

Article

## Outcomes for women with co-occurring disorders and trauma: Program-level effects

Joseph J. Coccozza, (Ph.D.)<sup>a,\*</sup>, Elizabeth W. Jackson, (Ph.D.)<sup>b</sup>, Karen Hennigan, (Ph.D.)<sup>c</sup>,  
Joseph P. Morrissey, (Ph.D.)<sup>b</sup>, Beth Glover Reed, (Ph.D.)<sup>d</sup>,  
Roger Fallot, (Ph.D.)<sup>e</sup>, and Steve Banks, (Ph.D.)<sup>a</sup>

<sup>a</sup>Policy Research Associates, Inc., Delmar, NY 12054, USA

<sup>b</sup>Cecil G. Sheps Center for Health Services Research, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

<sup>c</sup>Social Science Research Institute, Inc., University of Southern California, Los Angeles, CA 90089, USA

<sup>d</sup>Social Work and Women's Studies, University of Michigan, Ann Arbor, MI 48109, USA

<sup>e</sup>Community Connections, Washington, DC 20003, USA

Received 11 March 2004; received in revised form 8 July 2004; accepted 19 August 2004

### Abstract

Program-level effects at 6 months are reported from meta-analysis of a nine-site quasi-experimental study of comprehensive, integrated, trauma-informed, and consumer-involved services for women who have mental health problems, substance use disorders, and who have experienced interpersonal violence.

The average weighted effect size is significant for the treatment condition for improved post-traumatic symptoms ( $p < 0.02$ ), drug use problem severity ( $p < 0.02$ ), and nearly significant for mental health symptoms ( $p < 0.06$ ). There is significant heterogeneity in effect sizes across sites.

Program-level variables were examined in an effort to explain this heterogeneity. The findings indicate that sites which provided significantly more integrated counseling produced more favorable results in mental health symptoms ( $p < 0.01$ ) and both alcohol ( $p < 0.001$ ) and drug use problem severity ( $p < 0.001$ ). The same trend is observable for reductions in post-traumatic stress symptoms, although the difference does not attain statistical significance. © 2005 Elsevier Inc. All rights reserved.

**Keywords:** Women's services; Co-occurring disorders; Trauma; Service integration; Meta-analysis

## 1. Introduction

### 1.1. Overview

The Women, Co-occurring Disorders and Violence Study (WCDVS) represents the first major federal effort to address the significant lack of appropriate services for women with co-occurring mental health and substance use disorders who also have histories of physical and sexual

abuse. The primary goal of the WCDVS was the generation and application of empirical knowledge on developing new service approaches, and the effectiveness of these approaches for women who have traditionally been “high-end” users of publicly-funded services.

The Substance Abuse and Mental Health Services Administration (SAMHSA) provided support to nine sites across the country to develop, implement and evaluate new interventions. In particular, study sites were required to develop interventions with four key program characteristics. They were to: (1) provide a *comprehensive* range of services, that (2) were *integrated*, (3) *trauma-informed*, and (4) involved the *consumer/survivor/recovering women* (CSR) into the work of the project.

\* Corresponding author. Policy Research Associates, Inc., 345 Delaware Avenue, Delmar, NY 12054, USA. Tel.: +1 518 439 7415.

E-mail address: jcoccozza@prainc.com (J.J. Coccozza).

The study sought to assess the impact of this approach in comparison to those receiving services in usual care settings. This examination was conducted through a cross-site, multi-model intervention study with quasi-experimental comparison groups. At 6 months, the sample consisted of 2,006 women, 1,023 in the intervention condition and 983 in the usual care condition. The study used a common interview protocol focused on four primary outcome variables: mental health status, post-traumatic symptoms, drug use problem severity and alcohol use problem severity. A more detailed description of the study can be found in [McHugo et al. \(2005\)](#).

The primary purpose of this paper is to answer the following three research questions:

1. Are there any outcome differences on the four primary outcome measures at 6 months between the women in the intervention conditions who received integrated, comprehensive, trauma-informed and CSR-involved treatment and those in the care-as-usual conditions?
2. Do outcome effects vary across sites?
3. Can site variations be explained by differences in key program characteristics?

### 1.2. Background

Much of the foundation for the study was based on the growing recognition of the multiple and complex needs of women with co-occurring disorders and histories of interpersonal abuse. In order to more effectively address these needs, the study encouraged the intervention condition to provide and integrate a wide range of mental health, substance abuse, trauma-specific and other related services, to ensure that all services were trauma-informed and to involve consumer/survivor/recovering women in all aspects of the project.

The call for better service integration in response to persons experiencing multiple problems has existed for decades ([Kahn & Kamerman, 1992](#)). Supported by various federal agencies and private foundations, efforts to improve the integration of service delivery systems have been implemented in a number of areas including homelessness, mental health, substance abuse, education, children and family services, and health and others ([Dennis, Steadman, & Cocozza, 2000](#)). The common factor in all of these efforts has been the focus on better coordination of services for groups with multiple needs.

Despite the consensus surrounding the importance of better integrating systems for these populations, studies have failed to demonstrate that they produce better outcomes for participants. The nine-site demonstration on centralized mental health authorities ([Goldman, Morrissey, & Ridgely, 1994](#)) resulted in little impact on clients. Similarly, both the Fort Bragg Study in North Carolina ([Behr, 1988](#); [Bickman et al., 1995](#)) and the Stark County, Ohio ([Bickman, Summerfelt, & Noser, 1997](#)) work found

clinical outcomes, such as functioning and symptoms, were no different between youth with serious emotional disturbance served by more comprehensive, integrated interventions and those receiving usual care. ACCESS, a 5-year federally funded demonstration program that assessed the impact of integrated systems of care for homeless persons with mental illness, “found no evidence that technical support and allocation of funds for system integration efforts improved client outcomes” ([Rosenheck et al., 2001](#)).

Several of these studies suggest that a critical factor in explaining these findings may be the disconnect between systems integration and clinical outcomes. As [Bickman and colleagues \(1997\)](#) write, “The chain of logical connections between system reform and clinical outcomes is too large. It is difficult for actions at the systems level to manifest themselves through change in services and clinicians-client interactions”. In a somewhat similar view, [Rosenheck and colleagues \(2001\)](#) view the possibility that “bottom up” strategies for integrating service delivery that focus on the client level may be able to create “virtually” integrated services for clients regardless of the operation of the larger service systems.

This body of research influenced the direction of the WCDVS. During Phase I of the study, it became clear to the sites and the federal funding agencies that the study could not achieve all things (innovative service interventions, integrated services, and fully integrated systems of care) through a single 5-year initiative ([Policy Research Associates, 2000](#)). This concern, coupled with awareness of recent research findings and the growing focus on client-level integration, led to a major mid-course change in the goals of the WCDVS with emphasis being placed by the study on client-level integration. This shift involved the differentiation between *systems* integration and *services* integration. Consistent with [Kahn and Kamerman \(1992\)](#) a distinction was made between *administrative level strategies* aimed at changing *service delivery systems for a defined population as a whole* and *case-oriented strategies* designed to change *services delivery for individual clients*. Thus, systems integration is seen as involving a set of agencies that work together in order to better share resources, information, and clients to improve direct services to the target population ([Yessian, 1995](#)). Services integration on the other hand, is seen as a case-oriented integration designed to better serve an individual. In such an arrangement, services are integrated or coordinated, but relationships between agencies may not fundamentally change. Using this distinction, sites were instructed to deemphasize their efforts around systems-level integration, but to continue to stress better client-level services integration.

