in order to provide new information about which services yield the best outcomes for what types of clients, at what cost. This information helps address treatment issues such as the treatment gap, culturally competent treatment services, and recovery.

This analysis aims to compare the performance of the treatment systems located within cities in different geographic regions, an approach to performance measurement and outcomes monitoring that has not previously been done. Using the CSAT Target Cities multi-site data set, the goals of this analysis were to (1) identify and operationalize standardized performance measures that are applicable across treatment systems within cities; (2) illustrate a type of performance measurement analysis that can be conducted at the treatment system level; (3) assess the relative performance of the treatment systems/cities using the identified performance measures; and (4) highlight the uses, benefits, constraints and data requirements for system-level performance measurement monitoring and comparisons. The value of performance measurement analyses lies in their potential to identify the providers or treatment systems with better or worse client outcomes in order to highlight "best practices" that might be implemented across treatment

systems, or to identify aspects of care that should be targeted for improvement through technical assistance and other improvement strategies. Results from these types of analyses also hold the potential to contribute to decisions regarding the allocation and levels of funding for treatment systems or regions.

Patrick J. Coleman  
Project Director  
National Evaluation Data Services (NEDS)

## **ACKNOWLEDGMENTS**

We wish to acknowledge our reliance upon the overall guidance and direction of Ron Smith, the Center for Substance Abuse Treatment (CSAT) Government Project Officer for the NEDS contract. Caliber Associates is the prime contractor for NEDS in partnership with DeltaMetrics, The Lewin Group, the National Association of State Alcohol and Drug Abuse Directors (NASADAD), the National Development and Research Institutes (NDRI), the National Opinion Research Center (NORC), Science Applications International Corporation (SAIC), and UCLA Integrated Substance Abuse Programs (ISAP).

We wish to thank Dr. Susan Ettner, University of California, Los Angeles, for her valuable and insightful comments on an earlier draft of this paper. Thanks are also due to Substance Abuse and Mental Health Services Administration staff members who reviewed and commented on the earlier draft. We wish to acknowledge the project team for the Target Cities multi-site initiative, specifically, Dr. Richard Stephens and Heather Huff at the University of Akron, and Randolph Muck, the CSAT Government Project Officer. We also wish to acknowledge the Target Cities grantees and evaluators who contributed data to the Target Cities multi-site database.

Many individuals on the NEDS team contributed to this report through content and editorial reviews and final document preparation. Special thanks go to Michael Wolf-Branigin, Mary Layne, Beverly Hitchins, Marsha Morahan, Sharyn Berg, and Iris Mensing.

## **ABSTRACT**

This analysis compares the performance of the substance abuse treatment systems located within cities in different geographic regions. Performance is assessed based on the outcomes of clients initiating treatment, an approach to performance measurement and outcomes monitoring that has not previously been done. The analysis uses data from the CSAT Target Cities multi-site evaluation initiative and is based on a sample of 3,642 clients from six of the cities that participated in the second phase of the Target Cities initiative. Clients were categorized by both type of treatment and primary substance of use and a total of 11 outcome measures in the domains of drug/alcohol use, employment, and criminal activity were analyzed. Multivariate analyses were performed for each outcome measure for each category of client, using Ordinary Least Squares and/or Logistic Regression, as appropriate. The models included measures to adjust for non-treatment factors (e.g., client characteristics and severity of presenting symptoms at intake) that other studies have shown can influence post-treatment outcomes. Differences in the performance measures were compared across the cities before and after adjusting for client characteristics and severity measures. Results from the adjusted comparisons show consistent patterns across cities by both type of treatment and drug type. While most of the cross-city comparisons did not find statistically significant differences, Cities 4 and 6 appeared to perform less well than the reference city (City 1). Also, Cities 2 and 5 appeared to perform better relative to the reference city, though there were no apparent differences with respect to the performance of City 3. Findings from this analysis have implications for research, policy and practice, and highlight the uses, benefits, constraints, and requirements for performing such analyses at the treatment system level.

## **EXECUTIVE SUMMARY**

# EXECUTIVE SUMMARY

## 1. INTRODUCTION

This analysis complements and extends research conducted in the performance measurement arena within the substance abuse treatment field. There has been much recent attention focused on issues of accountability and cost containment in the substance abuse treatment field. In response, a range of performance measurement analyses have emerged that primarily attempt to compare the relative performance of specific treatment providers. Increasingly, more attention is being directed to systems of care as opposed to individual provider performance. The managers of treatment systems are interested in knowing the types, duration, intensity, and cost/funding levels of treatment that are most effective. Given the importance of identifying effective management strategies, this analysis uses quantitative measures to assess the relative performance of substance abuse treatment systems operating within six cities. Using the Center for Substance Abuse Treatment (CSAT) Target Cities multi-site evaluation database, the goals of this analysis were to identify and operationalize standardized performance measures that are applicable across treatment systems and cities; illustrate one approach to assessing and comparing the performance of substance abuse treatment systems; and highlight the utility, implications, constraints and data requirements for system-level performance measurement monitoring and comparisons.

## 2. METHODS

This analysis used data collected from six cities participating in the second phase of the CSAT Target Cities initiative. The analytic approach involved identifying performance measures in multiple outcome domains that are of interest to various stakeholders, conducting a multivariate analysis that uses case mix adjustment to partial out the effects of non-treatment influences on clients' outcomes, and examining the relative performance of cities for meaningful categories of clients. The analysis used data for 3,642 clients from six cities. Data were analyzed across five different types of treatment (non-intensive outpatient, intensive outpatient, methadone treatment, and short-term and long-term residential care) and four primary substances of use (alcohol, heroin, cocaine/crack, and marijuana). Outcome domains were selected based on their importance to policymakers and stakeholders. The outcome domains included drug/alcohol use, employment, and criminal justice involvement. A total of 11 performance measures were constructed across the three outcome domains: eight in the area of drug/alcohol, two in the employment outcome domain, and one in the criminal justice outcome domain. Descriptive and multivariate analyses were conducted in order to assess the performance of the cities before and after adjusting for differences in client demographics and the severity of

presenting symptoms across treatment populations. The cities were compared on each performance measure using either Ordinary Least Squares or Logistic Regression, as appropriate, across types of treatment and primary substance of use.

### **3. FINDINGS**

Across all cities, a total of 35 percent of clients received services in outpatient levels of care (20% in intensive outpatient and 15% in non-intensive outpatient). There was considerable variability across the cities, however, in terms of the proportions of clients served in inpatient versus residential treatment. Approximately 30 percent of clients across all cities were treated in inpatient settings (15% and 16%, respectively, for short- and long-term residential). The compositions of the populations receiving services in the six cities differed in gender, race, marital status, and the severity of their presenting symptoms. Thus, Cities 1, 2, and 3 had clients who reported more severe substance abuse problems than clients in Cities 4, 5, and 6. Clients also differed in the severity of their employment problems and criminal histories across cities. Examination of the average client outcomes, collectively, as well as across the cities, showed that clients improved on key measures of drug/alcohol use and problems as well as their employment status. Preliminary comparisons were made using performance measures in each outcome domain that were standardized for initial level of severity on the corresponding measure. No consistent patterns across outcome domains were observed for the cities; however, these initial comparisons did not standardize for client characteristics or severity on other measures.

Multivariate analyses were used, therefore, to examine the relative performance of the cities, given the variability in both client characteristics and severity at intake and clients' outcomes at follow-up. Models were constructed that allowed for the comparison of clients across cities by type of treatment and by primary substance of use. Generally, the cities did not differ significantly from the comparison city (City 1) on the majority of the performance measures. There were, however, some noteworthy differences. Clients receiving intensive outpatient treatment in City 4 had poorer substance abuse outcomes than clients in City 1 on four of the eight measures. This relatively lower than average performance is consistent with the descriptive findings which showed that City 4 had relatively small improvements in substance use compared to the other cities. Clients in City 5 in non-intensive outpatient treatment and short-term residential treatment had better substance use outcomes than clients treated in City 1.

While the findings described above also are consistent with the descriptive statistics, findings for City 6 illustrate the need for multivariate analyses when comparing treatment systems' performance. Though the descriptive statistics suggested that City 6 had better

performance with regard to employment, the multivariate analysis showed that clients in non-intensive outpatient treatment in City 6 actually had significantly less desirable outcomes than City 1 for both employment measures. Also, clients using heroin in City 3 had significantly better criminal activity outcomes. This finding also supports the use of multivariate analyses because the descriptive statistics initially indicated that City 3 had the highest proportion of clients with arrests/charges prior to follow-up.

In general, Cities 2 and 5 had better post-treatment outcomes, although this was slightly less consistent for City 2. Across all substances of use, Cities 2 and 5 each had more favorable outcomes for nine measures. City 3 proved to be the least consistent of all cities. While City 3 had more favorable performance for three outcomes across types of treatment, it also had less favorable performance for two outcomes.

#### **4. SUMMARY AND IMPLICATIONS**

In summary, this analysis reveals differences across cities with regard to the characteristics of the treatment populations served, the severity of the clients' presenting problems, and clients' outcomes on 11 performance measures. The multivariate analysis shows relative consistency across the different formulations of outcome measures. Despite the relative consistency of findings, the descriptive and multivariate results nevertheless illustrate the importance of using multivariate analyses to examine differences in client case mix across treatment populations when making assessments of relative performance. The analyses were intended to be illustrative of an analytic approach to system-level comparisons of performance rather than a vehicle that allows an analyst to make definitive statements about the specific cities included in the analysis. More specifically, given the data constraints associated with deriving performance measures and appropriate case covariates to control for significant non-treatment related factors, the results should be interpreted with caution.

The constraints of this analysis are related to the absence of certain services, provider, and community variables needed to develop fully specified models to address questions pertaining to why the clients from some cities appeared to have better outcomes than others. The analysis nevertheless has implications for treatment providers, policymakers, and researchers/evaluators. The results illustrate a novel approach to comparing the performance of treatment systems and further highlight the associated challenges and implications. In particular, the analysis points to the types of issues (e.g., performance measure identification and data collection of theoretically relevant case mix controls) policymakers could consider in designing outcome monitoring systems and evaluations of multi-site demonstrations. The analysis also suggests areas for further data collection and analysis, such as performance measure validation;

methodological issues related to combining performance measures for the development of global measures of system performance; and identification of a limited, parsimonious set of performance measures and case mix controls that could be used for comparing the performance of treatment systems.

## **I. INTRODUCTION**

# I. INTRODUCTION

This introduction is intended to orient readers to issues related to performance measurement in the substance abuse treatment field and the growing use of performance measurement analysis to examine differences in substance abuse treatment providers and treatment systems using post-treatment client outcomes. Specifically, this introduction will familiarize readers with the purpose and organization of this report; provide the context for the analysis, including a brief outline of the emergence of performance measurement analysis as a tool for monitoring substance abuse treatment provider and system performance through the study of clients' outcomes; and provide a review of the relevant literature, though there is little prior research directly applicable to this treatment systems-level analysis.

## 1. INTRODUCTION TO AND USE OF PERFORMANCE MEASUREMENT ANALYSIS

Performance measurement analyses can be used to measure and examine treatment provider-level as well as system-level performance along various dimensions, including provider processes and outcomes. Provider-level performance measurement analysis can be a useful tool to treatment services managers and system administrators, since results from such analyses can identify providers with either above or below average performance. That information is invaluable in efforts aimed at continuous quality improvement, best practices identification, and the development of performance-based contracting systems. System-level performance measurement permits comparisons of the performance of treatment systems located in differing geographic regions and also holds the potential to monitor and track changes in the national treatment system. For example, routine monitoring of data across treatment systems can identify the types of clients accessing treatment, those presenting for treatment, the nature of their presenting problems, services received and treatment outcomes.

Process performance measures refer to various aspects of the manner or frequency with which services are provided (e.g., number of screenings, counseling sessions, etc.). In contrast, this analysis is a performance measurement analysis that uses treatment outcomes as the performance measure. Examples of performance measures based on clients' outcomes in the substance abuse arena include measures of clients' substance use, employment status and arrest record following treatment. Independent variables, sometimes referred to as case mix controls, are used to statistically control for differences across providers or treatment systems with respect to the socio-demographic characteristics, severity and readiness of clients when they present for treatment. These variables are included in order to statistically control for (i.e., separate out) the effect of factors that influence treatment outcomes, but that are generally beyond the influence of

treatment providers. The value of performance measurement analyses lies in their potential to identify the providers or treatment systems with better or worse client outcomes in order to highlight "best practices" that might be implemented across treatment systems, or to identify aspects of care that should be targeted for improvement through technical assistance and other improvement strategies. Results from these types of analyses also hold the potential to contribute to decisions regarding the allocation and levels of funding for treatment systems or regions.

This report presents findings from an analysis that illustrates one approach to comparing the performance of substance abuse treatment systems within different cities. This analysis is based on multi-site data from the Center for Substance Abuse Treatment (CSAT) Target Cities initiative. The treatment systems of cities that participated in this initiative are typically located within the major metropolitan areas in these cities. These treatment systems primarily include substance abuse treatment providers that offer services at multiple levels of care. These treatment systems also include related social service providers (e.g., welfare offices, educational/vocational institutions, social service agencies, etc.) that provide additional services needed to address the multiple needs of clients presenting for treatment. In the remainder of this report, the treatment systems located within these metropolitan areas will be referred to simply as "cities."

This analysis of relative substance abuse treatment system performance is described as illustrative because no systematic comparisons of treatment system performance have been conducted to date, despite the emerging need for such analyses. The lack of research on this topic stems primarily from a dearth of available standardized and reliable information at the state and regional levels for developing "census-based" estimates to track changes in the systems' performance across systems/regions or over time. CSAT's Cooperative Agreements for State Treatment Outcomes and Performance Pilot Studies Enhancement: TOPPS I & II initiatives were created in direct response to the need for information to address treatment system performance at the state and regional, if not national levels. These initiatives created grants to states to provide funding for the design of special studies aimed at refining their outcome monitoring systems (OMS) or for the conduct of special performance measurement studies.

CSAT's planned implementation of a National Treatment Outcomes Monitoring System (NTOMS), aimed at monitoring the status of the national treatment system, is also intended to address this need. NTOMS, as currently outlined, will be based on a sample drawn from the National Survey of Substance Abuse Treatment Services (NSSATS) of the Office of Applied Studies (OAS) of the Substance Abuse and Mental Health Services Administration (SAMHSA). While data from NTOMS will be amenable to monitoring changes in the national treatment system, the data, as currently conceptualized, will not permit assessments of how the treatment

systems of particular states or geographic regions are performing over time or relative to one another (Department of Health and Human Services, 2002). Such information is of critical importance for identifying and implementing enhancements to treatment systems to improve services and clients' outcomes.

The majority of the previous performance measurement work in the substance abuse arena has been conducted at the individual treatment provider level, comparing providers in limited geographic regions or within states. This report differs from previous analyses in that it uses the treatment systems within cities as the unit of analysis. The analytic approach and findings from this study are expected to complement and extend the growing body of literature in the substance abuse treatment field related to performance measurement, continuous quality improvement, and performance contracting.

