

**Substance Abuse Treatment Outcomes: A Multi-Site Study of Male and
Female Prison Programs**

Bernadette Pelissier
Research Department
Federal Correctional Institution
PO Box 1000
Butner, NC 27509-1000
919-575-4541, Ext. 3656
bpelissier@bop.gov

Mark Motivans
Bureau of Justice Statistics
810 7th Street, N.W.
Washington D.C. 20531
202-514-4272
wMMotivan.WTGATE@worldtalk.bop.gov

Jennifer L. Rounds-Bryant
839 Woodgrove Street
Durham, NC 27703
drjem@ipass.net

Abstract

The present study examined whether there were program differences with respect to post-release outcomes in 20 federal in-prison substance abuse programs which used a cognitive-behavioral treatment approach. Recidivism and post-release drug use were examined for a sample of 1,343 individuals – 1,065 men and 278 women. Discrete time proportional hazards regression models showed that, after controlling for individual characteristics, no differences were detected among the 16 programs for men. In contrast, one of the four female programs had significantly higher drug use rates and one had significantly lower recidivism rates. Our results suggest that implementation of a treatment approach which has been shown to be effective – cognitive-behavioral treatment – can result in comparable outcomes across programs, despite differences in program implementation. However, specific types of variation in program implementation may affect outcomes.

Keywords: drug treatment, multi-site, prison

Substance Abuse Treatment Outcomes: A Multi-Site Study of Male and Female Prison Programs

Introduction

Problematic drug use has been identified as one of the nation's most serious health problems because of its negative effects on public safety, families and the economy (Horgan, Skwara, & Strickler, 2001). Yet this major public health problem has no clearly identified consistent treatment approach. This is in contrast to the way in which many public health problems (e.g., communicable diseases) are addressed: wide dissemination of empirically-proven, efficient, and cost-effective treatment (Campbell & Bryant, 2001; Koplan & Thacker, 2001; Malloy & Marr, 2001). Recently, both the drug treatment field and the field of corrections have made calls to follow the trend of evidence-based medicine, which espouses the development of a body of empirically based evidence on treatment effectiveness to disseminate to policy makers and program staff (MacKenzie, 2000; Prendergast & Podus, 2000). This goal has recently been exemplified by the National Institute on Drug Abuse's (NIDA) clinical trials network (Mathias, 2001).

Drug treatment delivered in the community and in prisons has generally been shown to be effective at reducing drug use and related negative behavior (Gerstein et al., 1997; Hubbard, Craddock, Flynn, Anderson, & Etheridge, 1997; Martin, Butzin, Saum, & Inciardi, 1999; Rhodes et al., 2001; Schildhaus et al., 1998; Simpson, Wexler, & Inciardi, 1999; Wexler, Melnick, Lowe, & Peters, 1999). Studies assessing the effectiveness of drug treatment include a range of treatment modalities such as methadone maintenance, outpatient drug-free treatment and therapeutic communities. In addition, the types of services delivered include various approaches such as the 12-step model and cognitive-behavioral models such as relapse prevention.

Meta-analyses have shown cognitive-behavioral treatment (including treatment provided specifically to drug users) to be effective among offender populations (Lipsey, Chapman, & Landenberger, 2001). An assessment of relapse prevention, a specific type of cognitive-behavioral treatment, has shown it to be effective for various types of substance users and in both inpatient

and outpatient settings (Irvin, Bowers, Dunn, & Wang, 2000). A meta-analysis of drug abuse treatment, primarily for treatment provided in the community, found that treatment was effective and that treatment modality was not related to effect size (Prendergast, Podus, Chang, & Urada, 2002). However, meta-analyses of drug treatment programs provided to criminal justice populations show some inconsistency in results regarding the types of drug treatment found to be effective, in part, because studies are coded in different manners. Several studies have found therapeutic communities to be effective (Mitchell et al., 2002; Pearson & Lipton, 1999). Another meta-analysis of criminal justice drug treatment programs found that therapeutic communities with aftercare, most of which have cognitive-behavioral components, are effective (Aos, Phipps, Barnoski, & Lee, 2001). One meta-analysis did not find drug-focused group counseling to be effective (Pearson & Lipton, 1999) whereas another found group counseling to be effective (Mitchell et al., 2002).

As for 12-step programs, most of the time too few studies were available to evaluate its effectiveness. It is important to note that in all these studies, the categories of treatment type are not mutually exclusive. Programs can include more than one treatment type and some programs had up to four different types of treatment (Pearson & Lipton, 1999). For example, a therapeutic community can contain a cognitive-behavioral approach but also incorporate a 12-step program approach.

Despite the literature which has shown cognitive-behavioral approaches to be effective in both prison- and community-based settings, to date there has been relatively little research which examines how variation in the implementation of drug treatment programs can affect treatment outcomes. The limited studies examine different components of program implementation. Some studies focus primarily upon program modality (Hser et al., 1998) whereas others examine both program modality as well as specific treatment components such as treatment revenue and client-staff load (Schildhaus et al, 2000; Schildhaus et al., 2002). Yet other studies examine post-treatment outcomes across different programs within the same modality and assess the role of

factors such as prominence of treatment plans, ancillary and support services, having a designated case manager, and having staff training specialists (Orwin & Ellis, 2000a & 2000b). All of these studies found an association between modality or other treatment components (e.g., having staff training specialists, the provision of ancillary services) and outcomes. However, one of these studies found that the specific treatment variables included in their analysis added little explanatory power (Schildhaus et al., 2000). That program characteristics other than modality are associated with post-treatment outcomes is consistent with other research which has demonstrated an association between program characteristics other than treatment modality (e.g., staff education, sources of funding) and treatment retention (Grella, Joshi, & Hser, 2000; Hser, Maglione, Joshi, & Chao, 1998; Orwin & Williams, 1999, Pelissier, Camp, & Motivans, 2003).

Although there is little research on program differences in treatment outcomes, it has been acknowledged that program-specific diversity is the state of the art in community-based drug treatment programs and that this diversity is not limited to program modality. Rather, the diversity is expressed by variation in treatment content, treatment philosophy, type of services, treatment duration and intensity, staff education and experience, client populations served, and sources of funding (Anglin and Hser 1990; Broome et al., 1999; Etheridge et al., 1997; Grella et al., 2000; Hser et al., 2001; Hser et al., 1998; Hubbard et al., 1989; Melnick, DeLeon, Hiller, & Knight, 2000; Orwin and Ellis, 2000a).