While the WCDVS study sought to provide and coordinate a wide variety of core services, emphasis was clearly placed on integrating mental health, substance abuse and trauma services.

A recent report by [SAMHSA \(2002\)](#) to Congress estimated that 7 to 10 million adults in the U.S. are experiencing

co-occurring mental and substance abuse disorders each year. The report's call for better integration of mental health and substance abuse services reflects findings that the two separate systems are not addressing the problems of co-occurring disorders effectively (Ridgely, Goldman, & Willenbring, 1990). Studies of integrated approaches at the clinical level for individuals with co-occurring disorders have shown an impact on clients. As summarized by Drake et al. (1998) a series of recent studies provide encouraging evidence for the effectiveness of integrated treatment of dual disorders.

For women, their exposure to violence has emerged as an additional critical factor in treatment. As discussed elsewhere (McHugo et al., 2005), violence in the lives of women has been described as "epidemic" (Browne & Bassuk, 1997) with 20–30% of women reporting victimization during their lifetime (Commonwealth Fund, 1997). Furthermore, research has demonstrated a relationship between mental illness, substance abuse, and trauma among women (Alexander, 1996). Given the prevalence and impact of physical and sexual abuse in the lives of women receiving mental health and substance abuse services, the need for accessible and effective programs for trauma survivors has been noted by consumer/survivors (Prescott, 2001), by mental health system administrators (NASMHPD, 1998), and by service providers (Harris, 1994). Given this, the WVCDVS required that the intervention groups implement and integrate trauma services along with mental health and substance abuse services as part of their intervention. While research on the impact of trauma programs is just beginning, findings on several of the standardized approaches for providing trauma specific services appear promising (Fallot & Harris, 2004; Kruse, 1999; Najavits, Weiss, Shaw, & Muenz, 1998).

In addition to integrating mental health, substance abuse and trauma services, intervention sites had to meet two other requirements. First, they had to provide a comprehensive range of services. The core set of services were defined to include: mental health treatment, substance abuse treatment, trauma-specific services, outreach and engagement, screening and assessment, parenting skills training, resource coordination and advocacy, crisis intervention, and peer-run services (McHugo et al., 2005). Second, study intervention sites were required to involve women who experienced co-occurring disorders and physical and sexual abuse in all aspects of their work. Within this project, these women asked to be referred to as Consumer/Survivor/Recovering persons, a term developed within the first 2 years of the study to reflect their histories as mental health service consumers, survivors of physical and sexual abuse, and recovering users of drugs and alcohol. This requirement is consistent with the position of various authors in the mental health field, who have noted that including people with direct experience with the services and the unique understanding of these systems that they bring can improve the quality of services (Campbell, 1997; Deegan,

1995; Fisher, 1994a, 1994b; McCabe & Unzicker, 1995; NASMHPD, 1998).

## 2. Methods

### 2.1. Overview

The 6-month results presented in this paper represent the first set of cross-site outcomes published from the WCDVS. Nine sites participated in the study. Each utilized a quasi-experimental research design and independently recruited women for the intervention and comparison conditions. An intent-to-treat approach is used in the study. Thus, for analysis purposes women are retained in the study condition (intervention or comparison) regardless of their subsequent continued participation in services during the study period.

Projects in the intervention conditions were required to provide an integrated, comprehensive range of services with a particular emphasis on mental health, substance abuse and trauma services and to involve CSRs in their efforts. While the nine intervention projects varied widely in how they actually developed and implemented their strategies, all of these program characteristics were consistently incorporated into their service delivery strategy (McHugo et al., 2005). Projects in the comparison condition provided usual care. These projects were selected by the grantees to represent similar programs, serving similar populations in the same geographical area. Unlike the intervention projects, they did not receive any additional support for SAMHSA under the grant to provide more comprehensive and integrated service that were trauma-informed and involved CSRs. Rather their programs and services reflected what was available in communities under usual conditions.

Women participated in interviews at baseline, 3, 6, 9, and 12 months. Outcome measures were administered at baseline, 6, and 12 months, while service use measures were collected at all time points. The following instruments were used to assess the primary outcomes:

*Mental health status* was assessed by the Global Severity Index (GSI) from the Brief Symptom Inventory (Derogatis, 1993). The Brief Symptom Inventory is a 53-item self-report scale that measures nine psychiatric symptom dimensions. The GSI is the aggregate severity measure with higher scores corresponding to more severe symptoms.

*Drug and alcohol problem severity* were assessed by the corresponding sections from the Addiction Severity Index (ASI; McLellan et al., 1992), which is a widely used omnibus measure with good psychometric properties. The Drug Composite Score (ASI-D) and Alcohol Composite Score (ASI-A) are based on reported use and perceived problem severity during the past 30 days. Higher scores indicate greater problem severity.

*Posttraumatic symptom severity* (PSS) was assessed by the Symptom Scale (Foa, Cashman, Jaycox, & Perry, 1997;

Foa, Riggs, Dancu, & Rothbaum, 1993). The PSS items represent the 17 cardinal symptoms of this anxiety disorder. Respondents rate how often each symptom has bothered them in the past month. The sum of the ratings provides an overall indicator of symptom severity.

For a more in-depth description of the study design and methodology, see [McHugo and colleagues \(2005\)](#).

## 2.2. Participants

Women enrolled in the study met a common set of eligibility criteria: age 18 or older, have both substance abuse and mental health disorders in the past 5 years (with at least one episode in the past 30 days), have had at least two prior treatment episodes, and have experienced physical or sexual abuse. At baseline, 2,729 eligible women completed the consent procedures and interviews. Of the 2,101 women who completed the interview at 6 months, 95 were interviewed more than 3 months late (i.e., more than 9 months after baseline). These women were dropped from the analysis. The final, analyzed sample size was 2,006. This number represents 95% of the 6-month achieved sample, and 74% of the original sample. Every woman included in the 6-month analysis completed her interview within 12 weeks of the data collection due date. The date of these interviews ranged from 6 weeks early to 12 weeks late.

At baseline, the 6-month sample ( $N=2,006$ ) resembled the original baseline sample ( $N=2,729$ ) in average age (36), ethnicity (49% Non-Hispanic White, 26% Non-Hispanic Black, and 17% Hispanic), education (50% with high school degree and 25% with GED), and marital status (38% married or partnered and 29% never married). However, there were some minor differences between the women who were included in the final 6-month sample and those who were either not interviewed at all or were interviewed too late to be included ( $N=763$ ). Further details on this sample are provided elsewhere ([McHugo et al., 2005](#)).