This analysis quantitatively examines the relative performance of six cities using performance measures (i.e., specific quantitative outcome measures/variables collected from clients) in the client outcome domains of drug/alcohol use, employment and criminality. The approach identified the most appropriate performance measures in these three domains of client functioning, assessed the performance of each city using these measures, and examined the relative performance of the cities using these measures. The results allow for the determination of whether particular cities appeared to be either over- or under-performing relative to the others.

This analysis illustrates how performance measurement analyses might be conducted using data from states, cities, or other geographic regions. Such analyses are expected to become possible in the future as the refinements of states' OMS are implemented. This would allow the collection of standardized data and the opportunity for system-level analyses. The report and analysis also highlight the basic data and OMS requirements needed to engage in systematic analyses of different treatment systems. In addition to demonstrating a single methodological approach to the comparative analysis of differing cities, this analysis points to several data-related challenges. These challenges illustrate the need for careful forethought and systematic long-range planning when designing and implementing demonstration initiatives and OMS changes. Finally, the analysis also calls attention to the need for coordination at the provider, community, state, regional, and national levels in order to facilitate systematic evaluations aimed at identifying lessons from successes and setbacks.

Results from this novel analysis on an issue of emerging importance (namely, monitoring and improving the overall performance of treatment systems) will be of primary interest and relevance to policymakers, as it demonstrates how data can be manipulated to generate information that can inform systemic policy decisions. The results and approach will serve to

guide and stimulate further thinking on the part of researchers, evaluators, and analysts charged with making evidence-based recommendations to policymakers. Substance abuse providers and counselors will also find the approach useful. This is the case because the analysis underscores some of the criteria that are likely to be used to evaluate the performance of treatment systems and clinics and influence funding decisions.

## **2. CONTEXT/RATIONALE FOR THE ANALYSIS**

Performance measurement analysis is especially timely and policy-relevant given the increased attention to issues of fiscal accountability and responsibility. The increasing requirements for accountability related to the expenditure of funds for publicly funded treatment, as well as the imperatives of containing costs, have contributed to increased emphasis on the following:

- Rigorous evaluations of the outcomes of substance abuse treatment to bolster arguments to policymakers that treatment is effective
- Assessments of how effective substance abuse treatment providers are improving the individual and social functioning of their clients (e.g., reduced drug use, increased employment, and lowered criminality)
- Analyses aimed at identifying the costs and effectiveness of alternative approaches and types of substance abuse treatment
- Analyses aimed at creating baseline measures of treatment system performance so that systems can be improved over time.

This analysis is one of the first contributions to the last of these four major lines of inquiry. Namely, it is an investigation of baseline measurements and analyses aimed at comparing and improving the performance of treatment systems.

Research on these issues has become increasingly important because the Federal government is encouraging the movement toward outcomes monitoring through initiatives such as the Government Performance and Results Act (GPRA) of 1993 and the Performance Partnership Block Grant Legislation. GPRA is the Federal government's effort, under the aegis of the Office of Management and Budget, to establish measurable performance goals to be reported as part of the budgetary process, thus linking Federal allocations with performance. One of the findings that spurred Congress to pass the Act was that "congressional policymaking, spending decisions and program oversight are seriously handicapped by insufficient attention to program performance and results" (GPRA, Public Law 103-62). Thus, among its purposes,

GPRA is intended to “improve Federal program effectiveness and public accountability by promoting a new focus on results, service quality, and customer satisfaction” (GPRA, Public Law 103-62). The Performance Partnership Block Grant legislation calls for the Secretary, in conjunction with States, to develop a “common set of performance measures that would be used for accountability” (Children’s Health Act of 2000, Public Law 106-310). In response, many states are refining and improving their outcome monitoring systems in order to more effectively evaluate the performance of the treatment providers with whom they contract and the performance of the treatment systems they operate. These efforts will allow states to better address accountability questions posed by various stakeholders including state and Federal payers, local community leaders, and individuals and families accessing services.

SAMHSA’s Center for Substance Abuse Treatment has made a variety of resources available to states for furthering the development of their outcome monitoring (CSAT’s Treatment Improvement Protocol [TIP] Series 14) and performance measurement systems (Cooperative Agreements for State Treatment Outcomes and Performance Pilot Studies Enhancement: TOPPS I & II). For example, under the CSAT-sponsored TOPPS I and TOPPS II, awards were made to 14 and 19 states, respectively, to conduct state performance and substance abuse treatment outcome pilot studies. The Single State Agencies (SSAs) responsible for providing publicly funded substance abuse treatment services in each state managed the pilot studies. The studies allowed states to propose innovative, state-specific strategies to measure, evaluate, and improve substance abuse treatment while simultaneously making enhancements to the infrastructure of their OMS. States funded under these initiatives have conducted a variety of studies that facilitate the planning and development of the infrastructure needed for collecting outcome monitoring data. They have also simultaneously conducted pilot studies on a variety of timely issues. Among the substantive issues addressed under TOPPS I are the impact of managed care initiatives on substance abuse treatment effectiveness, the usefulness of assessment using versions of the Addiction Severity Index (ASI), and the impact of culturally sensitive counseling on special populations. The goal of these state-specific studies is to guide the expansion and application of the studies to the entire state treatment system and to incorporate these strategies into the state and national databases.

Nineteen states collect client data under the TOPPS II initiative. They submit data for a comparable core set of elements collected on all clients at critical treatment points, including intake, discharge, and follow-up. Theoretically, depending on the representativeness of the individual state’s data, the cross-state data could be pooled and used to examine broader issues such as the collective or comparative performance of several states or regions. Under the TOPPS II initiative, CSAT has recommended that states employ case mix adjustment methods when conducting cross-provider analyses. Such a recommendation was necessary, given the growing

body of literature in the substance abuse treatment field that demonstrates that comparisons of relative provider performance should, at a minimum, account for differences in the severity of clients' presenting symptoms (i.e., client case mix). Thus, both the within-state studies and the across-state studies must explicitly recognize and grapple with the issue of adjusting or statistically controlling for client differences (case mix) across providers (and treatment systems) that are beyond the immediate influence of providers and the therapeutic services they deliver.

This analysis illustrates an approach to establishing baseline measures of treatment systems' performance that can be used to monitor changes over time. It also demonstrates how different treatment systems can be examined to determine how they are performing relative to one another using standardized measures of performance. The section that follows summarizes the limited previous research and discusses ongoing efforts to assess the relative performance of substance abuse treatment systems.

### **3. PRIOR RESEARCH/ONGOING ASSESSMENTS OF TREATMENT SYSTEM PERFORMANCE**

There is little empirical literature in the substance abuse treatment field that directly aims to assess and compare the performance of entire substance abuse treatment systems using client outcomes. However, much "process" evaluation is conducted through efforts such as the Health Plan Employer Data and Information Set (HEDIS) and other similar managed care report cards. This analysis is therefore best described as illustrating one approach to comparisons of treatment system performance.

Regardless of the lack of available empirical work that directly addresses this topic, a review of the small but growing body of literature on performance measurement is informative. The literature review revealed an apparently disparate body of research that, when viewed as a collective, highlights the types of factors/issues that must be considered when assessing the performance of treatment systems.

The principal challenge underlying any quantitative performance measurement analysis is to statistically separate the effects of the therapeutic services clients receive during treatment (and their impact/effect) from influences of client characteristics and severity or aspects of the environment that may also affect clients' outcomes. This is necessary given that the goal of these analyses is to answer the question: "All else being equal, were the services and approach to treatment used by one provider (or treatment system) more effective than that of another?" The answer to this question is a first step toward identifying which clinical providers use the

"best" practices, i.e., practices that hold the potential for improving the quality of services and resulting client outcomes.

A review of the literature shows that clients' outcomes (and apparent provider performance) can be influenced by factors operating at multiple levels (Reichman, Rogers, Ameen & Harwood, 2000). These levels include:

- Client-level factors
- Treatment-level factors
- Structural characteristics/features of providers
- Features of the providers' operating/regulatory environments
- Community-level factors
- State-level factors.

There are several illustrative examples of comparative analyses of provider performance that statistically control for client-level factors (Harwood, Field, & Tunkelrott, 2001; Koenig, Fields, Dall, Ameen, & Harwood, 2000; Ameen, Dall, Harwood, & Doyon, 1999; Hendryx, Dyck, & Srebnik, 1999; and Swindle et al., 1995). Performance measurement analyses in the literature typically control for clients' socio-demographic characteristics, severity of presenting problems, and readiness for treatment. For example, Harwood et al. found that clients who are female, African-American, or who have children are significantly less likely to comply with a referral to treatment from detoxification. Koenig et al. found that clients for whom counseling for employment problems was important or very important had greater improvement in days of employment at follow-up than those for whom counseling was not important or only moderately important. Koenig et al. also found that clients with less severe medical problems had statistically less improvement in days of paid employment following treatment.

A number of studies have attempted to simultaneously control for client case mix and provider-level factors (Rogers, Ameen, & Harwood, 2001; Swindle et al., 1995). Some studies attempted to control for provider-level characteristics such as length of stay in treatment, treatment intensity, and other characteristics of the treatment experience. For example, both Harwood et al. (2001) and Rogers et al. found length of stay in treatment to be significantly and positively correlated with successful referral to treatment from detoxification or a lower level of care from residential treatment (respectively). Swindle et al. controlled for program characteristics including intended and actual length of stay, dropout rates, tolerance of problem behavior, and rates of discharge to aftercare. They found that treatment episodes with longer intended and actual lengths of stay, lower 7-day dropout rates, greater tolerance of problem

behavior, and greater proportions of patients receiving immediate post-discharge mental health treatment, among other factors, had lower 180-day readmission rates.

There also appears to be empirical and theoretical support for the need to account for (i.e., statistically control for) community and environmental factors given their potential impact on clients' post-treatment outcomes. Joe, Singh, Finklea, Hudiburg, & Sells (1977) used race-ethnicity variables as proxies for community structure and poverty when studying during-treatment outcomes for 13 methadone maintenance and outpatient drug-free treatment providers participating in the Drug Abuse Reporting Program (DARP). These community structural variables included the proportions of blacks, Puerto Ricans and Hispanics (non-Puerto Rican) in the community; unemployment; two variables measuring the percentage of black and Hispanic families living in poverty; and property crime rates. The outcomes of interest included opioid and non-opioid use, employment, productive activities (i.e., job, attending school, homemaking), and criminal activities. Results showed that the community-level characteristics were predictive of clients' outcomes. Specifically, low opioid use after treatment was most strongly associated with a small percentage of Hispanics in the community. Similarly, more favorable outcomes on productive activities were associated with a lower percentage of Puerto Ricans in the community. Criminal activities were associated with low unemployment indices and a low percentage of blacks in poverty in the community. Criminal activities also were associated with all of the community structural variables.

Another study conducted by Singh (1981) examined the effects of community-level factors on outcomes of DARP methadone maintenance programs in 27 large cities across the country. Five factors, including socio-demographic variables, social disorganization, economic conditions, crime/law enforcement, and quality of life were hypothesized to influence treatment outcomes. Clients were assigned scores for each of five outcome measures, and the individual client scores were aggregated for each treatment provider with rank order correlations used to examine associations between these scores and various measures of community-level factors. The results showed that large community population size, low median education, low percentage of managers, proprietors, and government officials in the labor force (an economic indicator), and high percentages of African-Americans, youth, female heads of household, and families in poverty were correlated with opiate and non-opiate use. Criminality outcomes were most strongly correlated with high property crime indices, while there were no strong correlations between community measures and alcohol consumption by methadone clients. Also, a low percentage of youth in a community was found to positively influence outcomes for productive activities.

A more recent study by Howard, LaVeist, & McCaughrin (1996) used data collected by the University of Michigan from 346 outpatient substance abuse treatment providers to examine completion of and dropout rates from treatment. Among the predictors was an indicator of the social environment based on the urbanicity and underclass status of a community. Organizations were classified as located in an urban or rural area and further classified as located in a poor or non-poor setting. The results indicated that organizations located in more urban settings, whether poor or non-poor, in general, had better treatment outcomes than organizations located in rural, non-poor settings. Additionally, these environmental factors were found to be better predictors of treatment success than was race. The authors concluded that research should place more emphasis on the socio-environmental influences on outcomes and less emphasis on client race.

Although the results are not yet available, two studies currently being conducted under the CSAT TOPPS II initiative are directly relevant to the issue of assessing the performance of different substance abuse treatment systems. One study is being conducted by a consortium of states including Oklahoma, Washington and Maryland using data on client outcomes from government archival sources. Analysts in these states identified standardized outcome measures and measures of client factors that influence those outcomes. Using post-treatment outcomes, including total wages, arrests, readmission to treatment and death, parallel (i.e., non-pooled) analyses were conducted by each state. These analyses will be used to assess the performance of each state treatment system with respect to the outcomes and contributing factors, as well as to make cross-state comparisons. The authors of this study explicitly recognize the complexities associated with attempts to make cross-state comparisons of performance given differences in macro-environmental influences such as arrest rates and employment opportunities. In addition, the researchers are addressing the significant challenges associated with differing definitions of both outcome and predictor variables across states (Leeper, Arria, & Krupski, 2000).

The second system-level performance measurement study, also being conducted under TOPPS II, will attempt cross-state/region analyses using pooled data from a consortium of states (Utah, New Hampshire, Rhode Island, and Arkansas). These states are adhering to a standardized study protocol and using a core set of instruments to collect client data at intake, during treatment and at follow-up. In addition, participating providers have completed a survey instrument that collects data on the structure, organization, staffing patterns and financing of the organization. The pooled analyses will attempt to identify relative *provider performance* before and after adjusting for differences in client case mix, organizational characteristics, and environmental/geographic factors (e.g., unemployment and crime rates, the differential availability of alcohol, treatment availability) (Utah Consortium, 1998).

These two studies illustrate the growing importance and need for assessments of system-level performance. They also illustrate the data requirements and challenges associated with the development of standardized performance indicators of treatment systems and case mix controls.

The analysis described in this report involves assessing and comparing the performance of six cities. It builds on previous work and complements the ongoing TOPPS II studies system-level analyses mentioned above.

#### **4. OVERVIEW OF THE ANALYSIS**

The overarching goals of this analysis are to:

- Identify and operationalize standardized performance measures that are appropriate for assessing the performance of treatment systems/geographic regions
- Conduct an illustrative performance measurement analysis that uses the public treatment system within specific geographic regions (in this case, cities) as the unit of analysis
- Assess the individual and relative performance of the treatment systems/cities using the identified performance measures
- Highlight the uses, benefits, constraints, and data requirements for system-level performance measurement monitoring and comparisons.

This analysis is based on data collected during the second phase of the CSAT Target Cities initiative. Sufficient standardized data and performance measures (constructed using client outcome measures) were available from six cities to render the data set amenable to performance measurement analyses.

