Treatment Research in OZ—Is Randomization the Ideal or Just Somewhere Over the Rainbow?*

Steven S. Martin, Senior Scientist
James A. Inciardi, Professor and Director
Daniel J. O'Connell, Research Associate
Center for Drug and Alcohol Studies

MEETING THE TREATMENT needs of offenders within the correctional system promises an important societal investment in reducing the number of incarcerated drug-involved offenders and the concomitant burgeoning costs of incarceration and health care. Researchers have documented the high costs of drug-abusing offenders whose criminal activity, criminal justice costs, often poor health status, and use of expensive public health services all put heavy burdens on the taxpayer and society (Harwood, Fountain & Livermore 1998; Harwood et al. 1984; Rice et al. 1990; French, Salomé and Carney 2002). Analyses of 26,000 drug users in the National Aids Demonstration Research (NADR) studies found that those who had been incarcerated had significantly higher rates of drug use, multiple drug use, daily drug injections, and unsafe needle use (Inciardi et al. 1993). These and other data (e.g., Chaiken 1989; Leukefeld & Tims 1988; Simpson, Wexler & Inciardi 1999) suggest that chronic drug users are found in the greatest concentrations among prisoners. Thus, correctional institutions should be excellent field settings for identifying concentrations of drug users, implementing treatment programs to a "captive" population, and rigorously assessing drug treatment outcomes among those chronic drug users who are most "expensive" for society.

Emphasizing effective treatment outcomes is necessary because addiction treatment is a service that is largely funded by the public sector. Recent studies show that 70 percent of treatment funding comes from public coffers (Office of Applied Studies 1998). In the current climate of shrinking budgets, especially in state governments that fund most treatment, legislatures are increasingly seeking evidence that money spent on treatment is producing the desired effect. Treatment outcome studies must show that they reduce drug use. Also, directly or by implication, research needs to demonstrate that success in reducing drug use leads to reductions in criminal behavior, improvements in health status, and a decrease in the use of more costly health services—all of which, in turn, generate cost savings to other sectors of society.

Background on Research on Drug Treatment in Corrections

The need for drug treatment within a criminal justice framework is well documented (e.g., Inciardi 1993; Simpson et al. 1999). More contentious is how effective various modalities are, and whether the money spent is recouped later. Research focusing on the effectiveness of residential in-prison treatment has tended to show moderate but significant effects on recidivism and drug usage after release from prison (Gaes et al. 1999; Martin, Butzin & Inciardi 1995; Pelissier et al. 2001; Wexler et al. 1999). Persons receiving treatment in prison followed by continuing treatment in a halfway house show even more promising results than those who only receive in-prison treatment (Martin, Butzin & Inciardi 1995). A recent meta-analysis of 78 treatment outcome studies found that the treated groups reported significantly better outcomes than nontreated groups (Prendergast et al. 2002).

Studies examining the cost effectiveness of various treatment modalities have found substantial returns on money invested. A recent study by French and his colleagues found a cost-benefit ratio of 4:34 for programs studied in Washington State (French, Salomé & Carney 2002). One study (CALDATA) reported the cost-effectiveness of publicly supported treatment programs in California (California Department of Drug and Alcohol Programs 1994). The CALDATA Study reported 18-month savings from treatment of $1.5 billion, with the largest savings coming from reduction in crime, followed by significant reductions in health care costs (ER admissions declined by a third). Studies conducted to date thus indicate that treatment is both effective and cost effective.

Criminal justice research faces daunting hurdles in design and implementation, however, and much of the research cited above suffers from longstanding problems (Apsler, 1991): not having proper control or comparison groups in the design, relying solely on self-reports of drug use and crime, and not having enough individual level impact data. More recently, Gaes has suggested that in-prison treatment designs are plagued by a combination of selection and attrition bias that makes randomization difficult (Gaes 1998). The process of selection, even in a

supposed randomized design, often results in groups that differ from one another in important ways. Gaes suggests that researchers should be cautious in designing comparison groups and recognize potential bias as well as explicitly spelling out the selection and mechanisms involved in the treatment regimen. Apsler (1991) listed additional factors that singly or together would improve treatment outcome research: measures on the variability among treatment programs, long project periods, objective validation of self-report measures, the cooperation of the treatment programs, large samples, multiple measures of treatment experience, and multiple measures of outcomes.

While these criticisms have been taken into account by prison treatment researchers, the dilemmas of conducting field studies that can be rigorously evaluated have proven difficult to solve. Consequently, most assessments of program effectiveness have been solely -*rather* than *outcome* oriented or have not incorporated multiple outcome criteria. Many times when outcome studies have been attempted, they have involved short follow-up time frames, lack of randomization, and have included only limited use of comparison groups, standardized measurement instruments, multivariate models, and appropriate control variables (Forcier 1991; Prendergast et al. 2002; Rouse 1991; Wexler 1995; De Leon, Inciardi & Martin 1995). To cite just one example, in Prendergast et al.'s (2002) meta analysis, only 7.7 percent of studies had a comparison group that actually received *no* treatment. Most received a routine or alternative treatment. Additionally, Prendergast et al. note that 59 percent of the studies in their sample used a random or *quasi*-randomized design. That means that 41 percent of studies were not able to even attempt to randomize the selection of treatment and comparison groups, and it is unclear what is included in the quasi-randomized design in many of the remaining studies.