Without random assignment to conditions, it was possible that between-group differences would occur. Analysis of the outcome measures at baseline indicated that some differences did occur between the intervention and comparison programs by site (see [Table 1](#)). When significant between-group differences occurred at an individual site, standardized mean differences were used to estimate the effect size.

## 2.3. Program contrasts

### 2.3.1. Overview

All nine sites were similar in that the intervention services were expected to incorporate four key program elements: integration, comprehensiveness, trauma-informed treatment, and CSR involvement in treatment. At each site, these services were compared to similar programs, in the same region that represented “treatment as usual.” Despite consistency across sites in including the four program elements, the implementation and extent to which they were implemented varied across the intervention conditions. In addition, although agencies in the comparison condition were not provided any study support or funds to enhance their services there was no certainty that they had not incorporated these elements into their program since they were free to adopt any of the same program elements from other sources. For this reason, the collaboration and analysis of program data focused on identifying the differences between the intervention and the comparison conditions as they actually existed.

In order to assess the impact of the four key program characteristics on the study outcomes of the participants, it was necessary to assess: (1) whether the intervention conditions actually implemented the four key program elements; (2) whether these program attributes were present at a greater level in the intervention condition than in the usual-care comparison condition; and (3) whether these differences, if they did exist, had any impact on the outcomes experienced by study participants.

Table 1  
Differences in mean outcome measures between Intervention (I) and Comparison (C) programs at baseline ( $N = 2006$ )

| Site   | Mental health status (GSI) |             | Post-traumatic symptoms severity (PSS) |               | Alcohol problem severity (ASI-A) |             | Drug problem severity (ASI-D) |             |
|--------|----------------------------|-------------|--|---------------|----------------------------------|-------------|-------------------------------|-------------|
|        | I                          | C           | I                                      | C             | I                                | C           | I                             | C           |
| Site A | 1.06 (0.68)                | 1.03 (0.64) | 20.09 (9.94)                           | 18.29 (11.33) | 0.17 <sup>b</sup> (0.28)         | 0.28 (0.35) | 0.20 <sup>a</sup> (0.15)      | 0.24 (0.14) |
| Site B | 1.49 <sup>b</sup> (0.74)   | 1.21 (0.73) | 25.22 <sup>a</sup> (10.86)             | 22.06 (11.75) | 0.23 (0.35)                      | 0.16 (0.26) | 0.21 <sup>c</sup> (0.16)      | 0.12 (0.14) |
| Site C | 0.89 <sup>b</sup> (0.73)   | 1.28 (0.69) | 17.25 <sup>c</sup> (10.80)             | 24.44 (11.36) | 0.19 <sup>a</sup> (0.30)         | 0.32 (0.34) | 0.18 (0.14)                   | 0.18 (0.14) |
| Site D | 1.42 <sup>a</sup> (0.78)   | 1.71 (0.86) | 24.94 (12.00)                          | 26.15 (11.95) | 0.17 (0.28)                      | 0.13 (0.18) | 0.09 (0.13)                   | 0.10 (0.11) |
| Site E | 1.49 (0.85)                | 1.49 (0.79) | 25.64 (12.63)                          | 25.11 (12.60) | 0.12 <sup>c</sup> (0.21)         | 0.26 (0.33) | 0.11 (0.12)                   | 0.14 (0.13) |
| Site F | 1.46 (0.84)                | 1.44 (0.80) | 25.43 (12.34)                          | 24.28 (11.51) | 0.10 (0.23)                      | 0.08 (0.19) | 0.11 <sup>a</sup> (0.14)      | 0.15 (0.14) |
| Site G | 1.51 (0.76)                | 1.41 (0.71) | 26.31 (10.95)                          | 24.40 (11.29) | 0.29 (0.35)                      | 0.26 (0.35) | 0.17 <sup>c</sup> (0.18)      | 0.29 (0.16) |
| Site H | 1.55 (0.75)                | 1.64 (0.76) | 27.32 (11.37)                          | 27.8 (11.01)  | 0.15 (0.23)                      | 0.14 (0.24) | 0.11 <sup>a</sup> (0.11)      | 0.14 (0.13) |
| Site I | 1.13 (0.69)                | 1.00 (0.63) | 21.72 (12.72)                          | 19.06 (10.20) | 0.24 (0.33)                      | 0.30 (0.35) | 0.24 <sup>b</sup> (0.17)      | 0.16 (0.16) |

Note: Values are mean (standard deviation).

<sup>a</sup> ( $p < 0.05$ ).

<sup>b</sup> ( $p < 0.01$ ).

<sup>c</sup> ( $p < 0.001$ ).

To accomplish this, a set of eight program contrasts (Banks, McHugo, Williams, Drake, & Shinn, 2002) were organized and developed around the four key program elements and the required set of core services. Each contrast represented a key program element, a core service, or a combination of the two. One additional element of treatment, the extent of its focus on women's issues, was tested. This contrast was developed to capture what was thought to be an important treatment influence on women's improvement.

### 2.3.2. Development and results

A research workgroup with representatives from each site was convened to identify a means of measuring differences between the intervention and comparison conditions on these key program elements and services. For six of the eight contrasts, there was no single, direct way to measure these multidimensional model elements, so data were collected from three sources. First, observer ratings made by teams of site-level and research staff were compiled. These involved independent team members' ratings of various aspects of the key program elements that were offered in the intervention and comparison conditions. Second, ratings were used that had been scored during site visits conducted by the Coordinating Center. These ratings pertained to service delivery dimensions and the observed differences between the two conditions. Third, study participants' self-reports were examined to identify the nature of services and treatment received by these women during their first 3 months in the study.

These six program contrasts included:

1. *Resource Coordination and Advocacy*: services involving case management to advocate for the consumers and to identify and coordinate the various needed resources.
2. *Parenting*: services involving parenting skills training.
3. *Woman-Focused*: services involving strengths-based assessments and focus on empowering women and addressing needs specific to women's issues.
4. *CSR Involvement*: services involving CSRs in treatment (as peer-supports or facilitators of peer-run groups), as consumer advocates, and/or as advisors in the development or coordination of services.
5. *Trauma-Informed*: services involving trauma-specific group therapy where providers understand the influence of violence and abuse in women's lives.
6. *Integrated Treatment*: Various services involving concurrent treatment for mental health problems, substance abuse, and post-traumatic stress symptoms; Two other contrasts were derived from an analysis of the self-report data alone. These were:
  7. *Integrated Counseling*: Integration of the three treatment foci—mental health, substance abuse and trauma—within individual and group counseling.
  8. *Core Services Received*: Number of six core services received.

For the first six contrasts, a team of 14 raters (one or more persons from each site) reviewed the data assembled from the three possible sources. These data included the average value or level of each indicator (variable), for each condition (intervention or comparison), at each site. When appropriate, between-group differences were tested for significance. The 14 raters discussed the strengths and weaknesses of each indicator and the statistical results of the between-group comparisons. Based on this discussion, each of the fourteen raters independently coded each site as high (1) or low (0) contrast, depending on whether the majority of the tests indicated a significant difference between the two conditions.