The analytic approach involved first identifying which cities reported on the minimum set of key variables needed for the performance measurement analysis. Six were identified. Next, outcome domains were selected that are of importance to policymakers and that could be reasonably expected to change as a result of substance abuse treatment. The outcome domains selected were drug/alcohol use, employment, and criminal involvement. Within each outcome domain, performance measures were identified and/or created. It was important to ensure that the performance measures were comparably measured, were interpretable and had face validity. Where possible, multiple formulations of the performance indicators were identified and/or created. This was done in order to ensure that the results for any given performance measure were not a function of the manner in which the performance measure was constructed. This also

permitted an assessment of the degree to which there was consistency across the alternative formulations of the performance measures. A total of 11 formulations of the performance measures were constructed: eight in the domain of drug/alcohol use, two in the area of employment, and one in the criminal justice outcome domain. Details regarding the construction of the performance measures are contained in the Methods section of this report.

Using either Ordinary Least Squares or Logistic Regression, as appropriate, each performance measure was modeled using client-level data. Dummy variables/indicators were used for the cities to test for statistically significant differences. Client characteristics and measures of severity of presenting problems were controlled for as permitted by the data. The cities were compared using each of the selected performance measures. The results of these comparisons are summarized for each of the various types of treatment and types of clients.

## **5. ORGANIZATION OF THIS REPORT**

There are three additional chapters to this report. Chapter II describes the methods and analytic approach to conducting the analysis. It also includes a detailed description of the data source. Chapter III presents the findings from the cross-city comparisons of each of the 11 outcome measures by type of treatment and primary substance of use. Chapter IV summarizes the findings and constraints of this analysis and presents a discussion of the implications of this analysis for treatment providers, policymakers, and future research/evaluation using performance measurement.

## **II. METHODS**

## **II. METHODS**

In the following section, the data source and analytic approach used in this analysis are described. First, the nature and content of the data source – the multi-site database of the Target Cities initiative – are described. Next, the sample exclusion criteria and operationalizations of the variables used to specify the models are presented. Finally, descriptions of the analytic approach and statistical methods used to conduct the analysis are presented.

### **1. THE TARGET CITIES DATA SET**

This section of the report describes the source of the data and the composition of the database used in the analysis. The criteria for exclusion of records from the analysis database are also described.

#### **1.1 Data Source**

The data used in this analysis were obtained from the multi-site database of the Center for Substance Abuse Treatment's (CSAT) Target Cities Initiative. These data were compiled during the second phase of the multi-site initiative. Data collection was initiated after some sites had begun data collection efforts, thereby precluding the ability to modify the data collection instruments or design to collect completely standardized data across all of the participating cities. Consequently, the University of Akron, as the Coordinating Center for the effort, performed a crosswalk of the sites' data collection instruments and selected a core set of common variables for inclusion in the multi-site database. A multi-site evaluation planning team met on a regular basis to discuss issues surrounding the merits of, and barriers to including particular variables in the multi-site data set. The team agreed on a standardized set of concepts and accompanying variables. Given each site's familiarity with their instrumentation, each site made individual determinations regarding whether to submit individual variables based on the extent to which it represented the concept of interest. Thus, the final multi-site data set contained only data that were highly consistent with the concepts under consideration. Following receipt of the data, the University of Akron conducted internal validity and other quality checks on the data submitted. (See Stephens, Scott, & Muck, in press, for more detail.)

The primary data set used in this analysis contained demographic, substance use, and other severity status measures collected from clients at intake and follow-up for seven cities. This data set included elements of the Addiction Severity Index (ASI) along with additional elements of interest to local evaluators. This was the case except for one site that used a modified version of the Comprehensive Intake Assessment Instrument (CIAI). Another data set

included information on clients' treatment history, including treatment type. Neither the data set documentation and descriptions of the cities' study designs nor the published literature contained information regarding the criteria used for selecting providers within cities. One goal of the Target Cities initiative was to design and implement a central intake unit (CIU) in each city through which clients could access treatment. Clients who accessed treatment through a CIU were frequently provided with case management services, while clients in the comparison group accessed treatment directly from the provider and received standard care. The data were collected during somewhat different periods of time and for different durations across the cities and analysis groups. (Refer to Exhibit II-1 for a summary of the data collection periods by site.) The timing of the follow-up data collection also differed across the cities. Follow-up data were collected primarily at 6 and/or 12 months post-baseline, with one site also collecting data at 1 month post-baseline and another at 3 months post-baseline. (Refer to Exhibit II-2 for a summary of the follow-up data collection schedule.)

<b>EXHIBIT II-1 DATA COLLECTION PERIODS BY SITE</b>	
<b>City</b>	<b>Data Collection Period</b>
City 1	CIU group: August 1996-June 1998 Comparison: March 1995-December 1995; July 1996-July 1997
City 2	CIU group: n/a Comparison: n/a
City 3	CIU group: November 1996-February 1997 Comparison: November 1997-February 1998
City 4	CIU group: March 1997-July 1997 Comparison: November 1994-April 1995
City 5	CIU group: September 1996-November 1996; July 1997-September 1997 Comparison: September 1996-January 1997
City 6	CIU group: May 1995-February 1997 Comparison: March 1995-July 1996

<b>EXHIBIT II-2</b>				
<b>TIMING OF THE FOLLOW-UP INTERVIEWS</b>				
<b>City</b>	<b>Months After Baseline*</b>			
	<b>One</b>	<b>Three</b>	<b>Six</b>	<b>Twelve</b>
City 1			X	
City 2			X	X
City 3			X	
City 4			X	X
City 5		X	X	X
City 6	X			X

\* All baseline data were collected at client intake.

## 1.2 Analysis Database

The primary data set used for this analysis contained all intake and follow-up assessment data for each client. Each record represented one client. This analysis included both CIU and comparison group clients<sup>1</sup> because separate analyses of the Target Cities data showed little difference in outcomes between the CIU and comparison clients (Stephens et al., in press). The intake/follow-up assessment data set was merged with a data set containing treatment history data using a unique client identification number to match the records for each client. The treatment history database contained data for seven cities (derived from each city's local MIS system), although only six of the cities were included in both databases. The records without a corresponding match in the other database were deleted since the analysis required information from both data sets. In addition, records for clients with duplicate entries were deleted.<sup>2</sup> Accounting for these deletions, the merged data set contained 6,163 clients in six cities.

<sup>1</sup> One objective of the Target Cities initiative was to design and implement a central intake unit (CIU) in each city through which clients wishing to access publicly funded substance abuse treatment would be referred to care. Clients in the CIU group (referred to as the experimental group in Target Cities documentation) were frequently provided with case management services, while clients in the comparison group accessed treatment directly from the provider and received standard care.

<sup>2</sup> For clients with multiple records (as determined by duplicate unique client identification number): (1) if records were identical, one was deleted; (2) for records identical except for date of interview—the record with the later date was kept; (3) for almost identical records (identifying information matched and same fields were completed, though values differed slightly)—the record with the later interview date was kept; (4) for identical records included in both the comparison and CIU groups in the treatment history database—the record which corresponded to the single record in the intake/follow-up database was kept; (5) for records in both the control and CIU group data in both databases—the correct record could not be identified, and so these records were deleted; and (6) for duplicate records for which identifying information, such as date of birth, differed, two distinct records were assumed, and the unique client identification numbers were differentiated from each other.

The analysis data set contained information in eight areas: client demographics, substance use, crime and criminal justice system status, family/social environment, employment, education, physical health, and mental health, in addition to treatment history and referral data. Overall, while the Target Cities data set contained much of the information needed to assess and compare the performance across cities, it lacked some important information that would have been useful in further specifying the models used in the analysis. For example, the Target Cities data set did not contain any information that would permit determining clients' length of stay/duration of treatment. Additionally, the Target Cities data set lacked information on treatment intensity, the nature of services provided, and the number and qualifications of treatment staff. Finally, the Target Cities data did not permit identifying clients' movement across levels of care, information that is useful for addressing continuum of care issues. Lack of these data prevented analysis of whether and the extent to which these "treatment characteristics" were associated with outcomes. This information, while not critical to conducting comparisons of cities' performance, would be useful in explaining *why* particular clients being served by a given provider within a city had better/worse outcomes, or why clients in particular cities had better/worse outcomes. It does not, however, prevent determining *whether* outcomes were better/worse in various cities.

### 1.3 Exclusion Criteria

The merged data file contained data for 6,163 clients in six cities. Several exclusions were necessary prior to conducting the analyses. First, only clients who primarily used either alcohol, heroin, cocaine (including crack) or marijuana were included because sample sizes for clients with other primary substance use problems were too small to support a multivariate analysis. Thus, 679 records for clients with another primary substance of use were excluded. Second, given that the intent of this analysis was to examine post-treatment outcomes, 748 records were deleted for which neither a 6-month nor 12-month follow-up interview had been completed. Another 801 records were excluded because both days of substance use and days of substance use problems as measured at intake were either zero or missing. These clients were considered to be qualitatively different from the rest of the analysis population. An additional 285 records were excluded for clients for whom age could not be calculated or who were younger than 18 or 65 and older where there were apparent errors in data coding. Finally, eight additional records were excluded for cases in which type of care was inconsistent or in which the sample size for a type of treatment within a particular city was very small.<sup>3</sup>

---

<sup>3</sup> The following eight records were deleted: 3 clients in inpatient, therapeutic community care; 1 client in "other" modality; 1 client in intensive outpatient methadone treatment; 1 client in City 4 in short-term residential treatment; and 2 clients in methadone treatment in City 5.

The final analysis database used for analysis contained 3,642 clients across six cities. These clients received treatment in one of five types of treatment: non-intensive outpatient, intensive outpatient, methadone treatment, short-term residential, or long-term residential. All clients included identified either alcohol, heroin, cocaine (including crack) or marijuana as their primary substance of use.

Across the six cities, 368 and 497 clients received non-intensive outpatient and intensive outpatient treatment, respectively. An additional 310 clients received methadone treatment in three of the cities (the other three sites did not include methadone). Four of the six cities had a total of 394 clients in short-term residential care, and 364 clients across five cities received long-term residential treatment. Another 575 clients did not receive/show up for treatment. (Note that the type of treatment data were missing for many clients.) The largest group of clients, 1,419 across the six cities, identified cocaine (including crack) as their primary substance of use. Heroin was the substance of choice for 1,099 clients, while alcohol was the primary substance of choice for 924 clients. A small number of clients (190) reported marijuana as their primary substance of use.

## **2. ANALYTIC METHODS/APPROACH**

This section describes the approach to comparing the performance of the treatment systems. It also presents the parameters surrounding the specifications of the analytic model.

### **2.1 Analytic Approach**

The approach to comparing the performance of treatment systems involved first assessing whether the cities had the minimum comparable data available to be included in the analysis. Next, it was necessary to create analytically meaningful classifications of clients in order to make comparisons across cities. The data permitted categorizing clients by type of treatment and by substance use type. Clients were categorized by type of care in order to facilitate making appropriate comparisons across cities. Level of care was one particularly meaningful distinction because it “standardizes” for treatment intensity. For analytic purposes, outcomes were examined by city for each of the following types of treatment:

- Non-intensive outpatient (ambulatory treatment services including individual, family, and/or group services)
- Intensive outpatient (clients receive, at a minimum, two or more hours of treatment per day for three or more days per week)
- Methadone (pharmacological therapy for clients with heroin addiction)

- Short-term residential (treatment duration of 30 days or less)
- Long-term residential (treatment duration of more than 30 days).

Outcomes were also examined by categorizing clients based on their primary presenting problem/substance type (regardless of type of treatment). This categorization was also deemed necessary since it permits cross-city comparisons of clients with similar presenting problems. Outcomes were created and examined for clients whose primary substance of choice was:

- Alcohol
- Heroin
- Cocaine/crack
- Marijuana.

Next, it was necessary to select criteria (performance measures) against which the cities could be reasonably held accountable for effecting change. It was determined that funders, policymakers, the public and clients could reasonably expect participation in substance use treatment to improve clients' functioning in the domains of drug/alcohol use, employment, and criminal justice. Consequently, standardized performance measures were derived from these three outcome domains. Many of the measures included are elements of the Addiction Severity Index (ASI) that were asked using uniformly-worded questions by all but one of the cities.

Methodologically, it was important to minimize the possibility that the findings from the analysis were a function of variable construction and instead attributable to participation in substance use treatment. Consequently, where possible, several different formulations of the performance measures in each outcome domain were constructed. In the drug/alcohol outcome domain, a total of eight alternative formulations of performance measures were created. Collectively, these performance measures were intended to measure reduction in associated drug/alcohol-related problems and use. For the employment outcome domain, two performance measures were created that were intended to assess the extent to which clients demonstrated improvements following treatment with respect to their employment status. For the criminal justice outcome domain, an area of considerable concern to policymakers and the public, one performance measure was created that was designed to measure criminal activity at follow-up. (Additional detail on these performance measures and their construction is found in Section 2.2.)

The analytic challenge associated with any comparative performance measurement analysis is to ensure that the "units of analysis" (in this case, the treatment systems within cities) are compared in terms of changes in outcomes that are attributable to various aspects of the treatment process rather than the mix of client characteristics. First, a single performance

measure was constructed for each of the outcome domains (substance use, employment and criminal activity) that were standardized for severity at intake in that particular measure. The treatment systems were compared on these three preliminary measures. The three measures were selected because each captured change in clients' functioning from intake to follow-up in each of the three outcome domains of interest, changes that are expected to be affected by participation in substance abuse treatment. Additional detail on their construction is contained in Chapter III, Section 1.6. While adjusting for initial level of severity at intake, these performance measures could not be standardized for client characteristics and severity in other areas of functioning. Therefore, multivariate analyses were conducted in order to statistically control for as many client characteristics and severity-related influences on clients' outcomes as were permitted by the data.

Depending on the distribution of the outcome measure under consideration, either Ordinary Least Squares (for performance measures that were continuous) or Logistic Regression (for performance measures that were dichotomous) was used to partial out the effect of client characteristics and measures of severity of presenting symptoms at intake on the performance measures. The selection of client characteristics and severity measures (i.e., case mix controls) used in the analysis was guided by the review of the literature (Ameen et al., 1999; Rogers et al., 2001; Harwood et al., 2001; and Swindle et al., 1995). Each performance measure was modeled as a function of client demographics (age, gender, race/ethnicity and marital status), client severity at intake in the domains of drug and alcohol use and prior psychiatric treatment episodes, primary substance of use, various measures of criminal justice history/status, employment history, and level of care measures. Refer to Exhibit III-1 in Chapter III for a complete list of independent variables. For each performance measure, the model specifications were customized to include the most theoretically relevant variables. Alternative statistical techniques to a fixed effects model (e.g., random effects models or multi-level models) might also be employed; however, they typically require larger sample sizes, have more stringent data requirements, and are more difficult to interpret. Given that this analysis was intended to be illustrative, the more straightforward fixed effects approach was used. Refer to Appendix A for a list of the model specifications for each of the performance measures.