Even though program diversity appears to be the norm, empirical research has not yet adequately addressed the question of whether and how program variation in factors other than modality or treatment approach (e.g., relapse prevention, 12-step model) are related to treatment outcomes. Perhaps as a result of this, researchers have noted that neither program nor client characteristics adequately explain the considerable variation in treatment outcomes across and within treatment modalities (Orwin & Ellis, 2000a; Prendergast & Podus, 2000). The difficulty of assessing the effects of program characteristics is that even in similar types of programs there can be considerable variation in the inclusion and emphasis upon various program elements (Mears &

Kelly, 2002).

Purpose of Study

With the exception of one study (Rhodes et al., 2001), most evaluations of prison-based substance abuse treatment programs are not multi-site in nature. Therefore, there is a dearth of information on program variation within criminal justice settings. It cannot be assumed that variations in program implementation will manifest themselves similarly in community and prison settings because of the constraints provided by the controlled setting of prisons. Because the federal Bureau of Prisons (BOP) attempted to standardize program content, this study provides the opportunity to assess: 1) whether there is evidence of program variation within a prison system with multiple programs based upon the same treatment approach and, 2) whether programs implemented across a number of sites within such a context will show differences in outcomes after controlling for individual level characteristics. If program variations in outcomes are found, this study will attempt to identify program factors associated with increased or decreased favorable outcomes. It is recognized at the outset that such an attempt is made all the more difficult by the absence of theory available to guide the selection of factors and because of the possible complexity in identifying unique program factors.

This study follows up on the multi-site drug treatment outcome evaluation which found positive effects of prison-based treatment (Pelissier, Camp, Gaes, Saylor, & Rhodes, 2003; Pelissier et al., 2001; Rhodes et al., 2001). The TRIAD evaluation found that the likelihood of rearrest or revocation within 3 years after release among individuals released to supervision was 44% among male treated subjects as compared with 52% for male untreated subjects. The comparable figures for women were 24% and 30%, respectively. Furthermore the likelihood of drug use among men during the 3-year post-release period was 50% for treated subjects as compared with 59% for untreated subjects. The comparable figures for women were 35% and 43%, respectively (Pelissier et al., 2003; Pelissier et al., 2001). This study assesses whether the positive outcomes found for all programs grouped by gender are similar across programs.

There were 16 programs for men and four for women which used a cognitive-behavioral treatment approach in the multi-site evaluation. Although the intent was to standardize programs by disseminating program manuals, variation across programs is likely to occur for several aspects of program implementation such as staff experience and the incorporation of other program components such as program emphases and program policies related to rule enforcement.

Method

Program Description

Treatment approach, structure, content were the targets of the BOP's efforts to standardize its substance abuse treatment programs. The BOP's first three residential substance abuse treatment programs used the relapse prevention approach as a core element of its cognitive-behavioral treatment model (Marlatt & Gordon, 1985). These programs implemented a 1,000-hour, 12-month duration protocol. The treatment programs subsequently implemented (all those currently operational) the same treatment approach, but with a 500-hour, 9-month protocol.

Written treatment manuals and on- and off-site training were used in an attempt to create a consistent treatment approach across all the programs. The manuals prescribed content for 350 of the 500 required hours of treatment and contained lesson plans which included a presentation guide for the substance abuse treatment counselor and activity worksheets to be used during group sessions and for homework assignments. The manuals contained outlines of didactic presentations and group exercises for the following program modules: Screening and Assessment, Treatment Orientation, Cognitive Skill Building, Relapse Prevention, Interpersonal Skill Building, Criminal Lifestyle Confrontation, Wellness, and Transitional Programming. Treatment involved psycho-education and group process. The content for the remaining program hours was to be at the discretion of the individual substance abuse program coordinators. Much of the discretionary time was expected to be used in process groups extending the time spent on one or more of the program modules.

Other attempts at program standardization pertained to staffing and program monitoring.

All program directors were doctoral level psychologists and standard criteria were developed for the hiring of substance abuse treatment specialists, the primary treatment providers. Regional program coordinator positions were established with goals including the provision of assistance to programs in establishing a monitoring process to maintain program quality and integrity. Specific duties of these regional positions subsequently outlined in policy called for on-site consultation for advice on program implementation and operations, on-site training, technical assistance and technology transfer.

With the exception of one program, participants started treatment in cohorts of approximately 20 individuals who all progressed through the various stages of treatment together. Individual counseling was available, but was not provided on a frequent basis. Group sessions were usually limited to approximately 10 to 12 individuals.

From the opening of the BOP's first three substance abuse treatment programs in 1990, standard program eligibility requirements were put in place. Individuals had to volunteer for treatment, as well as have 15 to 36 months remaining to serve in prison, and have a history of moderate to severe drug abuse. The criteria excluded those with a history of violent or assaultive behavior during their current incarceration, those who were not fluent in English, and those who had serious mental health problems.

The treatment programs also shared other structural characteristics: (1) Treatment was unit-based and each unit housed approximately 100 individuals, (2) all residents lived together in one housing unit which was separate from the general population with the intent of building a treatment community, (3) staff-client ratios were 1:12 for the 500-hour protocol and 1:24 for the 1,000-hour protocol, (4) treatment comprised two 2-hour sessions, five days a week, with the other half of the day spent in general institutional activities such as work assignments, education programs, and recreational activities along with individuals in the general population, and (5) each program, by virtue of being housed in a federal prison, also shared many aspects of the federal correctional milieu. For example, rule violations such as a positive urinalysis test resulted in

program discharge in all cases because programs were required to follow general prison policies regarding rule violations.

Sample

The total sample included 1,343 individuals – 1,065 men and 278 women – who were included in the multi-site substance abuse treatment evaluation. These individuals were from 2 programs and were substance abuse treatment participants between 1990 and 1995. Most of these individuals were within one year of release from BOP custody when they completed the program.

Measures

Outcome Variables. We used two outcome measures in the analyses: recidivism and post-release drug use, both for a follow-up period of up to three years following release from prison. The data for these outcome measures were obtained from telephone interviews with Probation officers which were conducted at three points in time: 6 months, 18 months, and 3 years, unless the individual was revoked or completed supervision before the 3-year time point. The interviewers obtained information on the frequency of urine testing, date and type of drug for each positive urine test, and date and type of arrest for all arrests occurring within the follow-up time period. Approximately 89% of the subjects were released from prison to the supervision of a Probation officer. For those not released to supervision, arrest data were obtained from National Crime Information Center (NCIC) data. The outcome measure of drug use was limited to those released to supervision who also received urine tests.