Fourteen scores of '0' or '1' resulted for each program element. These fourteen scores were then averaged across raters. Average scores above 0.5 were coded as high contrast and average scores 0.5 or below were coded as low contrast. For each contrast, the level of inter-rater agreement was calculated using the kappa statistic.

Contrasts 1, 2, and 3 showed the strongest agreement among the raters ( $\kappa > .69$ ). Contrasts 4 and 5 were used despite weak kappas (.44 and .43) because of their salience to the intervention strategy employed in the WCDVS. Contrast 6, also thought to be a salient indicator, showed very weak inter-rater reliability ( $\kappa = .37$ ). This contrast was broadly defined as treatment integration (inclusion of the three key treatment foci) across several services. Discrepancy among the raters arose from the complexity of the rating task and conflicting information associated with one or more of the indicators assembled. The decision was made to test the original contrast, but to also test a more narrowly defined and potentially more reliable indicator using data obtained from client interviews (Contrast 7-Integrated Counseling).

Integrated Counseling was measured by summing the number of treatment foci (three possible: mental health, trauma and substance abuse) that women reported receiving in their individual and group counseling sessions at the 3-month follow-up interview. This provided a measure of client-level integration of these treatment issues—within or across these counseling modalities. Sites were designated high contrast if a significant difference existed between conditions. For the "high contrast" integrated counseling group, women in the intervention reported receiving an average of 2.04 to 2.68 of the key three foci, whereas women in the comparison agencies (same sites) reported an average of 1.68 to 2.23 treatment foci.

The final contrast, Core Services Received (CS), was also empirically defined. A value was calculated for each woman, and this value represented the number of six possible core services that she reported receiving at the 3-month follow-up interview.<sup>1</sup> Sites with a significant

<sup>1</sup> As discussed earlier, in order to meet the federal guideline that funded programs be comprehensive, each program was required to provide a series of core services. For three of these—outreach, assessment, and crisis intervention—there was no relevant data available from the interviews.

difference between the average score for women in the intervention program vs. women in the comparison program were classified as high contrast. Intervention women in the high contrast sites reported an average of 3.91 to 5.67 of the six measured core services, and the comparison women reported an average of 3.05 to 4.95.

For each of the contrasts, classifying a site as ‘high’ depended solely on whether the intervention provided significantly more of the program element than its comparison condition. If the intervention program in Site A provided significantly more Integrated Counseling than the comparison program in Site A on the Integrated Counseling contrast, it would be designated high contrast on Integrated Counseling. If that same intervention project provided a similar or lesser amount of Core Services, as compared to its usual care condition, then the site was designated low contrast on Core Services. For each contrast, the high contrast sites were grouped together and compared (by ANOVA) to the low contrast sites. If the high contrast sites, as a group, demonstrated significantly better outcomes, the

Table 2  
Six-month outcome effects (N = 2006)

| Outcome                                | Effect Size | p-value |
|--|-------------|---------|
| Alcohol problem severity (ASI-A)       | -.003       | .95     |
| Drug problem severity (ASI-D)          | .109        | .02     |
| Mental health status (GSI)             | .085        | .06     |
| Post-traumatic symptoms severity (PSS) | .110        | .02     |

element measured by the contrast could potentially be one explanation for the better outcomes.

2.4. Data analysis

Prospective meta-analysis was used to analyze the data. Banks et al. (2002) have described this approach in detail. This approach calculates individual and overall weighted effect sizes and models the heterogeneity of site outcomes using program-level contrast variables.

The method for estimating the site effect sizes depended on between-condition differences on the baseline

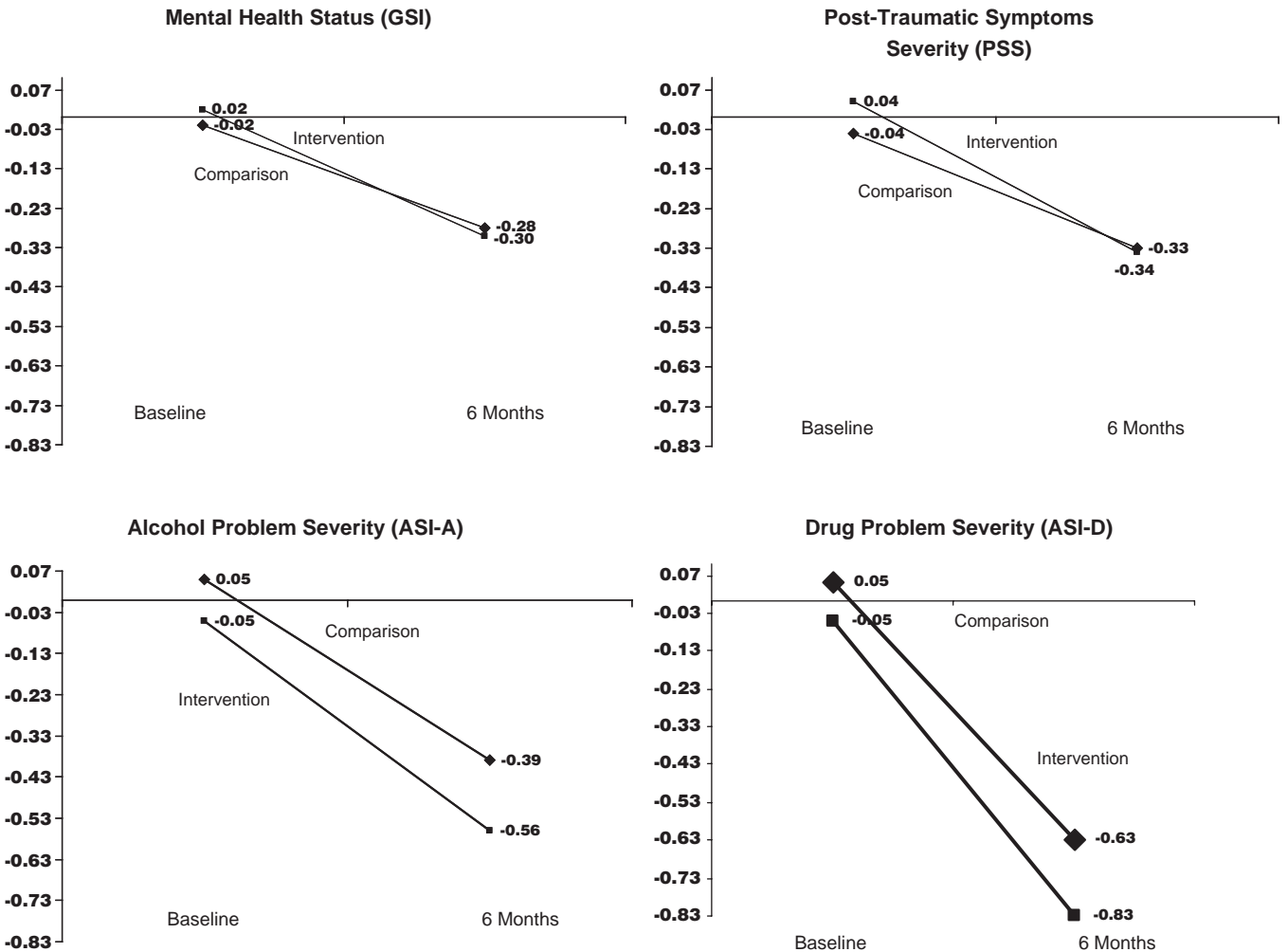


Fig. 1. Standardized change scores on outcome measures at baseline and 6 months for intervention and comparison conditions.