Results from the multivariate analysis provide insights into how the respective cities performed with respect to improving clients' functioning in the areas (outcome domains) of drug/alcohol use, employment, and criminality. Specifically, the use of multivariate analysis to adjust for case mix (i.e., differing demographic profiles and the degree of impairment at intake) provides insights into systems' performance after "accounting" for differences that are beyond the influence of the providers and clinicians providing treatment. The results are presented in the Findings section of this report.

---

## 2.2 Model Specification

### Performance Measures

This analysis measured clients' outcomes in three domains: drug/alcohol use, employment, and crime. Multiple measures of clients' outcomes were used for each domain. For drug/alcohol use, the selected performance measures were based on drug/alcohol problem days in the past 30 days and drug/alcohol use days in the past 30 days measured at follow-up. There are several alternative formulations of these outcome measures including change in problems/use from intake to follow-up, proportional change in problems/use, and a dichotomous measure (absence of problems/use at follow-up). For employment, the performance measures were days with paid employment in the past 30 days and change in days of paid employment from intake to follow-up. For crime, the performance measure was whether clients were arrested and charged with at least one of ten crimes since the last assessment interview. As indicated in Appendix A, the models for each outcome domain were similarly specified with problem severity at intake being included as a covariate in models for both change scores as well as level of functioning measured at follow-up. Refer to Exhibit III-2 for descriptive statistics for these measures.

There are eight drug/alcohol performance measures. The first is a measure of the number of days with drug/alcohol *problems* in the past 30, as measured at follow-up. The second is a measure of the number of days of drug/alcohol *use* in the past 30, as measured at follow-up. Both of these measures are integer valued variables with values ranging from 0-30 and are based on clients' identified primary substance of use. Clients with no problem days or no use days at intake are excluded from the respective measures. For example, for clients with a primary alcohol problem, days of alcohol use and days with alcohol problems were used to construct the drug/alcohol use and drug/alcohol problem measures, respectively. For the cases in which days of substance use at follow-up were missing, two dichotomous variables (use in past 30 days and use since last assessment interview) were examined and days of substance use were set equal to zero if either of those variables was "no."

The third performance measure is the change in number of days with drug/alcohol problems in the past 30 from intake to follow-up. This is calculated by subtracting the number of days with drug/alcohol problems in the past 30 at intake from the number of days with problems at follow-up. This measure is an integer valued variable ranging from -30 to 29. A negative change score indicates a more favorable outcome, while a positive change score indicates a less favorable outcome. A fourth measure, the change in the number of days of drug/alcohol use in the past 30 from intake to follow-up is calculated in the same fashion. The number of days of

substance use in the past 30 at intake is subtracted from the number of days of substance use at follow-up. This is also an integer valued variable ranging from -30 to 29.

Another performance measure is the proportional change in the number of days with drug/alcohol problems in the past 30 from intake to follow-up. This is calculated by dividing the change in days with drug/alcohol problems from intake to follow-up (calculation described above) by the number of days with drug/alcohol problems in the past 30 at intake. This measure is a continuous variable ranging from -1 to 29. A measure of the proportional change in the number of days of drug/alcohol use in the past 30 from intake to follow-up is constructed in the same manner. This is also a continuous variable ranging from -1 to 29.

The final drug/alcohol measures were dichotomous variables. One indicates whether a client has an absence of drug/alcohol problem days in the past 30 at follow-up (value=1) or any days of problems in the past 30 at follow-up (value=0). The other indicates whether a client has an absence of days of drug/alcohol use in the past 30 at follow-up (value=1) or any days of drug/alcohol use at follow-up (value=0).

There are two employment performance measures. The first is the number of days of paid employment in the past 30 days at follow-up. This is an integer valued variable ranging from 0 to 30. The second measure is the change in the number of days with paid employment in the past 30 from intake to follow-up. This measure is calculated by subtracting the number of days of paid employment in the past 30 at intake from the number of days of paid employment in the past 30 at follow-up. This is an integer valued variable ranging from -30 to 30. A positive change score indicates a more favorable employment outcome than does a negative one. City 2 did not collect data on these variables, and thus is excluded from the calculation and analysis of both employment measures.

The criminal justice performance measure is a dichotomous variable measuring whether the client had been arrested and charged with one of ten crimes since the initial assessment interview (value=1) or if he/she had not been arrested (value=0). The ten crimes included in this measure were shoplifting/vandalism, drug charges, assault, arson, burglary, forgery, robbery, prostitution, rape, and homicide/manslaughter. The duration of time since the last assessment interview was most usually 6 months, except in the cases of City 6 and select records for which 6-month data were not collected. In these cases, the duration of time since the last assessment interview is 12 months.

## **Independent Variables**

The independent variables used in this analysis represent client-level demographics, client substance use severity, criminal justice history, employment history, and type of treatment. The independent variables were selected to represent variables identified in the literature that influence clients' post-treatment outcomes. The case mix controls used in this analysis were selected primarily at the client level. While the literature suggests that it is also important to adjust for provider and treatment-level characteristics as a second step of the analysis to explain the observed differences in rankings across cities, the data set did not support constructing such measures. Refer to Exhibit III-1 for descriptive statistics for the variables used in this analysis. The independent variables used in the analysis were defined as follows.

## **City and Type of Treatment**

This set of variables allows for comparisons across the cities and types of treatment. The city in which a client was enrolled is indicated by a set of five dichotomous variables: (a) City 2=1 and others=0; (b) City 3=1 and others=0; (c) City 4=1 and others=0; (d) City 5=1 and others=0; and (e) City 6=1 and others=0. City 1 was used as the comparison city because it was generally close to the mean value for each outcome measure, and it had sufficient sample sizes for each client-level grouping. Type of treatment was included as a statistical control in those models that sub-set the data by primary substance of use. For those models, a set of five dichotomous variables was used to indicate type of treatment or no treatment received: (a) non-intensive outpatient care=1 and other=0; (b) intensive outpatient care=1 and other=0; (c) methadone treatment=1 and other=0; (d) short-term residential=1 and other=0; and (e) long-term residential=1 and other=0. The comparison condition was "received no treatment." The no treatment group is comprised of both control and CIU group clients who failed to continue to access treatment. These clients typically dropped out of treatment within 14 days of the intake assessment. Since previous analyses of the Target Cities multi-site data found no differences between the CIU and comparison groups in terms of outcomes, data for clients in these groups were pooled and analyzed together. It is also important to note that not all cities had clients enrolled in every type of treatment.

The models also included a dichotomous variable that indicated the timing of the follow-up interview from which the performance measure was constructed. For the large majority of the clients, 6-month post-baseline follow-up data were used. However, as mentioned earlier, City 6 did not collect 6-month follow-up data. Therefore, for this city and for any clients in Cities 2, 4 and 5 missing 6-month data but having 12-month data, 12-month post-baseline follow-up data

were used. This variable indicates where 6-month data (value=0) versus 12 month data (value=1) were used.

### **Client-level Demographics**

The regression models contain the standard client-level demographic variables routinely included in treatment outcomes and performance measurement analyses, particularly those that employ case mix adjustment methods. Variables included in this analysis include age, gender, race/ethnicity and marital status. Client age has been found to affect treatment outcomes (Ornstein & Cherepon, 1985; Leigh, Ogborne & Cleland, 1984). Age at intake is included in the models as both a continuous linear variable and as a quadratic term. The quadratic is included to capture any non-linear relationship between age and substance use. As mentioned previously, clients under age 18 or 65 years and older were not included in this analysis. Gender has also been found to be associated with client outcomes (Agosti, Nunes, Stewart, & Quitkin, 1991) and was constructed as a dichotomous variable where 1=male and 0=female. Race/ethnicity, also found to be associated with clients' outcomes (Agosti et al.), is represented by three dichotomous variables: (a) 1=black and 0=all others; (b) 1=Hispanic and 0=all others; and (c) 1=other race/ethnicity (i.e., non-white, black or Hispanic) and 0=white, black or Hispanic, with whites being the reference category. Marital status is represented by two dichotomous variables: (a) 1=never married and 0=other; and (b) 1=divorced, separated or widowed and 0=other, with married being the omitted category. Level of educational attainment is excluded from the models because City 5 did not collect any data on the item. The sensitivity of the models was tested to determine whether the inclusion/exclusion of level of educational attainment would affect the results. Analyses with a sample of clients in the cities with these data indicated that excluding educational attainment did not affect the results.

### **Client Severity**

In order to adjust for differences in client composition (i.e., case mix) across cities, measures of client severity at intake were included. The client severity measures included in the analysis were comprised of four variables measuring the extent of recent alcohol, heroin, cocaine (including crack), and marijuana use at intake, respectively; two variables measuring days with alcohol/drug problems at intake; and two variables measuring prior psychiatric treatment (inpatient and outpatient). With the exception of the two psychiatric measures, the data did not permit any inclusion of measures of drug/alcohol severity for the clients' lifetime (e.g., age of first regular use or dependence).

Four integer-valued variables (value 0-30) measure days of alcohol use in the past 30 days at intake, days of heroin use in the past 30 days at intake, days of cocaine (including crack) use in the past 30 days at intake, and days of marijuana use in the past 30 days at intake. These variables are included for all clients, regardless of primary substance of use. If clients were missing data for a measure other than the one corresponding to their primary substance, the value was assumed to be zero. Two continuous variables (value 0-30) measure days with alcohol problems in the past 30 days at intake and days with drug (any type excluding alcohol) problems in the past 30 days at intake. These variables are also included for all clients.

Prior receipt of psychiatric treatment in a lifetime is also used as a measure of client severity. These variables are included because research shows that clients with psychiatric comorbidities present additional challenges to substance abuse treatment. The models include two dichotomous variables. One measures whether the client has ever (in his/her lifetime) been hospitalized for psychiatric treatment (1=yes; 0=no). The other measures whether the client ever received (in his/her lifetime) outpatient treatment for psychiatric problems (1=yes; 0=no).

The models also include a set of three dichotomous variables indicating clients' primary substance of use. These variables are included in the analyses conducted by type of treatment. Primary substance of use is represented by: (a) 1=heroin and 0=other substance; (b) 1=cocaine/crack and 0=other substance; and (c) 1=marijuana and 0=other substance, with alcohol as the omitted category. As detailed previously, clients indicating a primary substance of use other than alcohol, heroin, cocaine/crack or marijuana were excluded from the data set used for analysis.

### **Criminal Justice History**

Prior involvement or status with the criminal justice system is measured by probation/parole status at intake, whether admission to treatment was prompted by the legal system, and history of arrests and charges for a set of six crimes. A dichotomous variable measures whether a client is on probation or parole (1=yes; 0=no) at the time of intake. Another dichotomous variable indicates whether a client's admission to treatment was prompted by the legal system (e.g., judge or probation/parole officer) (1=yes; 0=no). A set of six dichotomous variables measure whether a client was arrested and charged at least once with each of six different crimes during his or her lifetime. Those six variables, collected at intake, measure whether arrested and charged with shoplifting/vandalism (1=yes; 0=no), drug charges (1=yes; 0=no), assault (1=yes; 0=no), robbery (1=yes; 0=no), burglary (1=yes; 0=no), and homicide/manslaughter (1=yes; 0=no). Four crime measures collected at intake (arson, forgery, prostitution, and rape), though included in the outcome measure, are excluded from the model

because very few clients had ever been arrested and charged with those offenses. Analyses revealed that excluding these variables make no difference to the results.

### **Employment History**

Two continuous variables measure employment history and stability at intake. The first measures the number of days with paid employment during the past 30 days at intake (value 0-30). The other measures the number of months of employment for the client's longest full-time job, ranging from 0-420. It should be noted that this variable was operationalized somewhat differently for City 6. This city actually collected the number of months of employment for the client's last full-time job. These employment history variables were not collected by City 2 and are therefore only included in the models that predict employment outcomes at follow-up in order to retain City 2 in the remainder of the analyses.

### **III. FINDINGS**

### III. FINDINGS

The following section contains a description of the characteristics of the sample including client demographics, severity measures, and type of treatment. This reveals that there are quite distinct differences across the six cities in terms of the socio-demographic characteristics, the severity profile of clients initiating treatment, and the types and mix of treatment to which they were admitted. The section also contains descriptive statistics on each performance measure and its components. There are also distinct differences across the six cities in terms of the average client outcomes on various measures—cities are apparently achieving a range of respective outcomes, suggestive of the potential value of attempting to identify which ones are better/weaker and why. Of course, the clear differences across cities uncovered in the descriptive data speak directly to the necessity of performing case mix adjustment before attempting to reach conclusions about which cities have better/worse performance. The findings are summarized with respect to each city’s performance on the outcome measures by type of treatment and primary substance of use.

#### 1. SAMPLE DESCRIPTION

A total of 3,642 clients received an intake assessment in six Target City treatment systems. In general, the cities appeared to have served clients with differing socio-demographic and severity compositions, suggesting the need for multivariate analyses when making performance measurement comparisons. Furthermore, the cities used various types of treatment in different proportions, or at least the samples compiled for this data set had quite different compositions. The descriptive statistics for the independent variables are shown in Exhibit III-1.