Three different indicators of recidivism were used to gauge the robustness of the results. These three indicators of recidivism included: (1) arrest for a new offense or revocation of supervision among those supervised by a Probation officer, (2) arrest for a new offense only for those released to supervision by a Probation officer and (3) arrest for a new offense for all individuals, including both those released to supervision and those not released to supervision. Arrest was defined as the first occurrence of an arrest within 3 years after release from custody and when there were multiple charges associated with the first arrest, we analyzed the single most

serious offense.

The arrest or revocation indicator was selected because a technical revocation is a competing event with rearrest for a new offense. A revocation event triggered solely by violation of a condition of supervision can be separate from an arrest for a new offense, yet both events may share similar underlying antecedents (e.g., a return to drug use can trigger both a new offense or a technical violation). Therefore, this measure of recidivism combines new arrest and revocation in order to prevent possible biases in solely predicting arrests for a new offense.

On the other hand, as probation districts vary in their policies regarding revocation of supervision, particularly in response to positive urinalysis tests, combining arrest and revocation may confound behaviors of released offenders with responses of Probation officers. Thus, the other measures of recidivism were limited to arrest for a new offense. There were two different indicators of arrest for a new offense. One was limited to the supervised individuals and the other included both supervised and unsupervised individuals. The indicator including both supervised and unsupervised individuals allowed an assessment of the role of the supervision process itself.

The second outcome measure was drug use after release from prison. The drug use measure included four different violation of supervision categories: (1) positive urinalysis test for any illegal drug, (2) refusal to submit to a urine test, (3) positive breathalyzer test for alcohol, and (4) admission of drug use to the Probation officer. When a person refused a urine test, the assumption was that he or she would have had a positive urine test result. This outcome measure was limited to individuals released to supervision and tested for drug use. There were 56 men and 14 women released to supervision who were not included in the analyses of drug use because they had no urinalysis testing. Although a more appropriate measure of drug use would be the total number of drug use occurrences, our analyses are limited to the first occurrence. In some Federal judicial districts individuals are revoked after the first or second positive urine test while in other districts individuals are revoked only after repeated positive test results.

Predictor and Control Variables. Individual level predictor variables included

background and other factors which prior research and our previous evaluation study found to be important predictors of recidivism and drug use (Beck & Shipley, 1989; Gendreau, Little, & Goggin, 1996; Hepburn & Albonetti, 1994; Martin, Butzin, Saum, & Inciardi, 1999; Pelissier et al., 2003; Schmidt & Witte, 1988). These predictors included race, ethnicity, employment before incarceration, educational level, history of prior commitment, age at first commitment, age at release, having had a spouse with a drug problem, type of daily drug use in the year before arrest, drug and alcohol treatment history (used only in drug use models), DSM-III-R diagnoses of antisocial personality and depression (American Psychiatric Association, 1987), mental health treatment history, treatment completion, prison rule violation, and living arrangement after release. Factors indicative of level of supervision such as post-release treatment services and level of supervision received were included as control variables since these factors could affect the likelihood of detection for either of the two post-release outcomes. Two other in-prison control factors included the length of time between creation of the program and treatment entry for each participant as well as the amount of time served after discharge from the treatment program.

Most of the predictor measures were obtained from BOP automated data files and face-to-face interviews with individuals. The DSM-III-R diagnostic data were obtained using the Diagnostic Interview Schedule (Robins, Helzer, Croughan, & Ratcliff, 1981). Informed consent procedures were used and 90% of the individuals approached for participation agreed to be interviewed. Post-release predictor variables were obtained from telephone interviews with Probation officers for up to a 3-year period after release from BOP custody (see outcome measures).

Each program was included as a predictor variable to assess whether the outcome for one or more programs was statistically different after controlling for the individual level predictors. A description of all the variables and the various categories used for nominal variables are shown in Table I. Effects vector coding was used to code the program covariates and for most of the nominal and ordinal level predictor and control variables. In effects vector coding, each coefficient

represents the contrast of that category with the adjusted grand mean.

Most of the same predictors were used across the various outcome measures, with a few exceptions. For example, we used history of past drug and alcohol treatment in the models of drug use but not recidivism. Additionally, having a drug related rule violation within six months before release was used in the post-release drug use models whereas having one or more serious rule violations within six months before release was used in the recidivism models.

Analytic Strategy

We analyzed the data using survival modeling techniques because they model both the occurrence and timing of the event. This approach also controls for individuals who did not fail during the 3-year post-release observation period and for individuals who were not observed during the entire post-release period due to termination of supervision before the end of the 3-year period (Allison, 1984; Blossfeld, Hamerle, & Mayer, 1989). Early termination of supervision occurred because of incarceration for a detainer, death, or because the term of supervision was less than 3 years. Thus, the results reflect the effect of variables on the survival time, that is, time until recidivism or time until first evidence of drug use during the 3 years of post-release supervision.

Men and women were modeled separately because not only were the original outcome analyses conducted separately because of the lower recidivism and post-release drug use rates of women (Rhodes et al., 2001), but also because previous research suggests that the process of recovery may differ between men and women (Messina, Wish, & Nemes, 2000; Pelissier et al., 2003). The separate models ensure that gender differences would not be confounded with other program differences.

Discrete time proportional hazards regression with random effects was used to model time to failure (Jenkins, 1995; Meyer, 1990; Prentice & Gloeckler, 1978). This approach prevents biased estimates when there is the presumption that unmeasured factors are present which differentiate between the various programs (Allison, 1995; Hosmer & Lemeshow, 1999;

Hougaard, 2000). As we are first seeking to assess whether programs vary in post-program outcomes, the random effects models test whether there is evidence of significant variation in outcomes across the various treatment programs. Moreover, if there are unmeasured factors which differentiate programs then this would create a correlation between unmeasured and measured covariates violating the assumption of independent error terms. The use of a discrete time proportional hazards model with random effects allowed us to simultaneously take into account the censored nature of the data and unmeasured program differences. Due to limited variability across female programs, these analyses use fixed effects survival analyses.