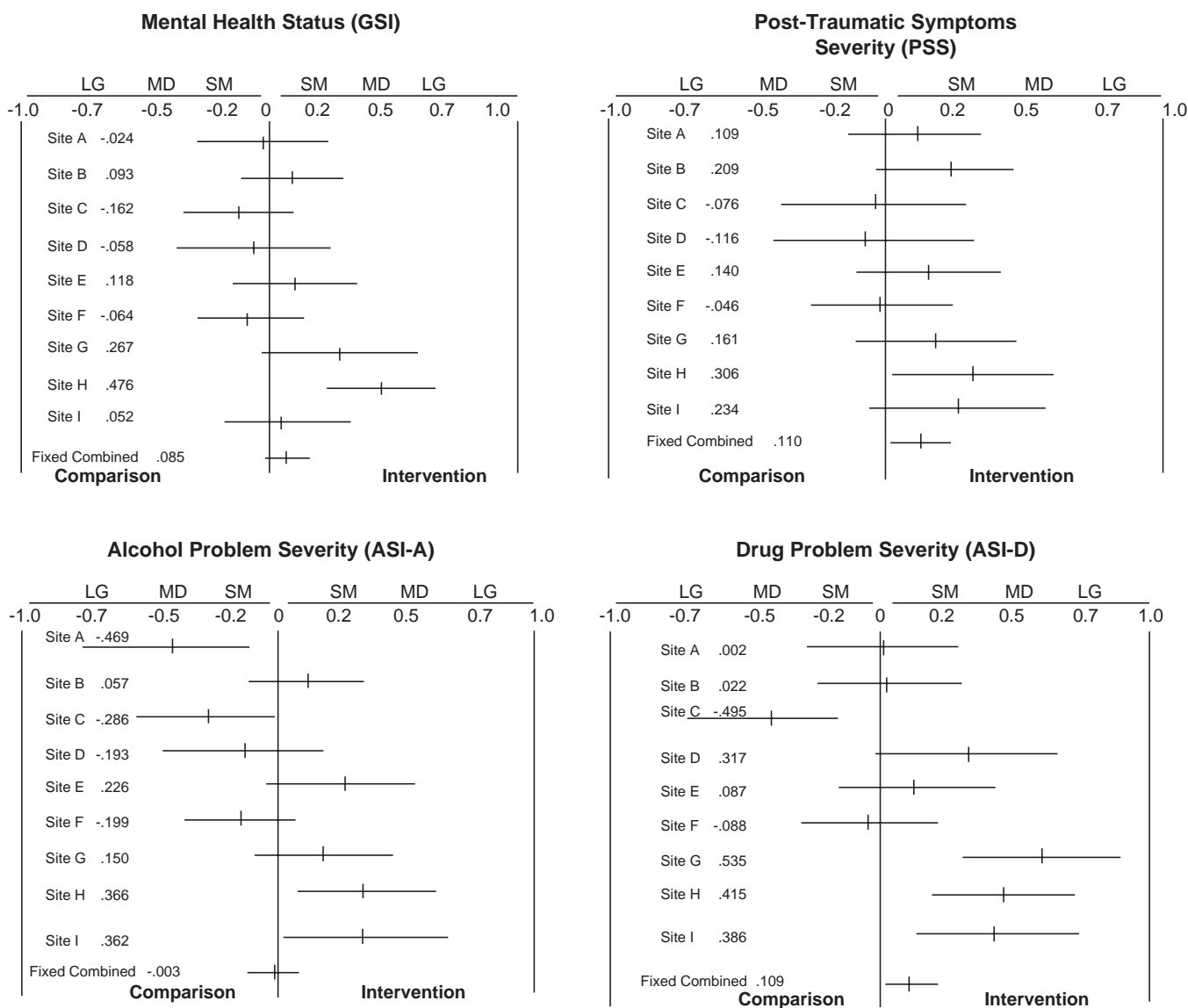


Fig. 2. Variations in effect sizes across study sites.

scores of the outcome measures. Given the large sample size, a significance test with corresponding *p*-value would be a weak criterion to assess between-condition differences, so we used standardized mean differences (SMD) instead. SMDs were considered significant if the mean difference was equal to or greater than an absolute value of 0.1. For sites showing baseline differences on outcome variables, standardized mean change scores were used to estimate the effect size. Otherwise, ANCOVA was used (Banks et al., 2002). The Comprehensive Meta-Analysis Program (Borenstein & Rothstein, 1999) was used to calculate all effect sizes, *Q*-values, and to test within and between group differences.

When an outcome demonstrated a significant overall effect size ( $p \leq 0.05$ ), the heterogeneity of site outcome was measured (*Q* statistic). A *p*-value of 0.05 was used to determine significance. When both of these measures

achieved significance, the program contrasts were used to model the variance. ‘High contrast’ site effects were grouped and compared to the ‘low contrast’ group effect size. If the ‘high contrast’ group achieved a significantly larger effect size than the ‘low contrast’ group, then it was inferred that the program element was related to the success of those high contrast sites.

### 3. Results

#### 3.1. Women’s improvement over time

Regardless of being in the intervention or comparison conditions, women’s outcomes improved between baseline and 6 months in every group and on every outcome measure. For both groups, average mean differences for

all four measures of mental health status, post-traumatic symptom severity, alcohol problem severity and drug problem severity were positive (signifying less symptoms and more improvement) at 6 months (table available from corresponding author). Thus, women in the WCDVS study experienced improvement on all outcomes regardless of whether they were in the intervention or comparison conditions between baseline and 6 months.

While all women improved at 6 months, the crucial research question for the WCDVS is whether those in the intervention condition demonstrated better outcomes than those in the comparison condition. Importantly, women in the intervention conditions appeared to improve more than those in usual care. As can be seen in Table 2, two outcomes showed significantly more change in the intervention than in the usual care groups. ASI-Drug showed an effect size of 0.109 ( $p=0.02$ ), and the post-traumatic symptom outcome (PSS) had an overall weighted effect of 0.110 ( $p=0.02$ ). While not quite statistically significant ( $p=0.06$ ), the mental health (GSI) outcome demonstrated an effect that was similar in size to the trauma and drug use outcomes (.085). These changes are shown graphically in Fig. 1. Outcome values are converted into standard scores to facilitate comparisons between measures that are calculated on different scales.