<b>EXHIBIT III-1</b>							
<b>DESCRIPTIVE STATISTICS FOR INDEPENDENT VARIABLES</b>							
	<b>City 1</b>	<b>City 2</b>	<b>City 3</b>	<b>City 4</b>	<b>City 5</b>	<b>City 6</b>	<b>Total</b>
<b>Client-level Demographics</b>							
Age (mean)	35.3	36.6	33.5	35.5	34.2	37.5	35.5
Male (%)	44.2	55.1	52.6	61.5	74.3	67.8	50.7
<b>Race/Ethnicity</b>							
White (%)	10.8	22.6	52.1	62.8	28.8	28.3	21.6
Black (%)	74.9	73.9	37.1	26.5	69.9	54.9	66.5
Hispanic (%)	12.9	1.4	7.2	2.3	0.7	5.3	9.0
<b>Marital Status</b>							
Married (%)	11.7	15.4	18.6	25.1	16.9	4.9	13.6
Never married (%)	60.0	50.0	41.2	43.8	54.7	0.0	51.8
Divorced/separated/widowed (%)	28.3	34.7	40.2	31.1	28.4	95.1	34.6

<b>EXHIBIT III-1 (CONT.) DESCRIPTIVE STATISTICS FOR INDEPENDENT VARIABLES</b>							
	<b>City 1</b>	<b>City 2</b>	<b>City 3</b>	<b>City 4</b>	<b>City 5</b>	<b>City 6</b>	<b>Total</b>
<b>Client Severity</b>							
Days of alcohol use (mean)*	6.7	12.1	10.0	5.3	7.9	5.1	7.4
Days of heroin use (mean)*	8.8	5.8	4.8	3.7	2.7	1.7	6.9
Days of cocaine use (mean)*	6.1	9.9	10.1	3.2	5.3	4.2	6.4
Days of marijuana use (mean)*	2.0	3.7	4.5	2.2	2.7	1.7	2.4
Days with alcohol problems (mean)*	6.4	10.4	10.4	4.7	7.1	8.9	7.2
Days with drug problems (mean)*	16.3	17.2	15.8	10.1	9.7	12.5	15.3
Prior inpatient psychiatric treatment (%)	17.5	17.3	13.9	19.1	14.2	17.8	17.3
Prior outpatient psychiatric treatment (%)	14.4	17.4	38.1	30.3	17.6	30.7	18.9
<b>Primary Substance of Use</b>							
Alcohol (%)	22.8	23.7	28.9	37.4	33.1	31.0	25.7
Heroin (%)	37.7	18.8	16.5	26.1	10.8	14.7	30.2
Cocaine (%)	36.5	52.8	48.0	22.8	39.2	49.0	39.0
Marijuana (%)	3.0	4.7	6.7	13.8	16.9	5.3	5.2
<b>Criminal Justice History</b>							
On probation/parole at intake (%)	17.8	22.5	19.1	47.9	43.2	23.7	22.9
Admission prompted by legal system (%)	9.9	16.2	42.8	53.2	32.4	20.8	18.5
Arrested & charged – shoplifting/vandalism (%)	22.9	13.4	24.5	35.3	33.1	22.5	23.2
Arrested & charged – drug charges (%)	35.4	30.2	33.7	42.0	32.4	44.1	35.7
Arrested & charged – assault (%)	14.9	8.7	16.7	19.8	14.2	22.9	15.1
Arrested & charged – robbery (%)	8.7	4.6	2.0	11.8	14.9	11.4	8.6
Arrested & charged – burglary (%)	9.3	10.3	7.8	19.6	12.8	13.1	10.8
Arrested & charged – homicide/manslaughter (%)	2.7	0.8	1.3	1.4	2.7	2.5	2.2
<b>Employment History</b>							
Days with paid employment in past 30 (mean)	3.2	n/a	7.5	5.9	8.2	1.3	3.8
Months of longest fulltime job (mean)	45.1	n/a	53.5	52.0	54.0	48.3	47.1
<b>Treatment Characteristics/Type of Treatment</b>							
Non-intensive outpatient (%)	9.2	9.6	14.1	43.9	26.4	10.4	14.7
Intensive outpatient (%)	12.1	41.1	12.9	7.7	11.5	38.2	19.8
Methadone (%)	19.2	8.8	n/a	14.7	n/a	n/a	12.4
Short-term residential (%)	26.9	7.5	n/a	n/a	28.4	n/a	15.7
Long-term residential (%)	8.7	6.7	22.4	24.2	n/a	50.6	14.5
No treatment received (%)	24.0	26.3	50.6	9.5	33.8	0.8	22.9

\* Number of days in the past 30 days with substance use/problems. Presented for all clients in the analysis database though clients with no use or problem days for their primary substance of use are excluded from the models of the corresponding outcome measures.

## **1.1 Client-level Demographics**

The mean age of clients included in the sample was 35.5 years, with little variability across cities. Approximately 80 percent of the clients were between the ages of 25 and 45. Slightly more than half (50.7%) of these clients were male, but this distribution was not consistent across cities. This variability is likely a function of the oversampling of female clients by some cities. In terms of race/ethnicity, clients were predominantly black (66.5%), with the exception of Cities 3 and 4 which had relatively low proportions of blacks in their samples (37% and 27%, respectively). On average across all cities, 21 percent of clients were white and 9 percent were of Hispanic origin. About one half of clients were never married (51.8%); 35 percent were divorced, separated, or widowed; and 14 percent of all clients were married. The average number of years of education completed was 11.5.

## **1.2 Criminal Justice History**

On average, 23 percent of all clients were on probation/parole at intake. There was considerable variability across cities, with Cities 4 and 5 having relatively higher proportions of clients on probation/parole (48% and 43%, respectively). In contrast, between 17 and 24 percent of the clients in the remaining cities were on probation/parole at intake. For the entire sample, while an average of 19 percent of clients' treatment was prompted by the criminal justice system, these proportions were significantly higher for Cities 3, 4, and 5. These cities had 43, 53, and 32 percent, respectively, of their clients whose admission was prompted by the legal system. Approximately 36 percent of clients had been arrested and charged with a drug-related offense during their lifetime.

## **1.3 Primary Substance of Use**

The cities also appeared to be serving clients with varying degrees of problem severity and differing types of presenting problems. As shown in Exhibit III-1, 39 percent of all clients in the sample reported cocaine/crack as their primary substance of use, although this proportion ranged from 23 percent of clients in City 4 to 53 percent of clients in City 2. Heroin was the substance of choice for 30 percent of the sample, with the proportion of users ranging from 11 percent in City 5 to 38 percent in City 1. Approximately 26 percent of all clients reported alcohol as their primary substance of use. The proportion of clients using alcohol was fairly consistent across cities, ranging from 23 percent in City 1 to 37 percent in City 4. On average, only five percent of clients across all cities reported primary use of marijuana, although in City 5, the city with the highest proportion of marijuana-using clients, 17 percent of clients used marijuana. City 1 had the smallest proportion of clients using marijuana (3%).

Seventeen percent of all clients reported ever receiving inpatient psychiatric treatment, while 19 percent had ever received outpatient treatment for psychiatric problems. These significant differences across treatment populations in the nature and severity of their presenting symptoms strongly suggest that multivariate analyses are needed to adjust for differences in the treatment population case mix when comparing the performance of the cities.

#### **1.4 Type of Treatment**

There were differences in the proportions of clients served by type of treatment across the cities. In the case of the Target Cities data this was planned to occur, because each of the cities had somewhat different evaluation objectives in mind when their sampling strategies were developed. However, such differences are a direct function of the variability of the demographic characteristics and nature and severity of the drug-related problems of the treatment populations. Differences also reflect how the respective system managers choose to design their treatment capacity in order to meet the perceived needs. Performance measurement analyses are needed to account for such fundamental differences in treatment system configuration. Multivariate analyses are therefore used to separate out the impact of treatment system configuration and client case mix since these factors mediate overall treatment system performance.

The majority of clients (46.9%) across all cities received substance abuse treatment services in outpatient settings (i.e., intensive outpatient, non-intensive outpatient or methadone treatment). Approximately 20 percent of clients across all cities received intensive outpatient treatment, while another 15 percent received non-intensive outpatient treatment. The proportion of clients who received intensive outpatient treatment ranged from only 8 percent in City 4 to 41 percent in City 2. The proportion of clients who received non-intensive outpatient treatment also varied across cities, from 9 percent of clients in City 1 to 44 percent of clients in City 4. While only three cities had clients who received methadone treatment, those cities had, on average, 12 percent of clients in methadone treatment, ranging from 9 percent in City 2 to 19 percent in City 1.

Smaller proportions of clients (30.2%) received treatment in residential treatment settings (short- or long-term residential treatment). Short-term residential and long-term residential treatment accounted for 16 and 15 percent of clients in the sample, respectively. Though only three cities had clients in short-term residential treatment, the proportion of clients in this type of treatment ranged from 8 percent in City 2 to 28 percent in City 5. The proportion of clients receiving long-term residential care also varied widely across the five cities. Seven percent of clients in City 2 received long-term residential treatment as compared to 51 percent of clients in City 6. Approximately 23 percent of clients in the sample either received an intake assessment and failed to show up for treatment or dropped out shortly after the intake assessment.

## 1.5 Substance Use Days/Substance Use Problem Days

Overall, the data also show that the CSAT Target Cities initiative served clients with significant drug/alcohol and associated social problems. Data presented in Exhibit III-2 show that clients used drugs/alcohol for an average of 16 days out of the past 30 when they presented for treatment. It is notable that the clients in Cities 1, 2, and 3 have an average of almost 18 days of substance use out of the past 30, while clients in Cities 4, 5, and 6 reported an average of 12 days of substance use in the last 30. Similarly, clients who presented for treatment reported experiencing an average of 19 days of drug/alcohol *problems* in the 30 days prior to intake. As with substance use in the past 30 days, Cities 1, 2, and 3 appear to have clients with more severe presenting problems than Cities 4, 5, and 6. Clients in Cities 1, 2, and 3 reported having more drug problem days (20.6) than those in Cities 4, 5, and 6 who reported an average of 16.4 problem days in the past 30 at intake. Substance abuse appears to be related to or to contribute to employment difficulties. The data show a significant negative correlation between both substance use days and substance use problem days at intake and the number of days of paid employment in the past 30 (data not shown). Descriptive results show that clients presenting for treatment reported an average of only 3.8 days of paid employment in the past 30 days at intake. Clients in Cities 3, 4, and 5 reported somewhat higher numbers of days of employment (7.5, 5.9, and 8.2, respectively) than did clients in Cities 1 and 6 (3.2 and 1.3, respectively).

The descriptive data also show that participation in treatment yielded strong benefits in terms of both reducing clients' drug/alcohol use and problem days and increasing their labor force participation between the pre- and post-treatment periods. Specifically, there was an average of a 12-day reduction in the number of substance use days in the past 30 from intake to follow-up. Clients in Cities 1, 2, and 3 exhibited the largest reductions (almost 13 days), as expected, given that those clients had the highest use at intake. Similarly, there was an average reduction of almost 13 days in the number of drug problem days from intake to follow-up. Clients in Cities 1, 2, and 3 reported the largest reductions. It is notable that City 5 had reductions that were similar to those of cities with more severely impaired clients, although its clients did not report being as severely impaired as those in Cities 1, 2 and 3. The data did not permit constructing a change score for arrests/charges. An average of 11.2 percent of clients were arrested and charged in the 6 months prior to the follow-up, with a range of 4.9 percent for City 2 to 21.2 percent for City 3.

<b>EXHIBIT III-2</b>							
<b>DESCRIPTIVE STATISTICS FOR COMPONENTS OF DEPENDENT VARIABLES</b>							
	<b>City 1</b>	<b>City 2</b>	<b>City 3</b>	<b>City 4</b>	<b>City 5</b>	<b>City 6</b>	<b>Total</b>
Substance use problem days at intake*	19.4	21.1	21.5	15.8	16.2	17.1	19.1
Substance use problem days at follow-up	6.6	4.7	4.9	8.5	2.0	8.5	6.3
Change in substance use problem days**	-12.8	-16.4	-16.6	-7.3	-14.2	-8.7	-12.8
No problem days at follow-up (%)	51.4	69.4	73.3	51.3	63.2	39.2	54.6
Days of substance use at intake*	17.2	18.9	17.2	11.4	14.3	10.3	16.4
Days of substance use at follow-up	4.8	6.2	3.6	5.8	3.4	4.3	4.9
Change in days of substance use**	-12.4	-12.8	-13.7	-5.5	-10.8	-6.0	-11.4
No substance use days at follow-up (%)	65.3	58.2	69.9	55.0	68.6	63.9	63.6
Days of paid employment at intake*	3.2	n/a	7.5	5.9	8.2	1.3	3.8
Days of paid employment at follow-up	6.8	n/a	11.1	9.1	10.6	7.1	7.5
Change in days of paid employment**	3.6	n/a	3.5	3.2	2.5	5.8	3.7
Arrested and charged with crime in lifetime (%) ***	58.9	46.3	61.7	69.6	62.8	67.8	59.1
Arrested and charged with crime in 6 months prior to follow-up (%) ***	11.1	4.9	21.2	14.6	9.5	12.7	11.2

\* Days at intake are presented only for those records included in outcome measures at follow-up.

\*\* Change in days is calculated by days at follow-up-days at intake.

\*\*\* Arrested and charged with a crime includes any one of the following: shoplifting/vandalism, drug charges, assault, arson, burglary, forgery, robbery, prostitution, rape and homicide/manslaughter. For all clients in City 6 and select clients in Cities 2, 4, and 5, measure is of arrests and charges in 12 months prior to follow-up.

## 1.6 Standardized Performance Measurements

While the descriptive tabulations are useful, they do not adjust for differences in client characteristics and severity at intake. Unadjusted comparisons of the cities' apparent performance are nevertheless possible using the descriptive data. Comparisons are possible that standardize for initial level of problem severity. They will not, however, account for differences in type of treatment, drug type, client characteristics, or severity measures. In the drug/alcohol outcome domain, cities are compared using the standardized change in drug/alcohol problem days. This measure is constructed as the average change in drug problem days from intake to follow-up divided by the average number of drug problem days at intake. A measure in the employment outcome domain is similarly constructed. The average change in the number of

paid employment days from intake to follow-up is divided by the average number of paid days at intake. In the criminal justice outcome domain, the average number of arrests in the 6 months prior to follow-up is standardized by dividing by the average number of lifetime arrests. These measures are presented in Exhibit III-3 below, along with the rankings of the cities in parentheses.

<b>EXHIBIT III-3</b>						
<b>STANDARDIZED MEASURES AND RANKINGS FOR SELECTED</b>						
<b>PERFORMANCE MEASURES</b>						
	<b>City 1</b>	<b>City 2</b>	<b>City 3</b>	<b>City 4</b>	<b>City 5</b>	<b>City 6</b>
Standardized change in substance problem days	-.66 (4)	-.78 (2)	-.77 (3)	-.46 (6)	-.88 (1)	-.51 (5)
Standardized change in days of paid employment	1.13 (2)	n/a	.47 (4)	.54 (3)	.34 (5)	4.40 (1)
Percentage change in criminal activity	.81 (3)	.89 (1)	.66 (6)	.79 (5)	.85 (2)	.81 (3)

Results from these standardized performance measurements showed no consistent patterns across outcome domains. For example, while City 3 appears to be an average performer on the drug and employment outcome domains, it ranks lowest on the criminal justice outcome measure. In contrast, City 5 appears to have the best performance in the drug/alcohol outcome domain and above average performance on the criminal justice measure but is below average on employment. City 1 appears to be an average (or at least “median”) performer on all three outcome domains. This city is used as the point of comparison in the multivariate analyses for exactly this reason. While these standardized comparisons yield some insights into relative city performance, the differences in cities identified earlier (i.e., in types of treatment, client demographics and severity measures) indicate that a multivariate analysis is warranted. It should also be noted that results based on single indicators are less reliable than those based on multiple indicators of the same outcome domain.

## **2. REGRESSION RESULTS**

The descriptive results showed that cities do differ with respect to both the characteristics and severity of clients (Exhibit III-1) and the outcomes of participation in substance abuse treatment (Exhibit III-2). Multivariate analyses were conducted in order to determine whether these observed differences in clients’ outcomes across cities were meaningful after controlling for client differences in characteristics and severity. In other words, the multivariate analyses permit a more rigorous assessment of whether the differences across cities are a function of the therapeutic services delivered to clients after adjusting for differences in difficulty of treating the different populations of clients served.

The cities were compared based on multiple performance measures (see Appendix A) for clients served in five types of treatment—non-intensive outpatient, intensive outpatient, methadone treatment, short-term residential, and long-term residential—as well as for clients with four different primary substances of use, including alcohol, heroin, cocaine/crack, and marijuana. Results from each regression model are presented in Appendix B.