The follow-up period was 36 months after release from prison. We examined the distribution of outcomes in each of the 36 monthly intervals to ensure that an event (e.g., recidivism, drug use) occurred in each time period. Because recidivism and drug use did not occur in each of the 36 months across all measures among both men and women, it was necessary to collapse the intervals into discrete categories with at least one event.¹ In presentation of the results, we do not discuss (or present) the effects of the individual level predictors or the control variables as our interest is primarily in assessing whether there is evidence of unmeasured covariates and whether or not there are program effects.²

To assess whether or not there were program differences in post-release outcomes, we entered each program as a predictor into the models. This allows not only the identification of whether or not one or more programs had statistically different outcomes from all the programs on average, but also identifies the direction of the variation (e.g., more or less favorable). A positive coefficient indicates a greater likelihood of recidivism or drug use and a negative coefficient indicates a smaller likelihood of recidivism or drug use. If program differences are found, the second step of the analysis will be to identify possible program factors which may be related to outcomes.

Results

We begin by providing a description of how the programs both differed and were similar

in various aspects of program implementation. This information provides a context for interpreting the results of the multivariate models. This is followed by a profile of the individuals in the sample and then with the results for the event history models.

Program Implementation Similarities and Differences

Although the intent was to have programs standardized through program content guidelines and similarity of settings and BOP policies, it is likely that programs did not equally implement the standardized content and that, even if they did, there was variation in program emphases and staff experience. Treatment staff perceptions of the programs provide one measure of similarities and differences between the programs. Staff perceptions were obtained through anonymous surveys administered on a yearly basis between 1993 and 1995 to all treatment staff at the programs involved in the multi-site evaluation of substance abuse treatment. Staff survey information was available from 19 of the 20 programs.³

A review of the staff survey data showed that there were many similarities across the programs but differences as well. Staff survey information indicated a high degree of similarity for various elements of program foci which confirmed that there was some degree of uniformity in implementation of a cognitive-behavioral approach. Ninety percent or more of the staff across the 19 programs reported a great or a very great program focus upon abstinence, a non-criminal lifestyle, interpersonal skills, problem-solving and physical health. However, there were some focus areas with variation between programs. For example, between 60 and 77 percent of staff overall reported feeling that there was great or very great focus upon values clarification, enduring personality changes and the therapeutic community environment and the percentages were similar between staff at male and female programs. However, these percentages ranged considerably from one program to another. For example, the percentage of staff perceiving a great or very great focus upon these elements ranged between 0 and 100 percent for the therapeutic community environment and between 20 and 100 percent for the other two items.

Regarding other differences, there was variation in the experience and education levels of

staff. Approximately 40 percent of the staff overall had three or more years of drug and alcohol counseling experience before coming to work in the prison-based programs. Within specific male programs the percentage of staff with 3 or more years of experience ranged between 0 and 100 percent. The comparable figures among the programs for women were 0 to 40 percent. Overall, sixty-percent of the staff reported having a master's or doctoral degree but the percentage at each program varied between 0 and 100 percent.

One of the major domains in which programs clearly varied in their implementation pertained to the expected participation and responsibility of program participants and program discharge rules. The percentage of staff in each program agreeing or strongly agreeing that participants are expected to demonstrate continued concrete progress toward their goals ranged between 0 and 67 percent. The percentage range across programs of agreement or strong agreement in emphasizing the use of peer pressure to induce conformity or in the degree to which program participants are sanctioned in programs by loss of privileges to a great or very great extent was 0 to 75 percent.

There was program variation in the various reasons for program removal except for violation of BOP policies such as drug use (over which programs had no control). For example, the variation in the percentage of program staff who felt great or very great importance is given to removal for violating programs rules was 0 to 100 percent. Comparable figures for the importance given to removal for lack of progress in the program were 0 to 67 percent.

Sample characteristics

There were total of 948 male and 245 female program participants in the analysis of arrest or revocation among supervised individuals.⁴ The sample size for this outcome measure in the 16 male programs ranged between 11 and 124 with eight programs having at least 50 participants. The sample size in the four female programs ranged between 12 and 121. There was variation in the sample size between each outcome measure because some measures were limited to supervised individuals and because the drug use measure was further limited to individuals who

received urinalysis testing. The largest sample size was for the measure of arrest for a new offense only among all individuals. For this measure there were 1,065 men and 278 women. In contrast, for the outcome measure of drug use there were 892 men and 231 women. The sample sizes for each of the models are provided in the Tables 2 and 3 with the outcome results.

Sixty-five percent of the men in the arrest or revocation sample were white, 8 percent were of Hispanic origin, and the mean age at release from prison was 37.5 years. Approximately 68 percent of the men had a prior commitment, 57 percent were employed full- or part-time in the month prior to imprisonment, and the mean education level was 12 years. In reference to daily drug and alcohol use in the year prior to arrest, about 31 percent reported having used illicit drugs other than marijuana (such as heroin or cocaine), 14 percent reported using marijuana, and 18 percent reported using alcohol only. More than 80 percent of the men completed the treatment program. After release from prison, 22 percent lived with a spouse and 18 percent lived in a common law arrangement.

Approximately one half of the women were white, 8 percent were of Hispanic origin, and the mean age at release was 35 years. The mean education level was 11 years, 38 percent reported that they had been employed full- or part-time in the month prior to incarceration, and 43 percent had a prior commitment. Furthermore, approximately 54 percent of the women in the sample reported using illicit drugs other than marijuana (such as heroin or cocaine) on a daily basis in the year before arrest, while 11 percent said the only illegal drug they had used was marijuana, and 10 percent said they had used alcohol only on a daily basis. Approximately 70 percent of the women completed treatment. After release and during supervision, 11 percent of the women lived with a spouse and 8 percent lived in a common law arrangement.

Recidivism models

The multivariate results for the three recidivism indicators among men, arrest for new offense and arrest or revocation, are shown in Models 1, 2 and 3 in Table 2. In all three models, none of the program coefficients were found to be statistically significant ($p < .003$).⁵ Thus, no male

program had a higher or lower likelihood of recidivism than all programs on average. The likelihood ratio assesses whether the addition of the random effect term was statistically significant. The likelihood ratio test, which was not significant for either model indicated that there was no evidence of significant unmeasured factors.

Table 3 provides the recidivism results for female programs. The results for the three indicators of recidivism among women showed that there were significant program differences for only one indicator, that of arrest for a new offense only among all individuals (Model 1). The coefficient for Female Site 3 was negative, indicating that women from this program were less likely to be arrested than women from all four programs on average.

Drug use models

The results for the drug use outcome measure among 892 men are reported in Model 4 of Table 2. After controlling for individual and treatment factors, none of the programs demonstrated a statistically significant effect on the likelihood of drug detection in the 3 years following release from prison. The random effect component of this model was also not found to be statistically significant, indicating no evidence of unobserved heterogeneity, that is, of unmeasured factors.