### 3.2. Variability in women's improvement by site

The second research question asks about variability in outcome effects across sites. The I-bar charts presented in Fig. 2 shows that effect sizes varied across the nine sites. Between-site heterogeneity was calculated according to the adjusted Hedge's formula. Significant heterogeneity occurred in three of the four outcomes: ASI-Alcohol ( $Q=38.7$ ,  $p<0.001$ ), ASI-Drug ( $Q=40.0$ ,  $p<0.001$ ), and GSI ( $Q=17.7$ ,  $p=0.02$ ). The trauma outcome (PSS) was not heterogeneous ( $Q=8.7$ ,  $p=0.37$ ). The heterogeneity across sites sometimes included intervention conditions which produced less positive outcomes than the usual care conditions. In two instances, the comparison condition produced results comparable to or even better than the intervention condition at some other sites. The effect size represents a comparison of the intervention and usual care conditions and as a result can be negative for a variety of

Table 3  
Integrated counseling: Group effect sizes

| Contrast group    | Mental health severity (GSI) | Post-traumatic symptoms severity (PSS) | Alcohol problem severity (ASI-A) | Drug problem severity (ASI-D) |
|-------------------|------------------------------|--|----------------------------------|-------------------------------|
| High contrast     | .207                         | .169                                   | .213                             | .347                          |
| Low contrast      | -.033                        | .053                                   | -.213                            | -.121                         |
| Effect difference | .240 <sup>a</sup>            | .116                                   | .426 <sup>b</sup>                | .468 <sup>b</sup>             |

<sup>a</sup> ( $p<.001$ ).

<sup>b</sup> ( $p<0.001$ ).

Table 4  
Core services received: group effect sizes

| Contrast group    | Mental health status (GSI) | Post-traumatic symptoms severity (PSS) | Alcohol problem severity (ASI-A) | Drug problem severity (ASI-D) |
|-------------------|----------------------------|--|----------------------------------|-------------------------------|
| High contrast     | -.027                      | .039                                   | -.120                            | .004                          |
| Low contrast      | .283                       | .235                                   | .200                             | .293                          |
| Effect difference | -.310 <sup>b</sup>         | -.196 <sup>a</sup>                     | -.320 <sup>a</sup>               | -.289 <sup>b</sup>            |

<sup>a</sup> ( $p<0.05$ ).

<sup>b</sup> ( $p<0.01$ ).

<sup>c</sup> ( $p<0.001$ ).

reasons, related to either the intervention or the comparison conditions or both.

### 3.3. Program-level factors contributing to the positive intervention effects

The third research question asks about key program characteristics that might explain site-to-site variability in outcome effects. Although only three of the outcomes were significantly heterogeneous, the program contrasts were applied to all outcomes to explore their ability to explain site outcome variance. Only two of the eight contrasts (Integrated Counseling and Core Services) demonstrated a pattern of significance across the four outcomes.

As shown in Table 3, the Integrated Counseling contrast was significant in the expected direction. For example, the effect size for high contrast cluster of sites is 0.347 on the drug problem severity outcome as opposed to  $-0.121$  for the low contrast cluster, an effect size difference of 0.468.

Whereas the Integrated Counseling contrast produced results that favored the high contrast cluster, the Core Services contrast produced the opposite effect—the low contrast cluster scored higher than the high contrast cluster (See Table 4). The results were counterintuitive.<sup>2</sup>

In an effort to better understand the relationship between the Integrated Counseling and Core Services Received contrasts, sites were placed into combined categories according to their status (high or low) on each contrast (See Table 5). The ASI-Drug outcome was used for this test because it was the only outcome that demonstrated a significant overall effect and significant site heterogeneity. After grouping the sites into the appropriate categories,

<sup>2</sup> Sensitivity analyses were performed on both contrasts for the one outcome that was both significant and heterogeneous—ASI-Drug. Sensitivity analyses are recommended by Banks et al. (2002) and involve dropping the most extreme or influential site effect sizes in order to assess whether the significant between-group difference is due to one site or to the group as a whole. For both contrasts, IC and CS, the between-group (high vs. low contrast) differences remained significant even after removing the extreme effect sizes from each group. This suggests that the between-group difference does not depend on any one site but represents a true group difference in both contrasts.

Table 5  
Combination of program elements and associated site effect sizes for ASI-Drug outcome

| Effect sizes                          | Combination of Program Elements (Contrasts) |          |          |           |
|---------------------------------------|---|----------|----------|-----------|
|                                       | Integrated Counseling / Services Used       |          |          |           |
|                                       | Low/Low                                     | Low/High | High/Low | High/High |
| Effect sizes                          | .022  | .002     | .535     | .317      |
| of individual sites in group          |   | -.495    | .415     | .087      |
| Weighted average of site effect sizes | .022  | -.174    | .458     | .252      |

weighted average effect sizes were calculated. The sites that were high on Integrated Counseling and low on Core Services produced the greatest effect size. The sites high on both elements also produced an overall effect size favoring the intervention; however, this effect size was not as large. In the two conditions where Integrated Counseling was low, the effect sizes were lower whether the Core Services contrast was low or high. Overall, the results suggest that Integrated Counseling may be the key ingredient in this study's effects. Being high on Core Services appears to contribute to beneficial intervention effects vis-à-vis usual care but only when paired with greater amounts of Integrated Counseling.

#### 4. Discussion

This study found that at the 6-month follow-up period, outcomes for women served in the intervention and usual care conditions improved. This finding is consistent across all four primary outcome measures of mental health status, post-traumatic symptoms, drug use, and alcohol use and across all sites. Regardless of whether the women were part of the intervention or comparison conditions, their outcomes improved over the 6-month period.

The primary reason for this study was to assess the extent to which participation in the intervention conditions resulted in better outcomes as compared to care-as-usual. Did women in the intervention conditions, which were designed to provide comprehensive, integrated, trauma-informed and CSR-involved approaches, show greater improvement than women who received services in the comparison conditions? The data presented in this paper suggest that they did.

On two of the four measures (post-traumatic symptoms and drug use severity), women in the intervention programs showed significantly greater improvement than those in usual care and on a third factor, mental health status, while not significant, intervention sites again did better. While the effect sizes are small, these findings are still somewhat surprising given the existing body of literature that has examined the impact of large scale, comprehensive, integrated service models and found little to suggest improved outcomes for individuals involved in these interventions.

The effect of the intervention is not consistent across sites. In fact, there is considerable variation across the nine participating sites. Thus, although as a whole the intervention sites displayed better outcomes, there was a significant amount of heterogeneity among sites. In some, women in the intervention conditions clearly did better than those in the comparison groups but in other cases there was little difference and in some cases women in usual care actually did better.