Cross-city analyses are performed by designating a “base” city against which the performance of all others is compared. The base city was generally close to the mean value on outcome measures across the six cities. The tables in Appendix B show the cities’ performance on each of the 11 indicators, presented by type of treatment and substance type. In the comparisons against the base city, positive values indicate performance that was better than the mean/comparison city, while negative values indicate performance that was worse than the mean/comparison city. Generally, the findings show that most cities were not significantly different from the base city on the majority of outcome measures after controlling for clients’ differences. The Ordinary Least Squares regression models explained up to one third of the variance in the models, depending on the type of treatment or substance type and performance measure under consideration. Also, review of the case mix controls (covariates) in the models finds no consistent patterns either across modalities or specific performance measures (regression results not presented). Still, some consistent patterns emerged. Some cities apparently performed either better or worse than the comparison city after adjusting for client characteristics and client severity. These city differences are discussed in greater detail below.

## **2.1 Findings By Type of Treatment**

Models were first constructed that made comparisons of clients’ outcomes across cities within type of treatment. The cities did not differ significantly from the comparison city on the majority of the performance measures. There are, however, some noteworthy differences. Clients receiving methadone treatment in City 2 exhibited significantly better substance abuse treatment outcomes for four of the eight measures as compared to clients in City 1 (refer to Exhibit B-3 for more detail). In contrast, clients receiving intensive outpatient treatment in City 4 consistently had less favorable substance use outcomes than did clients receiving treatment in the other cities. These findings are relatively robust, with City 4 exhibiting poorer outcomes on four of the alcohol/drug use performance measures (refer to Exhibit B-2 for more detail). These findings are also consistent with the descriptive statistics presented earlier for all types of treatment (refer to Exhibit III-2), and City 4’s standardized ranking for substance abuse problems (6 out of 6 cities) (refer to Exhibit III-3).

After adjusting for client characteristics and severity, clients in City 5 who received treatment in non-intensive outpatient and short-term residential treatment had better substance

use outcomes than clients treated in City 1, the comparison city (refer to Exhibits B-1 and B-4 for more detail). These findings are consistent with the observations from the descriptive tabulations and standardized rankings presented earlier (refer to Exhibits III-2 and III-3). In contrast, clients in non-intensive outpatient treatment in City 6 had significantly poorer outcomes on both measures of employment than did clients in City 1 (refer to Appendix B, Exhibit B-1 for more detail). This finding illustrates the utility of multivariate analyses when assessing relative treatment system performance. Specifically, while the descriptive tabulations would suggest that City 6 had superior performance relative to the other cities on the measure of change in employment days at follow-up (an increase of 5.8 days compared to an average of 3.2 days for the other cities and a standardized ranking of 1), this city apparently had a less severely impaired population, which accounted for its apparent superior performance (refer to Exhibits III-2 and III-3 for more detail). While five of the six cities served clients in long-term residential treatment, there were few differences across cities for these clients (refer to Exhibit B-5 for more detail). Also, there were no consistent differences with regard to City 3's performance relative to City 1.

## 2.2 Findings By Primary Substance of Use

Models were also constructed that compared clients' outcomes across cities within populations defined by primary substance of use. As with the analyses by type of treatment, the analyses by primary substance of use showed few statistically significant differences across cities. While most cities did not differ from the comparison city for most performance measures, there are some differences that stood out. Across all performance measures, Cities 2, 3, and 5 had more favorable substance abuse outcomes for clients whose primary substance of use was heroin. Of the eight substance use outcome measures, Cities 2, 3, and 5 had significantly better performance on four, three, and two of the measures, respectively (refer to Exhibit B-7 for more detail). Findings were similar for these cities for clients whose primary substance of use was cocaine/crack. However, some inconsistencies arose for Cities 2, and 3. While they had more favorable outcomes for the substance use *problems* measures, these cities had less favorable outcomes for the performance measures related to substance use *days* (refer to Exhibit B-8 for more detail). Also, heroin clients in City 3 had a more favorable criminal activity outcome. This finding again illustrates the utility of using multivariate analyses because such a finding was not evident from the descriptive findings. These findings showed that City 3 had the highest proportion of clients with arrests/charges in the 6 months prior to follow-up (refer to Exhibits B-7 and III-2 for more detail).

For clients whose primary substance of use was alcohol, City 4 exhibited significantly poorer outcomes than the rest of the cities. The consistency and robustness of this finding suggests that there may be something unique about this city that accounts for these differences

and that there may be some unmeasured characteristic that is not accounted for in the models. City 4 had less favorable outcomes for nine of the eleven performance measures, indicating more days of substance use and related problems and fewer days of employment at follow-up (refer to Exhibit B-6 for more detail). Clients in City 4 whose primary substance was heroin or cocaine/crack also had consistently less favorable outcomes for both substance use and employment measures (refer to Exhibits B-7 and B-8 for more detail). These findings are consistent with City 4's standardized ranking for substance problem days (refer to Exhibit III-3).

For heroin clients, City 6 had results similar to those of City 4. It performed worse on five of the eight substance use measures and both employment measures (refer to Exhibit B-7 for more detail). These findings were consistent with the descriptive analyses presented earlier and with the finding that there were few statistically significant covariates in the models. (Clients in City 6 had less improvement in terms of substance use and substance problem days—see Exhibit III-2.) However, City 6 had no significant differences for substance use or any other outcome measures relative to City 1 for cocaine/crack clients (refer to Exhibit B-8 for more detail). City 6 only differed from City 1 on one measure for alcohol clients (refer to Exhibit B-6 for more detail). Few differences in client outcomes were evident among clients using marijuana across cities (refer to Exhibit B-9 for more detail).

### **2.3 Summary**

In comparing outcomes across cities by both type of treatment and primary substance of use, most cities did not appear to be significantly different from the comparison city, although a few noteworthy patterns emerged. Overall, this finding suggests that the treatment systems are successful in serving clients with substance abuse problems and that the improvements clients make in the domains of principal importance (substance use, employment and criminal activity) are of relatively similar magnitudes. The analysis did find that some cities do appear to perform moderately better or worse than others on specific dimensions/measures. The most obvious finding was that Cities 4 and 6 consistently had less favorable treatment outcomes. Across all types of treatment, Cities 4 and 6 had less favorable performance on 7 and 5 of the 48 and 36 total outcome measures, respectively. These results were similar when the data were analyzed by primary substance of use; Cities 4 and 6 then had less favorable performance on 18 and 10 of the 48 total measures, respectively. In addition, Cities 2 and 5 generally tended to have more favorable treatment outcomes. Across types of treatment, City 2 had significantly better outcomes for 12 out of 45 performance measures, while City 5 had better outcomes on 8 of 36. Across substance of use, Cities 2 and 5 each had more favorable outcomes on 9 of the 36 and 48 total performance measures, respectively. The analyses found inconsistent results for City 3. While it had more favorable performance on 3 of the 36 measures across all types of treatment, it

also had less favorable performance for 2 measures. Similarly, across primary substance of use, City 3 had better outcomes for 6 of the 48 measures but worse outcomes for 2 measures. It should be noted that multiple performance measures were constructed within outcome domains and that the findings on any single performance measure are of less value than the overall/summary findings. This is the case since rankings on any single performance measure/analysis may be due to chance compared to consistent patterns that emerge across multiple indicators (refer to Exhibits III-4 and III-5 for more detail).

<b>EXHIBIT III-4</b>						
<b>SUMMARY* OF OUTCOME MEASURES BY TYPE OF TREATMENT</b>						
	<b>City 1</b>	<b>City 2</b>	<b>City 3</b>	<b>City 4</b>	<b>City 5</b>	<b>City 6</b>
Non-intensive outpatient	0	2/0/7	2/0/10	0/1/11	5/0/7	0/2/10
Intensive outpatient	0	2/0/7	1/0/11	0/4/8	0/0/12	0/3/9
Methadone	0	4/0/5	n/a	1/1/10	n/a	n/a
Short-term residential	0	2/0/7	n/a	n/a	3/1/8	n/a
Long-term residential	0	2/0/7	0/2/10	0/1/11	n/a	0/0/12
Total	0	12/0/33	3/2/31	1/7/40	8/1/27	0/5/31

\* Number of measures significantly better/worse/no different than City 1; significant at least at  $p < .10$

<b>EXHIBIT III-5</b>						
<b>SUMMARY* OF OUTCOME MEASURES BY PRIMARY SUBSTANCE OF USE</b>						
	<b>City 1</b>	<b>City 2</b>	<b>City 3</b>	<b>City 4</b>	<b>City 5</b>	<b>City 6</b>
Alcohol	0	0/0/9	0/1/11	0/9/3	0/0/12	0/1/11
Heroin	0	4/0/5	4/0/8	0/4/8	2/0/10	0/7/5
Cocaine/crack	0	4/3/2	1/1/10	0/5/7	3/0/9	0/0/12
Marijuana	0	1/0/8	1/0/11	0/0/12	4/0/8	0/2/10
Total	0	9/3/24	6/2/40	0/18/30	9/0/39	0/10/38

\* Number of measures significantly better/worse/no different than City 1; significant at least at  $p < .10$

## **IV. SUMMARY AND IMPLICATIONS**

## IV. SUMMARY AND IMPLICATIONS

This final section offers the reader a summary of the purpose, scope of the analysis, findings, and relevance of this analysis. This summary is intended to assist the reader in organizing the data presented, identifying the most salient findings, and understanding their potential impact on the substance abuse treatment field. Accordingly, implications and cautions of this analysis are also presented, given the novelty of both the approach and topic.

### 1. SUMMARY OF FINDINGS

The overall goals of this analysis were to (1) identify and operationalize standardized performance measures that are applicable across treatment systems and/or geographic regions (cities); (2) conduct an illustrative performance measurement analysis that used the treatment systems within specific geographic regions as the unit of analysis; (3) assess the relative performance of the cities using the identified performance measures; and (4) highlight the uses, benefits, constraints, and data requirements for system-level performance measurement monitoring and comparisons.

Results from the analysis show that the clients served by the six cities differ along important dimensions, suggesting that multivariate techniques that adjust for case mix (i.e., differences in client characteristics and severity at intake) are warranted when comparing the performance of the treatment systems. Overall, 20 percent of the clients across all treatment systems received intensive outpatient services, and 15 percent received non-intensive outpatient services. There was considerable variability across cities, however. For example, while only 8 percent of the sample received intensive outpatient services in City 4, 41 percent of clients in City 2 received intensive outpatient services. Also, while 30 percent of the clients across all cities received treatment in inpatient settings (16% and 15%, respectively, for short- and long-term residential services), the range across cities was broad for both types of treatment. For example, only three cities in this sample had clients who were served in a short-term residential setting, and the proportion of clients receiving this type of treatment ranged from 8 percent for City 2 to 28 percent for City 5.

Equally important were the noteworthy differences across cities with respect to demographic characteristics and, in particular, the severity of the clients' symptoms when they presented for substance abuse treatment. In summary, while there was little variability across cities in terms of the average age of clients served, the populations differed considerably in terms of gender, racial composition, and marital status. The cities also differed on key measures of the severity of clients' presenting symptoms. At intake, clients in Cities 1, 2, and 3 had an average

of almost 18 days of substance use out of the past 30, while clients in Cities 4, 5, and 6 reported an average of 12 days of substance use in the last 30. A similar pattern was found for substance problems in the past 30 days as reported at intake. Cities 1, 2, and 3 appeared to have clients with more severe presenting problems than Cities 4, 5, and 6. Clients in Cities 1, 2, and 3 reported more substance problem days (20.6) than those in Cities 4, 5, and 6 (an average of 16.4 problem days). Results also showed that clients presenting for treatment had an average of only 3.8 days of paid employment in the past 30 days at intake. As with the drug/alcohol measures, there was also noteworthy variability across the cities with respect to clients' employment problems.

The descriptive results are consistent with findings from large national studies such as the National Treatment Improvement Evaluation Study (NTIES) and the Drug Abuse Treatment Outcome Studies (DATOS), showing that clients did benefit from participation in substance abuse treatment at all the cities included in this analysis. For example, there was an average of a 12-day reduction in the number of drug use days in the past 30 from intake to follow-up. There was also an average reduction of almost 13 days in the number of drug problem days from intake to follow-up. As with the client characteristics and severity measures, some cities appeared to perform better/worse on some of the performance measures, despite the overall improvements. City performance was compared on selected measures in each outcome domain that accounted for initial differences in client severity but not client demographics. Results showed inconsistent patterns across the three outcome domains. As a result of the variability in treatment population characteristics, comparisons of city performance necessitate the use of multivariate techniques to statistically account for treatment population differences in order to ensure that appropriate comparisons and conclusions are made.

The multivariate analyses showed that the cities did not differ significantly from the comparison city for most performance measures either by type of treatment or primary substance of use. In summary, the analyses found that two cities (Cities 1 and 3) were generally median performers. Two cities (Cities 4 and 6) tended to perform worse than the median, and two cities (Cities 2 and 5) tended to perform better than the median when there were statistically significant differences in performance. City 4 consistently exhibited worse outcomes than the comparison city. Clients in City 4 whose primary substance of use was alcohol had less favorable outcomes for 9 of the 11 performance measures. Clients in City 4 whose substance of choice was heroin or cocaine/crack also had consistently less favorable substance use and employment outcomes. These findings for City 4 were also consistent with the descriptive statistics and standardized rankings (City 4 ranked 6 out of 6 cities on the standardized substance use measure).

Other findings illustrated the need for multivariate analyses when comparing treatment system performance. Clients receiving non-intensive outpatient treatment in City 6 had significantly poorer outcomes on both employment measures. In contrast, the descriptive tabulations suggested that City 6 had superior performance with regard to employment outcomes. Thus, this city likely had an easier case mix, which accounted for its apparent superior performance. Heroin clients in City 3 had a more favorable criminal activity outcome relative to clients in City 1, which also illustrated the utility of multivariate analyses. This finding was not initially evident from the descriptive statistics that found that City 3 had the highest proportion of clients with arrests/charges in the 6 months prior to follow-up.

Results from this analysis illustrate an approach to operationalizing performance measures and assessing and comparing the performance of treatment systems, a challenge that has not yet been addressed in the substance abuse field. The approach has important implications for analysis and will ultimately inform both policy and substance abuse treatment practice. The implications are discussed in some detail below. There are also constraints to the analysis, and the findings should be interpreted with caution. As indicated earlier, this analysis is intended to be illustrative of an approach to comparing treatment systems on key performance measures. The Target Cities database was selected because it was the most amenable to conducting a comparative analysis of treatment systems. Thus, while it contained many of the important case mix controls identified in the literature (Ameen et al., 1999; Rogers et al., 2001; Harwood et al., 2001; and Swindle et al., 1995), the data set did not contain the full range of theoretically relevant independent variables that would be included in a prospectively designed performance measurement analysis. For instance, comprehensive system-level performance measurement analyses would also include measures of provider characteristics, length of stay, intensity of treatment, and treatment/environmental characteristics. These measures would facilitate explaining the observed differences in treatment system performance. The absence of these measures from the Target Cities data set by definition constrained the analysis to one that is largely illustrative of the approach and challenges underlying analyses of treatment system performance.