The results for drug use for the 231 women who were released to supervision and had urine drug tests showed that the coefficient for one program – Female Site 1 – was statistically significant (see Table 3). Female program participants at Site 1 had a greater likelihood of post-release drug use than individuals on average from all programs.

Discussion

This study adds to the discourse on multi-site comparisons of substance abuse treatment programs. The results of this study indicated that, for the 16 male programs evaluated, the overall positive results found in the initial evaluation (Pelissier et al., 2003; Pelissier et al., 2001; Rhodes et al., 2001) were consistent from one program to another. On average, after controlling for individual level characteristics, men who participated at any one of the 16 male treatment

programs were not any more likely to have an arrest for a new offense, arrest for a new offense or revocation, or drug detection than all men on average. This finding occurred in the context of attempts to standardize program content but where there were noted differences and similarities in program implementation and staff experience. It is noteworthy that these same programs showed program differences in participant retention (Pelissier, Camp, & Motivans, 2003). Therefore, these results suggest that the common use of a relapse prevention theoretical approach can yield uniformly positive outcomes despite differences in program implementation. This can reassure program directors and policymakers of the robustness of the cognitive-behavioral model (Mowbray & Herman, 1991).

In contrast to the findings for men, among the four programs for women, the results indicated that participants from one program had a reduced likelihood of being arrested for a new offense only, while participants from another program had an increased likelihood of post-release drug use. It is noteworthy that the program with a significantly higher likelihood of drug use had near significant coefficients for two of the three indicators of recidivism, the two involving arrest for a new offense only, and that the coefficients were in the same direction as those for drug use. This provides evidence, albeit tenuous, of some similarity in patterns of outcome differences across the women's programs for the two domains of outcome - recidivism and drug use.

The results for the three different indicators of recidivism among women appear inconsistent because the coefficient for Female Site 3 was significant for only one indicator, arrest for a new offense among supervised and unsupervised women. However, the z test showed that the difference between the coefficients for the two measures of new offense only – supervised individuals only versus supervised *and* unsupervised individuals – was not significant (Clogg, Petkova, & Haritou, 1995). The lack of significance when using arrest for a new offense and revocation as an indicator may be confounded by the fact that there is considerable variation among Federal judicial districts on revocation, particularly for drug use. Since revocation reflects not only an individual's behavior but also the response of the Probation officer and/or judge, it is

not surprising that the results are not identical for the three different recidivism indicators. When looking at the measure of arrest or revocation among women, 51 percent of the recidivists were individuals who were revoked for violation of a condition of supervision and not because of a new criminal offense.

Although random effects models could not be developed separately for women's programs due to the small number of programs (e.g., four), we speculate about possible explanations for the results by retrospectively examining available program data using the staff survey information previously reported and information from site visits. Differences were identified for the two female programs, one with a greater likelihood of post-release drug use and one with a lower likelihood of recidivism. These differences suggest that there were relatively unique program structures which may have affected participant outcomes. Staff survey responses at the program with a lower likelihood of recidivism indicated that the program stood out because of its uniform emphasis on clients taking personal responsibility for their behavior as evidenced by the program's focus on treatment progress and consistent adherence to program and institution rules. However, this type of focus did also occur in some of the men's programs. Perhaps the variation in these few items, in and of themselves, were not sufficient to result in differential outcomes and that it was this variation along with other, as of yet unidentified, program differences which led to these results.

Staff survey data for the program with a greater likelihood of post-release drug use characterized the program as one where staff were less experienced than staff at all the other programs, including all men's programs, and where the staff provided less individual counseling than at other programs. Additionally, observation notes from site visits conducted by research staff indicated that this program stood out because treatment staff had to perform a significant amount of correctional duties and because the residential units housed only half of the treatment participants along with other non-treated individuals.

These results provide some support for the premise that cognitive-behavioral treatment,

which has generally been shown to be effective, can be equally effective across a relatively large number of programs despite some program variation. Although prior research indicates that differences in treatment outcomes across programs modalities exist (Anglin & Hser, 1990; Hser et al., 1998; Hubbard et al., 1989), our results suggest that variations in program implementation within programs of the same modality may have less of an impact on treatment outcomes when treatment content is standardized using an empirically validated treatment protocol and when program variation is limited. The finding of differences among programs for women may be due to the fact that there appeared to be greater variation in implementation within these four programs.

A limitation of this study is that the effect for each program represents the outcomes for participants at that program across time. Potential differences in programs may have been masked by intra-program differences across time. Because changes in elements of program implementation sometimes occurred during the treatment of a particular cohort of program participants, it would be difficult to identify demarcations of significant intra-program changes. Furthermore, such an effort would require knowledge about what program elements affect outcomes and thus the identification of crucial markers. One example of such a factor is the extent to which the disseminated program modules with standardized program session guidelines were actually used during treatment sessions. In 1994, 48 percent of the program staff in male programs and 31 percent of the program staff in female programs reported using these guidelines in at least 60 percent of their group sessions. However, in 1995 the percentage increased to 77 percent and 73 percent for male and female programs, respectively. Program factors may also require further clarification and verification with program perceptions obtained from program participants or assessment of program characteristics by observers.

Because this study found variation in outcomes only across female programs where multivariate analyses were not feasible, it was not possible to comprehensively examine the role of program characteristics. Future multi-site studies where program outcomes may vary are needed

to help distinguish the relative importance of treatment approaches (e.g., content) from other program factors such as staff experience, specific program goal emphases, emphasis on program adherence to rules and upon program progress. These studies will need to compare program implementation variations within the same modality and assess which of these treatment program factors result in more or less favorable outcomes. In addition, it would be useful to conduct studies within programs serving both men and women as well as programs serving only women. Our finding that no program differences occurred among the men's programs even though programs varied in specific policies, program emphases, and staff experience may not be replicated in non-prison settings or in other settings where the program variations may manifest themselves in a different fashion. What factors other than program content, if any, are related to program outcomes? In what environment or modality are such factors relevant to program outcomes?