This variability is consistent with the manner in which the intervention services were developed. In accordance with the federal guidelines, all intervention projects were required to incorporate four key characteristics into their programs (integrated services, comprehensive range of services, trauma-informed service, and CSR involvement). However, the guidelines allowed them great flexibility in how these were actually implemented. SAMHSA did not stipulate which model or strategy should be used to attain these characteristics. As a result, programs used a variety of different approaches. Although conceptually similar in striving to attain the required key program characteristics, sites used a number of different services and approaches. For example, in some sites services were provided primarily on an outpatient basis while in others they occurred in a residential setting; some interventions were provided in a setting that was primarily a substance abuse agency, others in a mental health agency environment; and while all sites provided trauma-specific services, four different models were used (McHugo et al., 2005; Veysey & Clarke, 2004).

Similarly, comparison conditions were not uniform. General guidelines provided by SAMHSA helped to ensure that comparisons were selected in the same geographical area and served similar women. Nonetheless, the characteristics of the comparison conditions varied widely, as did the elements of "usual care."

The final research question addressed in this article was whether the cross-site variations in outcomes can be explained by differences in key program characteristics. In order to assess whether program variations between the intervention and comparison conditions accounted for any of the observed outcome changes, a series of program contrasts were developed and measured. These eight contrasts attempted to capture programmatic differences between the intervention and comparison conditions at each site on a number of dimensions related to the intervention.

One program contrast in particular, Integrated Counseling, was found to be significantly related to three of the four outcomes measured across the sites. This program-level variable measured how many of the three key treatment foci—mental health, post-traumatic symptoms and substance abuse—women reported receiving. Sites in which there was a high contrast between the intervention and comparison conditions showed a larger effect size on four outcome variables. These findings suggest that the effects of the interventions are conditioned on providing integrated counseling to women who have co-occurring disorders and

violence histories. Other program contrasts such as resource coordination and advocacy or more general concepts such as integrated treatment were found to vary between intervention and comparison conditions at some sites, but these dimensions were *not* associated with the cross-site heterogeneity in study outcomes. A second program contrast that was significantly related to improvement in outcomes was the number of core services received in the treatment program. Considering the two significant program contrasts together, sites in which women received more services did not produce better results than their comparison counterparts unless they also received more integrated counseling services. This finding appears to underscore the importance of integrating of mental health, substance abuse and trauma treatment for women with co-occurring disorders.

Why is integrated counseling so important? Our earlier review of research findings suggests that this may be an example of an effective “bottom up” strategy for integrating services that focus on the client level (Rosenheck et al., 2001). While it adds trauma services to the mix, this finding also appears consistent with emerging research that suggests the positive impact of client-level integrated services for individuals with co-occurring mental and substance use disorders (Drake et al., 2001).

There are a number of limitations to this study that caution against drawing firm conclusions. The study relied on a quasi-experimental, non-random design that occurred within the context of a large, multi-site trial. Intervention conditions were consistent in their key program characteristics, but varied greatly in actual programs and services implemented to meet them. Comparison conditions were also not uniform and not static. While projects in the comparison condition were not provided the federal funds to enhance services based on the intervention model being tested, anecdotal information suggests that some of these sites began incorporating elements of the intervention over the life of the study. For example, independent of the WCDVS, there was a growing awareness over the years of the project of the role played by interpersonal violence in the lives of women, thus leading some projects in the comparison condition to introduce trauma services. Nonetheless, the use of program contrasts in this study, which focused on assessing actual rather than presumed differences, provided some insulation from this possible drawback.

The use of a prospective meta-analytic approach and program contrast is well suited to a study involving multiple sites with varied interventions. Yet these methods are relatively new (Banks et al., 2002) and procedures for developing and measuring program contrasts can and should be improved over the strategies employed in the WCDVS. Nonetheless, the results presented here support the usefulness of this approach. Two of the eight program contrasts explained variability in outcomes across sites and expand our understanding of the conditions under which the interventions work. Yet these two contrasts were based on client self-reports of services received, whereas the

other six contrasts were based on ratings completed by informants only from the intervention sites, who were not blinded to study hypotheses or study conditions, thus raising questions regarding the impact of the measurement itself. Finally, the findings reported here are based on 6-month outcome data. Whether these initial findings hold up over the full 12-month follow-up period will be addressed in subsequent papers.

Despite these limitations, the findings reported here suggest that (1) outcomes for women with multiple and interrelated needs can be improved by comprehensive, integrated, trauma-informed, and CSR-involved services and that (2) these effects are much more pronounced when services emphasize integrated counseling. Not addressed here is the role that person-level characteristics of the women who participated may play in understanding these effects. Because randomization was not employed in this study, the presence of selection artifacts would be a plausible alternative explanation for the positive results reported here. These artifacts occur when sites enroll women in their intervention condition with different problem constellations and experiences related to outcomes than those women enrolled in the comparison condition. This issue is addressed in a person-level analysis of WCDVS outcomes at 6 months reported in a companion article (Morrissey et al., 2005).

## Acknowledgments

This study was funded under Guidance for Applicants (GFA) No. TI 00-003 entitled *Cooperative Agreement to Study Women with Alcohol, Drug Abuse and Mental Health (ADM) Disorders who have Histories of Violence: Phase II* from the Department of Health and Human Services, Public Health Service, Substance Abuse and Mental Health Services Administration’s three centers: Center for Substance Abuse Treatment, Center for Mental Health Services and Center for Substance Abuse Prevention (March 2000). The assistance of project staff at the following participating sites (listed in alphabetical order by state) is gratefully acknowledged: Los Angeles, California: PROTOTYPES Systems Change Center, Vivian Brown, Principal Investigator, Stockton, California: Allies: An Integrated System of Care, Jennie Heckman, Principal Investigator, Thornton, Colorado: Arapahoe House – New Directions for Families, Nancy VanDeMark, Principal Investigator, Washington, DC: District of Columbia Trauma Collaboration Study, Roger Fallot, Principal Investigator, Avon Park, Florida: Triad Women’s Project, Margo Fleisher-Bond, Co-Principal Investigator, Colleen Clark, Co-Principal Investigator, Boston, Massachusetts: Boston Consortium of Services for Families in Recovery, Hortensia Amaro, Principal Investigator, Cambridge, Massachusetts: Women Embracing Life and Living (WELL) Project, Norma Finkelstein, Principal Investigator, Greenfield, Massachusetts: Franklin County

Women's Research Project, Rene Andersen, Principal Investigator, New York, New York: Portal Project, Sharon Cadiz, Principal Investigator. The Coordinating Center was operated by Policy Research Associates (PRA), located in Delmar, New York, in coordination with the National Center on Family Homelessness of Newton, Massachusetts and the Cecil G. Sheps Center for Health Services Research at the University of North Carolina (UNC) in Chapel Hill, North Carolina.

The interpretations and conclusions contained in this article do not necessarily represent the position of the WCDVS Coordinating Center, participating study sites, participating Consumer/Survivor/Recovering persons, or the Substance Abuse and Mental Health Services Administration and its three centers (Center for Substance Abuse Treatment, Center for Mental Health Services, Center for Substance Abuse Prevention).