The implications of this analysis include important design considerations for outcome monitoring systems and the evaluation of treatment systems. This analysis highlights the data infrastructure requirements necessary to conduct evaluations of clients' post-treatment outcomes across treatment systems or geographic regions. It also sets the context for future, more systematic studies and analyses that would compare the performance of treatment systems within and across cities.

## **2. IMPLICATIONS FOR TREATMENT PROVIDERS**

Providers and counselors need to be aware of the standards to which treatment systems and individual providers are being held. Publicizing the performance standards may contribute to greater accountability and quality improvement. This is possible since providers and counselors are ultimately responsible for delivering the services and affecting clients' outcomes. Feedback to providers and counselors regarding the relative performance of their cities could also serve as a means of continuous quality improvement. For example, explicit recognition that some cities may be performing better than the average could serve to motivate providers and counselors to improve their own performance. Such recognition of above average performance of cities could also serve to further convey the message that treatment can and does indeed work. Furthermore, to the extent that performance measurement analyses are linked to the financing and the resources available to providers, administrators may want to be made aware that their performance is linked to and contributes to overall system resources.

Among the constraints of this specific illustrative analysis is the lack of provider and treatment characteristics. Such data would have been a useful addition, as a second step of the analysis, to begin to identify the reasons for the observed rankings. The inability to include such measures as a second step of the analysis is especially important in that it highlights the need to identify and collect a parsimonious minimum set of "case mix controls" to ensure that comparisons of systems either across time or geographic regions are standardized. Specifically, key indicators at the client and provider levels are of critical importance to facilitate valid comparisons and explain "why" there are observed differences in performance across treatment systems. Such information is a critical component in identifying and disseminating best practices and implementing quality assurance measures. Better measurement and specification of the models for predicting client outcomes would also lead to more rigorous, scientific, and credible evidence of the benefits of substance abuse treatment and de-stigmatize treatment as ineffective and unworthy of support.

This analysis highlights the importance of identifying the expectations for treatment systems, designing data collection instruments, and using data collection designs that gather the data needed to evaluate the systems' performance on an ongoing basis. This analysis finds considerable consistency in performance measures across cities. This finding is important in that it identifies commonalities in the delivery systems across different locales and geographic environments. Thus, while approaches of the cities included in this analysis were customized to serve the needs of geographic regions, populations, and the individual needs of clients, there does appear to be relative consistency in the outcomes and, to some degree, the determinants of clients' outcomes. This consistency has practical implications since it suggests a need to identify

and monitor the performance measures that are most sensitive to changes over time and that are likely to be a function of changes in either treatment patterns or the organization of the service delivery system. The identification and monitoring of key indicators over time also holds the potential for monitoring national trends in substance abuse treatment outcomes, depending on the selection of cities. For example, collecting standardized indicators across cities or treatment systems would permit monitoring changes and regional variability in treatment access. Such information is necessary for making systemic changes to ensure that priority populations (e.g., those with co-occurring disorders or the homeless) are receiving necessary services.

Furthermore, attention could also be directed toward identifying key indicators that can be used not only to establish baseline measures against which changes can be measured over time, but that also permit comparisons with other treatment systems. It is recommended that these performance measures meet certain basic requirements as outlined by Burnam (1996). Performance measures must be well-suited to populations and purposes, have good psychometric properties, have minimal burden and cost to collect, and be clearly interpretable. In practical terms, after determining the goals for the system, the data collection system/ infrastructure needs to be examined and modified, if necessary, to collect data related to the performance indicators.

### **3. IMPLICATIONS FOR POLICYMAKERS**

This analysis has implications for policymakers in that it illustrates an approach to assessing the performance of treatment systems. The approach and findings show some of the different performance measures that can be constructed from information collected by treatment providers. The analysis also points to the significant benefits that can result from *a priori* consideration of the goals of treatment systems and the identification and establishment of key performance measures that could be used to monitor performance over time and across treatment systems. Comparative analyses of this type highlight the potential benefits of coordinating performance measurement indicators across systems. Specifically, lessons learned from treatment systems with particularly good performance can be used to inform the practices and designs of other systems. It also accentuates the importance of prospectively identifying the ways in which evaluation data will be used and ensuring that the data collected will directly map to the intended analyses. Such forethought will contribute to rigorous analyses with direct policy applications, including setting realistic and attainable goals that can be monitored and changed over time as the treatment system evolves.

#### **4. IMPLICATIONS FOR RESEARCHERS/EVALUATORS**

Several areas for additional data collection and analysis with respect to assessing the performance of treatment systems are possible. First, additional systematic evaluation and analysis is needed on identifying the most reliable and valid performance measures within important outcome domains that can be used to measure the performance of treatment systems. There is a considerable body of literature in the performance measurement arena that uses the Addiction Severity Index (ASI) as the basis for developing performance measures. These measures have been used at the individual provider level, but have not previously been systematically applied to the assessment of treatment system performance. In this analysis, several performance measures in the drug/alcohol outcome domain were used, and relative consistency was found across the different formulations of the performance measures. Additional research that identifies performance measures that are applicable and useful for an analysis of treatment systems is needed. These measures could possibly be based on the ASI composite scores or its respective individual components.

A related line of inquiry is the identification of a core parsimonious set of independent variables (case mix controls) that can be used when conducting comparative analyses of the performance of treatment systems. Such a list might include the minimum indicators to be included in models in order to ensure appropriate comparisons.

Future analyses that explore alternative ways of combining performance measures to create overall/global measures of the performance of treatment systems are possible. This analysis summed up the performance of cities across the 11 performance measures. Analyses that explore alternative ways of creating global measures and indices, including possible weighting schemes, could be conducted. Such analyses would be especially useful because, as indicated earlier, the use of multiple performance indicators reduces the possibility that findings are due to chance and enhances the validity of the findings. Given the direct policy and practical applications of performance measurement research, efforts to enhance rigor and ensure the validity of findings are especially important.

Many of the analytic issues and questions raised by this preliminary analysis can not be addressed until an analytic database is available that contains comparably collected data across differing treatment systems. Information from the Cooperative Agreements for State Treatment Outcomes and Performance Pilot Studies Enhancement: TOPPS I & II initiative or subsets of states funded under that initiative may support some of the analytic issues identified above. Furthermore, the widespread adoption of the ASI by many states is leading to the routine collection of standardized information, at least at intake, through their outcome monitoring

systems. State OMS data could be combined with administrative data (notably, employment, earnings, welfare, and criminal activity) to begin to address the relative performance of treatment systems.

## **5. CONSTRAINTS OF FINDINGS**

The principal intent of this analysis was to illustrate an approach to, and the utility and implications of, comparing the relative performance of the treatment systems operating in differing cities. As such, the analysis used a data source (the CSAT Target Cities initiative) that was most amenable to cross-system comparisons. While the data set does indeed permit cross-city comparisons, the ability to rigorously illustrate the full range of analyses and specify models in a manner that captures the most relevant factors affecting the cities' performance was limited by the available data and differences in the respective cities' study designs. The constraints of the data and analysis are especially important because they will sensitize the reader to, and underscore, critical issues that state and Federal policymakers and providers need to consider when designing evaluations of either demonstrations or enhancements to outcome monitoring systems. The overarching cautions stemming from the constraints imposed by the data and analysis pertain to the interpretation of this analysis rather than the utility of the approach to comparisons of treatment systems.

One of the most important constraints of these findings is that the Target Cities multi-site data collection efforts were not organized and undertaken prospectively. Rather, the multi-site database was compiled after cities had already begun their own data collection efforts. Thus, cities submitted data for inclusion in the multi-site data set only to the extent that they had them available (i.e., had already collected them). Also, each site used different data collection instruments and had unique study designs. Although five of the six cities included elements of the ASI in their data collection instruments, each city also had its own set of distinct assessment questions and response items. Therefore, no city was able to submit data for every single variable included in the multi-site data set. In addition, the data for those variables reported by all cities are not standard due to differences in question wording and/or response categories. As noted earlier, attempts were made to minimize the effect of these differences through the review of variable annotations and the selection of those variables that were measured in a standardized manner across the six cities. Each city collected its initial data over somewhat different periods of time and for different durations. Furthermore, follow-up data collection periods were not standardized across the cities. While the majority of the cities collected data at 6 months post-baseline, the other time points at which follow-up data were (or were not) collected varied considerably across the cities. This lack of standardization with respect to variables, response

categories, study design, the timing of the study implementation, and follow-up periods introduces a fairly significant potential source of variation into the analyses.

Related to the variation in specific data items and responses is the fact that each city also had a slightly different emphasis for its Target Cities project, and the data collected reflect their differing study foci. For example, some cities over-selected female participants, while the criminal justice system was an integral part of the project in another city. Due to these differing project constructions across cities, participants were not available for certain types of treatment in a city, or smaller than ideal sample sizes were available for types of treatment and/or cities. For example, only three of the six cities provided data for clients in methadone treatment. These differences also presented challenges in interpreting the findings. For instance, the Target Cities documentation did not have any information on how females were over-selected in the cities—whether this was through the selection of single-sex treatment providers or other means. The literature (Grella, Perry, & Anglin, 1996) remains inconclusive with regard to whether there are significant differences in treatment outcomes for attendees of single-sex versus co-ed treatment facilities. There was also no information regarding each city’s methadone treatment approach. While the literature supports high dosage, long-term maintenance for realizing the most favorable treatment outcomes, the cities included may have different policies concerning detoxification and short-term treatment. These differences have important implications for the findings. Thus, while the Target Cities data set and accompanying documentation are very detailed and contain a wealth of information, it was not possible to obtain complete information on all aspects of each city’s design and associated implementation issues. Consequently, while significant differences across cities were observed, there is not sufficient information to interpret or to make educated speculations with regard to “why” these apparent differences exist.

In spite of the considerable amount of documentation that accompanied the Target Cities data set, the documentation lacked information about issues such as provider recruitment, provider characteristics, and characteristics of the treatment systems within each of the six cities. For some cities, there was no information on the total number of providers represented or the number of providers by type of treatment. The documentation did not permit discerning whether the providers selected for inclusion in the Target Cities initiative/database are representative of the treatment systems in each city. The development of standardized data collection procedures and a core minimum set of data elements holds advantages both with respect to increasing the analytic utility of the data as well as reducing the documentation burden associated with different measures for variables.

Another issue concerns the fact that a large number of clients reported no use or days with substance use problems at intake. The lack of provider-level data and characteristics

impeded an ability to discern whether this phenomenon of high levels of abstinence at intake is due to requirements of providers (e.g., some providers won't accept clients for treatment until they have been clean for some period of time), clients served in criminal justice settings, or other factors such as differences in the data collection instrument, training of the interviewers, or systematic differences between clients.

In addition, the Target Cities multi-site data set does not contain any data that allowed length of stay or duration of treatment to be calculated, which is an important variable in treatment success. While the data set contains an admission date for each client, discharge dates are not included in the data set. This lack of information constrains the ability to address whether the outcomes may differ across cities due to length of stay. Similarly, there was no way to determine any movement by clients between levels of care for the given treatment episode. The data set also lacks information on treatment intensity and the nature of services provided. Therefore, this analysis is unable to determine any differences across cities in terms of hours of care, client-to-staff ratios, and other important factors that would affect client outcomes. These are all variables critical to assessing outcomes and aiding in systems improvements, especially if efficiencies are to be attained to aid in closing the treatment gap. These constraints of the data set further limit the ability to interpret the patterns observed. These "constraints" do, however, call attention to specific needs and requirements for outcome monitoring systems and multi-site, preferably cross-site, evaluation designs.

In summary, the principal constraints of this analysis stem from the lack of measures on the full range of relevant factors that potentially affect clients' post-treatment outcomes. This analysis was able to include the most important client-level case mix controls, but did not have information related to providers, their characteristics, or the social/ecological environments of the cities. The analysis was able to discern whether cities differed significantly on important performance measures. Additional information with regard to how representative the providers within cities were is also critical. Furthermore, while the analysis was able to identify some consistent differences across cities, information such as length of stay, treatment intensity/quality, and provider characteristics would have allowed it to address the issue of "why" some cities perform better than others. These results suggest a need for additional quantitative and qualitative analyses to address these issues.

## **REFERENCES**

## REFERENCES

- Agosti, V., Nunes, E., Stewart, J.W., & Quitkin, F.M. (1991). Patient factors related to early attrition from an outpatient research clinic: A preliminary report. *International Journal of Addictions*, 26, 327-334.
- Ameen, A.Z., Dall, T.M., Harwood, H.J., & Doyon, V.A. (1999). *Comparisons of substance abuse provider effectiveness: An analysis using case mix adjustment*. Unpublished manuscript.
- Burnam, M.A. (1996). Measuring outcomes of care for substance use and mental disorders. *New Directions for Mental Health Services*, 71, 3-17.
- Children's Health Act of 2000. Public Law 106-310. October 17, 2000.
- Claus, R., & Dailey, R.M. (In press, 2002). The Target Cities participants: From centralized intake to treatment entry. In R. Stephens, C. Scott, & R. Muck (Eds.), *Clinical assessment and substance abuse treatment: The Target Cities experience* (pp. 16-48). Rockville, MD: Center for Substance Abuse Treatment, Substance Abuse and Mental Health Services Administration.
- Department of Health and Human Services (DHHS), Substance Abuse and Mental Health Services Administration (SAMHSA). Request for Proposal (RFP) No. 270-02-7109. "National Treatment Outcomes Monitoring System (NTOMS)." March 21, 2002.
- Government Performance and Results Act of 1993 (GPRA). Public Law 103-62. August 3, 1993.
- Grella, C.E., Perry, S.M., & Anglin, M.D. (1996). *Drug and alcohol treatment programs for women: Client characteristics, program characteristics, and treatment outcomes*. Prepared under NEDTAC Contract No. 270-94-0001. Fairfax, VA: Caliber Associates.
- Harwood, H.J., Field, T., & Tunkelrott, T. (2001). *Evaluating the performance of substance abuse detoxification providers using treatment engagement as an outcome measure* (NIDA Contract No. P50 DA10233). Falls Church, VA: The Lewin Group.
- Hawkins, J.D. (1983). Community characteristics associated with treatment outcome. In J.R. Cooper, F. Altman, B.S. Brown, & D. Czechowicz (Eds.), *Research on the treatment of narcotic addiction: State of the art* (pp. 680-800). Washington, DC: US Government Printing Office.
- Hendryx, M.S., Dyck, D.G., & Srebnik, D. (1999). Risk-adjusted outcome models for public mental health outpatient programs. *HSR: Health Services Research*, 34, 1, 171-195.