As a result of the “disconnect” between drug treatment research and practice, typical treatment programs do not always implement empirically-proven treatment models (Fletcher & Battjes, 1999). One of the goals of the Clinical Trials Network of NIDA is to transfer the results of multi-site clinical trials to treatment providers and patients. It is likely that within-modality studies are crucial to understanding the effects of program implementation on outcomes because different modalities have different client populations. However, as noted by Hser et al (1998), there is virtually no research on the association between program philosophy and client outcomes. Meta-analyses of drug treatment programs have examined primarily program modality and treatment approaches. The challenge is to develop a theoretical perspective which can identify factors other than program modality and assess which of these factors are relevant to treatment program outcomes and which may also explain similarities and differences across and within community- and prison-based treatment programs. The development of evidence-based corrections (MacKenzie, 2000) will need to base treatment decisions on “credentialed” knowledge (Latessa, Cullen, & Gendreau, 2002) such as research from meta-analyses. An important

component of these meta-analyses will involve the ability to analyze a sufficient number of studies with similarly defined program characteristics.

References

- Allison, P. (1984). *Event history analysis*. Beverly Hills, CA: Sage Publications.
- Allison, P. (1995). *Survival analysis using the SAS system*. Cary, NC: The SAS Institute.
- American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.
- Anglin, M., & Hser, Y. (1990). Treatment of drug abuse. In M. Tonry & J. Wilson (Eds.), *Drugs and Crime* (Vol. 13, pp. 393-460). Chicago: The University of Chicago Press.
- Aos, S., Phipps, P., Barnoski, R., & Lee, R. (2001). *The comparative costs and benefits of programs to reduce crime*. Olympia, Washington: Washington State Institute for Public Policy.
- Beck, A. J., & Shipley, B. E. (1989). *Recidivism of prisoners released in 1983*. Washington, DC: Bureau of Justice Statistics, US Department of Justice.
- Blossfeld, H., Hamerle, A., & Mayer, K. (1989). *Event history analysis: Statistical theory and application in the social sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Campbell, A. L., & Bryant, K. A. (2001). Routine childhood immunizations. *Primary Care*, 28(4), 713-738.
- Clogg, C. C., Petkova, E., & Haritou, A. (1995). Symposium on applied regression: Statistical methods for comparing regression coefficients between models. *American Journal of Sociology*, 100(5), 1261-1293.
- Fletcher, B. W., & Battjes, R. J. (1999). Introduction to the special issue: Treatment process in DATOS. *Drug And Alcohol Dependence*, 57, 81-87.
- Gerstein, D., Datta, R., Ingels, J., Johnson, R., Rasinski, K., Schildhaus, S., & Talley, K. (1997). *Final report: National treatment improvement evaluation survey*. Rockville MD: Substance Abuse and Mental Health Services Administration, Center for Substance Abuse Treatment.
- Grella, C., Joshi, V., & Hser, Y. (2000). Program variation in treatment outcomes among women

- in residential drug treatment. *Evaluation Review*, 24(4), 364-383.
- Gendreau, P., Little, T., & Goggin, C. (1996). A meta-analysis of the predictors of adult offender recidivism: What works! *Criminology*, 34(4), 575-596.
- Hepburn, J. R., & Albonetti, C. A. (1994). Recidivism among drug offenders: A survival analysis of the effects of offender characteristics, type of offense, and two types of intervention. *Journal of Quantitative Criminology*, 10(2), 159-179.
- Horgan, C., Skwara, K. C., & Strickler, G. (2001). *Substance abuse: The nation's number one health problem*. Princeton, New Jersey: Robert Wood Johnson Foundation.
- Hosmer, D., & Lemeshow, S. (1999). *Applied survival analysis: Regression modeling of time to event data*. New York: John Wiley and Sons.
- Hougaard, P. (2000). *Analysis of multivariate survival data*. New York: Springer.
- Hser, Y., Anglin, M. D., & Fletcher, B. (1998). Comparative treatment effectiveness: Effects of program modality and client drug dependence history on drug use reduction. *Journal of Substance Abuse Treatment*, 15(6), 513-523.
- Hser, Y., Maglione, M., Joshi, V., & Chao, B. (1998). *Effects of treatment program and client characteristics on client retention*. Rockville, MD: Center for Substance Abuse Treatment, National Evaluation Data and Technical Assistance Center.
- Hubbard, R., Marsden, M., Rachal, J., Harwood, H., Cavanaugh, E., & Ginzburg, H. (1989). *Drug abuse treatment: A national study of effectiveness*. Chapel Hill, NC: University of North Carolina Press.
- Hubbard, R. L., Craddock, S. G., Flynn, P. M., Anderson, J., & Etheridge, R. M. (1997). Overview of 1-year follow-up outcomes in the Drug Abuse Treatment Outcome Study (DATOS). *Psychology of Addictive Behaviors*, 11(4), 261-278.
- Irvin, J. E., Bowers, C. A., Dunn, M. E., & Wang, M. C. (2000). Efficacy of relapse prevention: A meta-analytic review. *Journal of Consulting and Clinical Psychology*, 67(4), 563-570.
- Jenkins, S. (1995). Easy estimation methods for discrete-time duration models. *Oxford Bulletin of*

- Economics and Statistics*, 57(1), 129-138.
- Koplan, J. P., & Thacker, S. B. (2001). Fifty years of epidemiology at the Centers for Disease Control and Prevention: Significant and consequential. *American Journal of Epidemiology*, 154(11), 982-984.
- Latessa, E. J., Cullen, F., & Gendreau, P. (2002). Beyond correctional quackery: Professionalism and the possibility of effective treatment. *Federal Probation*, 66(2), 43-49.
- Lipsey, M. W., Chapman, G. L., & Landenberger, N. A. (2001). Cognitive-behavioral programs for offenders. *Annals of the American Academy of Political and Social Science*, 578, 144-157.
- MacKenzie, D. L. (2000). Evidence-based corrections: Identifying what works. *Crime and Delinquency*, 46(457-471).
- Malloy, C. D., & Marr, J. S. (2001). Evolution of the control of communicable diseases manual: 1917 to 2000. *Journal of Public Health Management and Practice*, 7(5), 97-104.
- Marlatt, G. A., & Gordon, J. R. (1985). *Relapse prevention: Maintenance strategies in the treatment of addictive behaviors*. New York, NY: Guilford Press.
- Martin, S. S., Butzin, C. A., Saum, C. A., & Inciardi, J. A. (1999). Three-year outcomes of therapeutic community treatment for drug-involved offenders in Delaware: From prison to work release aftercare. *The Prison Journal*, 79(3), 294-320.
- Mathias, R. (2001). NIDA clinical trials network begins first multisite tests of new science-based drug abuse treatments. *NIDA Notes*, 15(6), 1-5.
- Mears, D. P., & Kelly, W. R. (2002). Linking process and outcomes in evaluating a statewide drug treatment program for youthful offenders. *Crime and Delinquency*, 48(1), 99-115.
- Melnick, G., DeLeon, G., Hiller, M. L., & Knight, K. (2000). Therapeutic communities: Diversity in treatment elements. *Substance Use & Misuse*, 35(12-14), 1819-1847.
- Messina, N., Wish, E., & Nemes, S. (2000). Predictors of treatment outcomes in men and women admitted to a therapeutic community. *Am J Drug Alcohol Abuse*, 26(2), 207-227.