## References

- Alexander, M. J. (1996). Women with co-occurring addictive and mental disorders: An emerging profile of vulnerability. *American Journal of Orthopsychiatry*, 66, 61–70.
- Banks, S., McHugo, G. J., Williams, V., Drake, R. E., & Shinn, M. (2002). A prospective meta-analytic approach in a multisite study of homelessness prevention. *New Directions for Evaluation*, 94, 45–58.
- Behr, L. (1988). An integrated state system of services for seriously disturbed children. In *Chronic mental illness in children and adolescents*. Washington, DC: American Psychiatric Press.
- Bickman, L., Guthrie, P., Foster, E. M., Lambert, E. W., Summerfelt, W. T., Breda, C., & Heflinger, C. A. (1995). *Managed care in mental health: The Fort Bragg experiment*. New York: Plenum.
- Bickman, L., Summerfelt, W., & Noser, K. (1997). Comparative outcomes of emotionally disturbed children and adolescents in a system of services and usual care. *Psychiatric Services*, 48, 1543–1548.
- Borenstein, M., & Rothstein, H. (1999). *Comprehensive meta-analysis: A computer program for research synthesis*. Englewood, NJ: Biostat.
- Browne, A., & Bassuk, S. (1997). Intimate violence in the lives of homeless and poor housed women: Prevalence and patterns in an ethnically diverse sample. *American Journal of Orthopsychiatry*, 67, 261–278.
- Campbell, J. (1997). How consumers/survivors are evaluating the quality of psychiatric care. *Evaluation Review*, 21, 357–363.
- Commonwealth Fund. (1997). *Facts on abuse and violence: The Commonwealth Survey of the Health of Adolescent Girls*. New York, NY: Louis and Harris Associates, Inc.
- Deegan, P. E. (1995). *Coping with voices: Self help strategies for people who hear voices that are distressing*. Lawrence, MA: The National Empowerment Center.
- Dennis, D. L., Steadman, H. J., & Coccozza, J. J. (2000). The Impact of Federal Systems Integration Initiatives on services for mentally ill homeless persons. *Mental Health Services Research*, 2, 164–174.
- Derogatis, L. R. (1993). *Brief system inventory (BSI): Administration, scoring and procedures manual*, (4th ed.). Minneapolis, MN: National Computer Systems.
- Drake, R. E., Goldman, H. H., Leff, H. S., Lehman, A. F., Dixon, L., Mueser, K. T., & Torrey, W. C. (2001). Implementing evidence-based practices in routine mental health service settings. *Psychiatric Services*, 52, 179–182.
- Fallot, R., & Harris, M. (2004). Pilot studies of the Trauma Recovery and Empowerment Model. (unpublished manuscript).
- Fisher, D. B. (1994a). Health care reform based on an empowerment model of recovery by people with psychiatric disabilities. *Hospital and Community Psychiatry*, 45, 913–915.
- Fisher, D. B. (1994b). A new vision of healing as constructed by people with psychiatric disabilities working as mental health providers. *Psychosocial Rehabilitation Journal*, 17.
- Foa, E. B., Cashman, L., Jaycox, L., & Perry, K. (1997). The validation of a self-report measure of posttraumatic stress disorder: The posttraumatic diagnostic scale. *Psychological Assessment*, 9, 445–451.
- Foa, E. B., Riggs, D. S., Dancu, C. V., & Rothbaum, B. O. (1993). Reliability and validity of a brief instrument for assessing posttraumatic stress disorder. *Journal of Traumatic Stress*, 6, 459–473.
- Goldman, H. H., Morrissey, J. P., & Ridgely, M. S. (1994). Evaluating the Robert Wood Johnson Foundation Program on Chronic Mental Illness. *The Milbank Quarterly*, 72, 37–47.
- Harris, M. (1994). Modifications in service delivery and clinical treatment for women diagnosed with severe mental illness who are also survivors of trauma. Special Issue: Women's Mental Health. *Journal of Mental Health Administration*, 21, 397–406.
- Kahn, A. J., & Kamerman, S. B. (1992). *Integrating services integration: An overview of initiatives, issues, and possibilities*. New York: National Center for Children in Poverty.
- Kruse, S. (1999). Triad Women's Project: Triad pilot groups: Results of statistical analysis. Louis de la Parte Florida Mental Health Institute, University of South Florida.
- McCabe, S., & Unzicker, R. E. (1995). Changing roles of consumer/survivors in mature mental health systems. *New Directions for Mental Health Services*, 66, 61–73.
- McHugo, G. J., Kammerer, N., Jackson, E. W., Markoff, L. S., Gatz, M., Larson, M. J., Mazelis, R., & Hennigan, K. (2005). Women Co-occurring Disorders, and Violence Study: Evaluation design and study population. *Journal of Substance Abuse Treatment*, 28, 91–107.
- McLellan, A. T., Kushner, H., Metzger, D., Peters, R., Smith, I., Grissom, G., Pettinati, H., & Argeriou, M. (1992). The fifth edition of the Addiction Severity Index. *Journal of Substance Abuse Treatment*, 9, 199–213.
- Morrissey, J. P., Ellis, A. R., Gatz, M., Amaro, H., Glover, B., Savage, A., Finkelstein, N., Mazelis, R., Brown, V., Jackson, E. W., & Banks, S. (2005). Outcomes for women with co-occurring disorders and trauma: Person-level effects. *Journal of Substance Abuse Treatment*, 28, 121–133.
- Najavits, L., Weiss, R., Shaw, S., & Muenz, L. (1998). Seeking safety: Outcome of a new cognitive-behavioral psychotherapy for women with post-traumatic stress disorder and substance dependence. *Journal of Traumatic Stress*, 11, 437–456.
- NASMHPD. (1998). *Position Statement on Consumer Contributions to Mental Health Service Delivery Systems*. Paper presented at the National Association of State Mental Health Program Directors, Alexandria, VA.
- Policy Research Associates. (Summer 2000). *I. The tapestry: Weaving a fabric for change*. Delmar, NY: Policy Research Associates Inc.
- Prescott, L. (2001). *Consumer/survivor/recovering women: A guide for partnerships in collaboration*. Delmar, NY: Policy Research Associates, Inc.
- Ridgely, S., Goldman, H., & Willenbring, M. (1990). Barriers to the care of persons with dual diagnoses: Organizational and financing issues. *Schizophrenia Bulletin*, 16, 123–132.
- Rosenheck, R., Morrissey, J. P., Lam, J., Calloway, M., Stolar, M., Johnsen, M., Randolph, F., Blasinsky, M., & Goldman, H. H. (2001). Service delivery and community: Social capital, service systems integration, and outcomes among homeless persons with severe mental illness. *Health Services Research*, 36, 691–709.
- Substance Abuse and Mental Health Services Administration. (2002). Report on co-occurring disorders recommends integrated treatment. *SAMHSA News*, 10.
- Veysey, B. M., & Clarke, C. (2004). Introductions: Responding to physical and sexual abuse in women with alcohol, other drug, and mental disorders. *Alcoholism Treatment Quarterly*, 22, 1–18.
- Yessian, M. R. (1995). Learning from experience: Integrating the human services. *Public Welfare, Summer*, 34–42.