- Howard, D.L., LaVeist, T.A., & McCaughrin, W.C. (1996). The effect of social environment on treatment outcomes in outpatient substance misuse treatment organizations: Does race really matter? *Substance Use & Misuse*, 31, 5, 617-638.
- Humphreys, K., Moos, R.H., & Cohen, C. (1997). Social and community resources and long-term recovery from treated and untreated alcoholism. *Journal of Studies on Alcohol*, 58, 231-238.
- Joe, G.W., Singh, B.K., Finklea, D., Hudiburg, R., & Sells, S.B. (1977). *Community factors, racial composition of treatment programs and outcomes* (DHEW Publication No. ADM 78-573). Washington, DC: National Institute on Drug Abuse.
- Koenig, L., Fields, E.L., Dall, T.M., Ameen, A.Z., & Harwood, H.J. (2000). *Using case mix adjustment methods to measure the effectiveness of substance abuse treatment: Three examples using client employment outcomes* (NEDS Contract No. 270-97-7016). Fairfax, VA: Caliber Associates.
- Leahy P., Stephens, R., Huff, H., & Kaye, R. (In press, 2002). Methodological issues in the development of the Target Cities multi-site databases. In R. Stephens, C. Scott, & R. Muck (Eds.), *Clinical assessment and substance abuse treatment: The Target Cities experience* (pp. 16-48). Washington, DC: Center for Substance Abuse Treatment, Substance Abuse and Mental Health Services Administration.
- Leeper, T., Arria, A., & Krupski, T. "ICS Data Analysis Discussion." Presentation given at the Treatment Outcomes and Performance Pilot Studies Enhancement TOPPS II Grantee Meeting. Chicago, IL, October 30-31, 2000.
- Leigh, G., Ogborne, A.C., & Cleland, P. (1984). Factors associated with patient dropout from an outpatient alcoholism treatment service. *Journal of Studies on Alcohol*, 45, 4, 359-362.
- McCaughrin, W.C., & Price, R.H. (1992). Effective outpatient drug treatment organizations' program features and selection effects. *International Journal of the Addictions*, 27, 11, 1335-1358.
- Moos, R.H. (1996). Understanding environments: The key to improving social processes and program outcomes. *American Journal of Community Psychology*, 24, 1, 193-201.
- Ornstein, P., & Cherepon, J.A. (1985). Demographic variables as predictors of alcoholism treatment outcome. *Journal of Studies on Alcohol*, 46, 5, 425-432.
- Reichman, M., Rogers, S., Ameen, A.Z., & Harwood, H.J. (2000). *Review of literature on ecological determinants of substance abuse treatment outcomes for the TOPPS II Utah Consortium* (prepared under contract to the TOPPS II Utah Consortium).

- Rogers, S., Ameen A.Z., & Harwood, H.J. (2001). *An analysis of relative referral success for residential treatment providers in the state of Oklahoma* (prepared under the CSAT TOPPS II project). Falls Church, VA: The Lewin Group.
- Singh, B.K. (1981). The effects of community structure on during-treatment outcomes of methadone maintenance programs. *The International Journal of the Addictions*, 16, 7, 1183-1196.
- Stephens, R.C., Kaye, R., & Chen, H. (1999). Establishing a Target Cities model in Cleveland. *The Journal of Psychoactive Drugs*, 31, 3, 219-224.
- Stephens, R.C., Scott, C.K., & Muck, R.D. (In press, 2002). *Clinical assessment and substance abuse treatment: The Target Cities experience*. Rockville, MD: Center for Substance Abuse Treatment, Substance Abuse and Mental Health Services Administration.
- Substance Abuse and Mental Health Services Administration. (1995). *Developing state outcomes monitoring systems for alcohol and other drug abuse treatment*. Treatment Improvement Protocol (TIP) Services 14, DHHS Publication No. (SMA) 95-3021. Rockville, MD: U.S. Department of Health and Human Services.
- Utah Consortium. (1998). TOPPS II proposal.
- Zold-Kilbourn, P., Tucker, T.C., & Berry, J. (1999). Improving substance abuse treatment for indigent clients in Detroit. *The Journal of Psychoactive Drugs*, 31, 3, 233-239.

**APPENDIX A**  
**REGRESSION MODEL SPECIFICATIONS**

## APPENDIX A

### REGRESSION MODEL SPECIFICATIONS

Performance Measure	Independent Variables Included in Model
Number of substance use problem days in past 30	Age, age <sup>2</sup> , sex, race/ethnicity, marital status, days of alcohol use in past 30 at intake, days of heroin use in past 30 at intake, days of cocaine use in past 30 at intake, days of marijuana use in past 30 at intake, days with alcohol problems in past 30 at intake, days with drug problems in past 30 at intake, prior inpatient psychiatric treatment in lifetime, prior outpatient psychiatric treatment in lifetime, probation/parole status at intake, admission prompted by the legal system, arrested and charged with shoplifting/vandalism in lifetime, arrested and charged with drug offenses in lifetime, arrested and charged with assault in lifetime, arrested and charged with robbery in lifetime, arrested and charged with burglary in lifetime, arrested and charged with homicide/manslaughter in lifetime, whether 12-month follow-up data were used, type of treatment (for models by drug of use only), drug of use (for models by type of treatment only)
Change in number of substance use problem days in past 30	Same as for number of substance use problem days in past 30
Proportional change in number of substance use problem days in past 30	Same as for number of substance use problem days in past 30
Absence of substance use problem days in past 30	Same as for number of substance use problem days in past 30
Number of days of substance use in past 30	Same as for number of substance use problem days in past 30
Change in number of days of substance use in past 30	Same as for number of substance use problem days in past 30
Proportional change in number of days of substance use in past 30	Same as for number of substance use problem days in past 30
Absence of substance use days in past 30	Same as for number of substance use problem days in past 30
Number of days of paid employment in past 30	Same as for number of substance use problem days in past 30 in addition to days of paid employment in past 30 at intake and number of months of longest full-time job
Change in number of days of paid employment in past 30	Same as for number of substance use problem days in past 30 in addition to days of paid employment in past 30 at intake and number of months of longest full-time job
Arrested and charged with one of 10 crimes in past 6 or 12 months	Same as for number of substance use problem days in past 30 excluding variable indicating whether 12 month follow-up data were used

**APPENDIX B**  
**REGRESSION FINDINGS**

## APPENDIX B REGRESSION FINDINGS

<b>EXHIBIT B-1 OUTCOME MEASURES COMPARED TO CITY 1* FOR CLIENTS IN NON-INTENSIVE OUTPATIENT TREATMENT</b>						
	City 1	City 2	City 3	City 4	City 5	City 6
Substance use problem days	0				+2	
Change in problem days	0		+2		+2	
Proportional change in problem days	0					
Absence of problem days	0	+1				
Days of substance use	0				+1	
Change in days of use	0		+1		+1	
Proportional change in days of use	0			-2		
Abstinence/absence of days of use	0					
Days of paid employment	0	n/a				-2
Change in days of paid employment	0	n/a				-2
Arrested & charged with crime	0	+1			+1	

\*+= more favorable outcome than City 1; -= less favorable outcome than City 1; +1/-1 significant at  $p < .10$ ; +2/-2 significant at  $p < .05$ ; blank cell indicates no significant difference; 0 indicates base city; n/a indicates data were not collected by city

<b>EXHIBIT B-2</b>						
<b>OUTCOME MEASURES COMPARED TO CITY 1*</b>						
<b>FOR CLIENTS IN INTENSIVE OUTPATIENT TREATMENT</b>						
	City 1	City 2	City 3	City 4	City 5	City 6
Substance use problem days	0			-2		-2
Change in problem days	0			-1		-1
Proportional change in problem days	0			-2		
Absence of problem days	0	+1	+1			-2
Days of substance use	0					
Change in days of use	0					
Proportional change in days of use	0			-2		
Absence of days of use	0					
Days of paid employment	0	n/a				
Change in days of paid employment	0	n/a				
Arrested & charged with crime	0	+2				

\*+= more favorable outcome than City 1; -= less favorable outcome than City 1; +1/-1 significant at p<.10; +2/-2 significant at p<.05; blank cell indicates no significant difference; 0 indicates base city; n/a indicates data were not collected by city

<b>EXHIBIT B-3</b>						
<b>OUTCOME MEASURES COMPARED TO CITY 1*</b>						
<b>FOR CLIENTS IN METHADONE TREATMENT</b>						
	City 1	City 2	City 3	City 4	City 5	City 6
Substance use problem days	0	+2	n/a		n/a	n/a
Change in problem days	0	+2	n/a		n/a	n/a
Proportional change in problem days	0		n/a	+2	n/a	n/a
Absence of problem days	0	+2	n/a		n/a	n/a
Days of substance use	0		n/a		n/a	n/a
Change in days of use	0		n/a		n/a	n/a
Proportional change in days of use	0		n/a		n/a	n/a
Absence of days of use	0	+2	n/a		n/a	n/a
Days of paid employment	0	n/a	n/a		n/a	n/a
Change in days of paid employment	0	n/a	n/a	-1	n/a	n/a
Arrested & charged with crime	0		n/a		n/a	n/a

\*+= more favorable outcome than City 1; -= less favorable outcome than City 1; +1/-1 significant at p<.10; +2/-2 significant at p<.05; blank cell indicates no significant difference; 0 indicates base city; n/a indicates data were not collected by city

<b>EXHIBIT B-4</b>						
<b>OUTCOME MEASURES COMPARED TO CITY 1*</b>						
<b>FOR CLIENTS IN SHORT-TERM RESIDENTIAL TREATMENT</b>						
	City 1	City 2	City 3	City 4	City 5	City 6
Substance use problem days	0	+2	n/a	n/a	+2	n/a
Change in problem days	0		n/a	n/a	+2	n/a
Proportional change in problem days	0		n/a	n/a		n/a
Absence of problem days	0	+2	n/a	n/a		n/a
Days of substance use	0		n/a	n/a		n/a
Change in days of use	0		n/a	n/a	+2	n/a
Proportional change in days of use	0		n/a	n/a		n/a
Absence of days of use	0		n/a	n/a		n/a
Days of paid employment	0	n/a	n/a	n/a		n/a
Change in days of paid employment	0	n/a	n/a	n/a		n/a
Arrested & charged with crime	0		n/a	n/a	-1	n/a

\*+= more favorable outcome than City 1; -= less favorable outcome than City 1; +1/-1 significant at  $p < .10$ ; +2/-2 significant at  $p < .05$ ; blank cell indicates no significant difference; 0 indicates base city; n/a indicates data were not collected by city

<b>EXHIBIT B-5</b>						
<b>OUTCOME MEASURES COMPARED TO CITY 1*</b>						
<b>FOR CLIENTS IN LONG-TERM RESIDENTIAL TREATMENT</b>						
	City 1	City 2	City 3	City 4	City 5	City 6
Substance use problem days	0	+1			n/a	
Change in problem days	0				n/a	
Proportional change in problem days	0				n/a	
Absence of problem days	0	+2			n/a	
Days of substance use	0		-1		n/a	
Change in days of use	0			-1	n/a	
Proportional change in days of use	0				n/a	
Absence of days of use	0		-1		n/a	
Days of paid employment	0	n/a			n/a	
Change in days of paid employment	0	n/a			n/a	
Arrested & charged with crime	0				n/a	

\*+= more favorable outcome than City 1; -= less favorable outcome than City 1; +1/-1 significant at  $p < .10$ ; +2/-2 significant at  $p < .05$ ; blank cell indicates no significant difference; 0 indicates base city; n/a indicates data were not collected by city

<b>EXHIBIT B-6</b>						
<b>OUTCOME MEASURES COMPARED TO CITY 1*</b>						
<b>FOR CLIENTS WITH ALCOHOL AS PRIMARY SUBSTANCE OF USE</b>						
	City 1	City 2	City 3	City 4	City 5	City 6
Substance use problem days	0			-2		
Change in problem days	0			-2		
Proportional change in problem days	0					
Absence of problem days	0			-2		-2
Days of substance use	0			-1		
Change in days of use	0			-1		
Proportional change in days of use	0			-2		
Absence of days of use	0			-2		
Days of paid employment	0	n/a		-1		
Change in days of paid employment	0	n/a		-1		
Arrested & charged with crime	0		-2			

\*+= more favorable outcome than City 1; -= less favorable outcome than City 1; +1/-1 significant at  $p < .10$ ; +2/-2 significant at  $p < .05$ ; blank cell indicates no significant difference; 0 indicates base city; n/a indicates data were not collected by city

<b>EXHIBIT B-7</b>						
<b>OUTCOME MEASURES COMPARED TO CITY 1*</b>						
<b>FOR CLIENTS WITH HEROIN AS PRIMARY SUBSTANCE OF USE</b>						
	City 1	City 2	City 3	City 4	City 5	City 6
Substance use problem days	0	+2	+1	-2	+1	-2
Change in problem days	0	+2	+1	-2	+1	-2
Proportional change in problem days	0					-1
Absence of problem days	0	+2	+2			-2
Days of substance use	0					
Change in days of use	0					
Proportional change in days of use	0					-1
Absence of days of use	0	+2				
Days of paid employment	0	n/a		-1		-2
Change in days of paid employment	0	n/a		-1		-2
Arrested & charged with crime	0		+2			

\*+= more favorable outcome than City 1; -= less favorable outcome than City 1; +1/-1 significant at  $p < .10$ ; +2/-2 significant at  $p < .05$ ; blank cell indicates no significant difference; 0 indicates base city; n/a indicates data were not collected by city

<b>EXHIBIT B-8</b>						
<b>OUTCOME MEASURES COMPARED TO CITY 1*</b>						
<b>FOR CLIENTS WITH COCAINE/CRACK AS PRIMARY SUBSTANCE OF USE</b>						
	City 1	City 2	City 3	City 4	City 5	City 6
Substance use problem days	0	+2			+2	
Change in problem days	0	+2			+2	
Proportional change in problem days	0			-2	+1	
Absence of problem days	0	+2	+2			
Days of substance use	0	-2		-2		
Change in days of use	0	-2		-2		
Proportional change in days of use	0			-2		
Absence of days of use	0	-2	-2	-1		
Days of paid employment	0	n/a				
Change in days of paid employment	0	n/a				
Arrested & charged with crime	0	+2				

\*+= more favorable outcome than City 1; -= less favorable outcome than City 1; +1/-1 significant at  $p < .10$ ; +2/-2 significant at  $p < .05$ ; blank cell indicates no significant difference; 0 indicates base city; n/a indicates data were not collected by city

<b>EXHIBIT B-9</b>						
<b>OUTCOME MEASURES COMPARED TO CITY 1*</b>						
<b>FOR CLIENTS WITH MARIJUANA AS PRIMARY SUBSTANCE OF USE</b>						
	City 1	City 2	City 3	City 4	City 5	City 6
Substance use problem days	0					
Change in problem days	0					
Proportional change in problem days	0				+2	
Absence of problem days	0	+2	+1			
Days of substance use	0				+2	
Change in days of use	0				+2	
Proportional change in days of use	0					
Absence of days of use	0				+2	
Days of paid employment	0	n/a				-1
Change in days of paid employment	0	n/a				-1
Arrested & charged with crime	0					

\*+= more favorable outcome than City 1; -= less favorable outcome than City 1; +1/-1 significant at  $p < .10$ ; +2/-2 significant at  $p < .05$ ; blank cell indicates no significant difference; 0 indicates base city; n/a indicates data were not collected by city