- Meyer, B. (1990). Unemployment insurance and unemployment spells. *Econometrica*, 58(4), 757-782.
- Mitchell, O., MacKenzie, D. L., & Wilson, d. B. (2002). *The effectiveness of incarceration-based drug treatment on offending and drug use: An empirical synthesis of the research*. Paper presented at the American Society of Criminology, Chicago, Illinois, November.
- Mowbray, C. T., & Herman, S. E. (1991). Using multiple sites in mental health evaluations: Focus on program theory and implementation issues. *New Directions for Program Evaluation*, 50(summer), 45-57.
- Orwin, R., & Ellis, B. (2000a). *The national treatment improvement evaluation study: Multilevel reanalysis of treatment outcomes*. Fairfax, VA: Center for Substance Abuse Treatment, National Evaluation Data and Technical Assistance Center.
- Orwin, R., & Ellis, B. (2000b). *Treatment Components and Their Relationships With Drug and Alcohol Abstinence*. Fairfax, VA: Center for Substance Abuse Treatment, National Evaluation Data and Technical Assistance Center.
- Orwin, R., & Williams, V. (1999). *The National Treatment Improvement Evaluation Study: Retention Analysis*. Fairfax, VA: Center for Substance Abuse Treatment, National Evaluation Data and Technical Assistance Center.
- Pearson, F. S., & Lipton, D. S. (1999). A meta-analytic review of the effectiveness of corrections-based treatments for drug abuse. *The Prison Journal*, 79(4), 384-410.
- Pelissier, B. M. M., Camp, S. D., Gaes, G. G., Saylor, W. G., & Rhodes, W. (2003). Gender differences in outcomes from prison-based residential treatment. *Journal Of Substance Abuse Treatment*, 24, 1-12.
- Pelissier, B. M. M., Camp, S. D., & Motivans, M. A. (2003). Staying in treatment: How much difference is there from prison to prison? *Psychology of Addictive Behaviors*, 17(2).
- Pelissier, B. M. M., Wallace, S. B., O'Neil, J. A., Gaes, G. G., Camp, S., Rhodes, W., & Saylor, W. (2001). Federal prison residential drug treatment reduces substance use and arrests

- after release. *American Journal of Drug and Alcohol Abuse*, 27(2), 315-337.
- Prendergast, M., & Podus, D. (2000). Drug treatment effectiveness: An examination of conceptual and policy issues. *Substance Use & Misuse*, 35(12-14), 1629-1657.
- Prendergast, M., Podus, D., Chang, E., & Urada, D. (2002). The effectiveness of drug abuse treatment: A meta-analysis of comparison group studies. *Drug And Alcohol Dependence*, 67, 53-72.
- Prentice, R., & Gloeckler, L. (1978). Regression analysis of grouped survival data with application to breast cancer data. *Biometrics*, 34, 57-67.
- Rhodes, W., Pelissier, B. M. M., Gaes, G. G., Saylor, W., Camp, S., & Wallace, S. B. (2001). Alternative solutions to the problem of selection bias in an analysis of federal residential drug treatment programs. *Evaluation Review*, 25(3), 331-369.
- Robins, L., Helzer, J., Croughan, J., & Ratcliff, K. (1981). National Institute of Mental Health Diagnostic Interview Schedule. *Archives of General Psychiatry*, 38, 381-389.
- Schildhaus, S., Gerstein, D. R., Brittingham, A., Gray, F., Dugoni, B., Rubin, R., & Williams, E. (1998). *Services research outcomes study*. Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Statistics.
- Schildhaus, S., Gerstein, D. R., Dugoni, B., Brittingham, A., & Cerbone, F. (2000). Services research outcomes study: Explanation of treatment effectiveness, using individual-level and programmatic variables. *Substance Use & Misuse*, 35(12-14), 1910-2000.
- Schildhaus, S., & Shaw-Taylor, Y. (2002). *Explaining substance abuse treatment outcomes using individual and programmatic variables*. Fairfax, Virginia: National Evaluation Data Services.
- Schmidt, P., & Witte, A. D. (1988). *Predicting recidivism using survival models*. New York: Springer-Verlag.
- Simpson, D. D., Wexler, H. K., Inciardi, J. A., & (Eds.). (1999). Drug treatment outcomes for correctional settings, part 1 [Special issue]. *The Prison Journal*, 79(3), 291-371.

Wexler, H. K., Melnick, G., Lowe, L., & Peters, J. (1999). Three-year reincarceration outcomes for Amity in-prison therapeutic community and aftercare in California. *The Prison Journal*, 79(3), 321-336.

Footnotes

¹We conducted sensitivity analyses by collapsing the monthly intervals into several different interval combinations and found that the results did not vary substantially. The choice of intervals reflects the dual priorities of ensuring model stability and maximizing the number of intervals to be used in the analysis.

² Tables with results including the individual level predictor and control variables are available upon request.

³ The numbers of respondents varied by the number of staff at each program and the thus the number of respondents per program ranged from 3 to 11.

⁴ While descriptive statistics are provided only for the men and women included in the analysis of arrest or revocation, the distribution of characteristics is similar for the samples included in the other analyses.

⁵ The Bonferroni correction was used to assess significance of a coefficient at $p=.05$. Thus, the significance level required for a program to attain a significance level of .05 was .003 ($.05/16$) for men and .012 ($.05/4$) for women.

Authors Note

This research was funded in part by the National Institute on Drug Abuse through an interagency agreement with the Bureau of Prisons (BOP MOU-201-0). We thank anonymous reviewers and Barry Brown for comments on a previous version of this article.

Dr. Bernadette Pelissier, TRIAD Project Director, has been with the Federal Bureau of Prisons for 19 years. She has been director of the multi-site TRIAD drug treatment evaluation project and is directing an evaluation of a residential prison-based sex offender program.

Dr. Mark Motivans is a statistician with the Bureau of Justice Statistics and manages the Federal Justice Statistics Program.

Dr. Jennifer Rounds-Bryant is currently a faculty member at Capella University and is a licensed practicing Clinical Psychologist in North Carolina. She has 12 years of experience in the research and treatment of substance abuse and dependence. Her career has been that of the scientist-practitioner psychologist, working for research institutes, in community-based treatment programs, and for both Federal and state prison systems doing both research and treatment with substance abusers. Her current role as faculty at Capella University adds the third dimension of educator to her experience in the substance abuse treatment field.

Table 1

Description of Individual Level Predictors and Control Variables Used in Multivariate Analyses

Demographic/Background

Non-white; Reference category = white

Hispanic; Reference category = non-Hispanic

Employed in month before incarceration; Reference category = not employed

Number of years of education

Prior commitment; Reference category = no prior commitment

Age at time of first commitment to prison

Age at time of release from current incarceration

Spouse ever had a drug problem; Reference category = spouse never had a drug problem

Drug and Alcohol Use/Treatment History

Daily drug/alcohol use in the year before arrest

Used alcohol only; Used marijuana only; Used illicit drug other than marijuana;

Reference category = No daily alcohol/drug use

No prior drug treatment; Reference category = prior drug treatment

No prior alcohol treatment; Reference category = prior alcohol treatment

Mental Health/Treatment History

Depression and/or Antisocial Personality Diagnosis (DSM-III-R)

Depression only; Antisocial personality only; Both depression and antisocial personality;

Reference category = neither depression nor antisocial personality

No prior mental health treatment; Reference category = prior mental health treatment

Table I (Continued)

Treatment/Behavior in Prison

Completed drug treatment program; Reference category = did not complete

Serious rule violation within 6 months before release Reference category = no rule violation

Drug related rule violation within 6 months before release; Reference category = no rule violation

Post-Release Behavior

Living Arrangements after release

Lived with spouse; Lived with common law partner after release;

Reference category = did not live with spouse or common law

In-Prison (Control variables)

Number of months between when a program started and when individual entered program

Number of months between discharge from drug treatment and release from custody

Post-Treatment Services (Control variables)

Mean number of face-to-face visits with Probation officer per month - first 6 months after release

Mean number of urinalysis tests administered per month - first 6 months after release

Mean number of collateral contacts with Probation officer per month - first 6 months after release

Participated in self help groups after release; Reference category = did not participate

Post-release treatment services received

Received transitional services only; Received group or individual counseling only; Received post-release treatment and transitional services only; Received transitional services and individual or group counseling; Received transitional services and both individual and group counseling;

Reference category = no post-release treatment/transitional services

Table 2
Random Effects Discrete Time Hazard Models for Four Outcomes: Male Site Effects

	Model 1	Model 2	Model 3	Model 4				
	Rearrest for a New Offense Supervised and Unsupervised 1065	Rearrest for a New Offense Supervised Only 948	Rearrest or Revocation Supervised Only 948	Drug Detection Supervised Only 892				
SiteCovariates	B ¹	SE ¹	B	SE	B	SE	B	SE
Male Site 1	0.25	0.17	0.20	0.19	0.01	0.16	-0.10	0.18
Male Site 2	0.24	0.18	0.10	0.19	0.02	0.16	0.20	0.17
Male Site 3	-0.51	0.33	-0.46	0.40	-0.24	0.26	0.13	0.25
Male Site 4	0.19	0.27	0.31	0.27	0.10	0.24	-0.55	0.31
Male Site 5	0.01	0.22	-0.14	0.23	0.09	0.18	0.24	0.20
Male Site 6	-0.31	0.20	-0.51	0.23	-0.35	0.19	-0.08	0.19
Male Site 7	0.45	0.18	0.49	0.19	0.27	0.17	0.08	0.19
Male Site 8	0.09	0.33	0.01	0.35	0.10	0.28	-0.42	0.38
Male Site 9	0.15	0.25	0.31	0.33	0.05	0.29	0.38	0.30
Male Site 10	0.06	0.27	0.02	0.31	-0.09	0.26	-0.05	0.31
Male Site 11	0.13	0.23	0.14	0.24	0.03	0.26	-0.05	0.22
Male Site 12	0.14	0.19	0.07	0.21	0.06	0.17	0.13	0.19
Male Site 13	0.02	0.30	0.10	0.35	0.12	0.28	-0.06	0.28
Male Site 14	0.09	0.17	-0.12	0.20	-0.14	0.17	0.08	0.17
Male Site 15	-0.70	0.56	-0.16	0.57	0.01	0.43	0.47	0.46
Male Site 16	-0.29	0.43	-0.36	0.44	-0.03	0.33	-0.38	0.43
Constant	-1.86 *	0.45	-1.68 *	0.52	-1.93 *	0.44	-1.67	0.49
ln_varg	-14.66	205.01	-14.57	203.02	-1.97 *	0.91	-1.44	0.63
Sample size	1065		948		948		892	
LR Test ²	-.00001218		-.00001353		1.32580		2.81083	

¹ B = Unstandardized regression coefficient, SE = Standard error

² Test of model with gamma random effect term vs. model without random term

* p < .003 determined by computing Bonferroni correction for 16 contrasts

Table 3
Fixed Effects Discrete Time Hazard Models for Four Outcomes: Female Site Effects

	Model 1		Model 2		Model 3		Model 4	
	Rearrest for a New Offense Supervised and Unsupervised 278		Rearrest for a New Offense Supervised Only 245		Rearrest or Revocation Supervised Only 245		Drug Detection Supervised Only 231	
Site Covariates	B ¹	SE ¹	B	SE	B	SE	B	SE
Female Site 1	0.82	0.42	1.05	0.52	0.58	0.34	0.94 *	0.34
Female Site 2	0.43	0.50	0.40	0.72	-0.13	0.51	-0.26	0.49
Female Site 3	-1.46*	0.64	-1.92	0.93	-0.55	0.41	-0.71	0.39
Female Site 4	0.21	0.35	0.47	0.46	0.10	0.29	0.03	0.28
Constant	-3.52	1.19	-3.69 *	1.54	-3.26 *	1.05	-3.61 *	1.19
Sample size	278		245		245		231	
LR X^2 / df ²	41.18 (24)*		43.05 (33)		78.65 (33)*		88.53 (35)*	

¹ B = Unstandardized regression coefficient, SE = Standard error

² Test of reduced model (constant only) against full model

* $p < .0125$ determined by computing Bonferroni correction for 4 contrasts.