While the problems with prison treatment studies are well known, what is less often discussed is why correctional research has proven to be so difficult. This paper is an attempt to shed light on why some textbook examples of research methods are largely unworkable and in some cases may be counterproductive in the criminal justice field. The treatment outcome for offenders in the Delaware study described in this paper highlights both problems and practical solutions to some of the above difficulties that are being applied to evaluate a treatment program in an often "uncontrolled" real world setting. The case study demonstrates the need for an effective process evaluation to understand what cannot be a *priori* "controlled" in the "experiment."

We go on to posit a "mixed mode" outcome analysis strategy that includes comparing five client groups, two of which are randomly selected and three of which are not. Then, using the existing data, we posit some research hypotheses and give examples of an outcome that has been modeled in a multivariate analysis designed to control for known group differences. Although non-random group selection makes statistical judgment of significant effects open to question, such analyses may sometimes be necessary to make comparisons in field experiments.

Methods and Results of the Random Experiment

The study focuses on evaluating aspects of a multi-stage therapeutic community (TC) treatment program that was started as a research demonstration project in 1990 and which is now a continuing program in the Delaware correctional system. The use of TCs has expanded rapidly in prisons and community corrections settings. By the year 2000 over 300 TCs were operating in 47 states, and TCs currently operate in 54 countries (Rockholz 2000). In Delaware, there is an integrated continuum of corrections-based TC treatment that works in three stages tied to an inmate's changing correctional status: prison → work release → parole (Inciardi, Lockwood & Martin 1991, 1994). The effectiveness of such a continuum of correctional TC treatment with a focus on the work release stage has been shown to be more effective than in-prison treatment without the treatment continuum (Martin et al. 1999; Butzin et al. 2002; Wexler et al. 1999).

The original goal of the Delaware research was to examine the feasibility and clinical efficacy of a therapeutic community "work release" center for drug-involved felony offenders who had spent a number of years in prison. The issues of feasibility and efficacy were especially important, since the work release TC (CREST) represented the first attempt anywhere at developing a correctional work release program built on a therapeutic community model. The research design to evaluate CREST was primarily experimental, involving a randomized trial of the drug–involved inmates assigned to CREST with a group of drug-involved inmates assigned to regular work release.

Specifically, the design included a randomly selected sample of conventional work release residents with a past history of heavy drug use (the COMPARISON group). These releasees have relative freedom during working hours, but are held in secure dormitories after 10 p.m. Most attend AA/NA meetings at the work release center and have access to an on-site counselor but have little other treatment. The true "experimental" contrast in the study was between this COMPARISON group and the RANDOM-CREST group, a random sample of work release clients with a history of past heavy drug use who were assigned to the CREST TC on a random basis. So, subjects coming to work release with a history of past drug use but no prison TC experience were randomly assigned to one of these two groups.

These two groups are compared in terms of relapse and recidivism measures 12 months after completing work release. The basic hypotheses can be stated as: Drug-involved offenders receiving treatment in a TC are more likely to remain arrest-free and be less drug-involved than those who do not have treatment. Other baseline characteristics thought to be related to relapse and recidivism are controlled in the model. The baseline measures are self-report items. Dichotomous baseline measures include gender, previous drug treatment, and ethnic group (White/NonWhite). Frequency of drug use was derived from questions asking frequency of use of each of the following: injecting or noninjecting cocaine, heroin, speed, crack, PCP, hallucinogens, and non-prescribed sedatives, stimulants, tranquilizers, analgesics or other opiates in the six months prior to prison. The maximum reported use of any drug was recorded on a scale of 0 (no use) to 6 (use more than once a day). Continuous baseline measures were number of prior arrests, number of previous incarcerations, and age. An examination of baseline characteristics in Table 1 suggests that the randomization was effective in producing reasonably equivalent groups. The only difference that approaches significance is percentage "White."

To examine the effect of treatment group in the standard randomized design (treatment verses comparison group), we report the results of regression analyses predicting to: 1) recidivism (logistic regression predicting the likelihood of remaining arrest-free) and 2) degree of relapse (OLS regression predicting the frequency of drug use) one year after leaving work release.

Figure 1 shows the predicted probabilities (shown as percentages) of arrest-free within each group one year after leaving work release, controlling for the mean effects of the other covariates. The other covariates that are significant in the model are age (older more likely to be arrest-free) and number of previous times arrested and number of previous times imprisoned (the more previous arrests and more times in prison, the less the probability of being arrest-free). It is apparent that the RANDOM-CREST group is significantly more likely to be arrest-free at follow-up (58 percent) compared to the COMPARISON group (43 percent).

Figure 2 presents the results of an OLS regression predicting scores on the dependent variable Frequency of Drug Use one year after work release

TABLE 1
Baseline Sample Characteristics by Group:
Delaware Therapeutic Community Continuum

	Comparison Group	Random Crest
N	248	182
Age	29.8	29.2
Number of Arrests	9.6	9.7
Times in Prison	3.2	2.9
Illegal Drugs Used	5.7	5.8
Drug use before prison	3.9	4.7
Male percent	81	77
White percent	30	25
Prior Treatment percent	75	79
Arrest-free at 12 Months percent	43	56

FIGURE 1
Percent Arrest Free 12 Months After Leaving Work Release,
Randomized Model
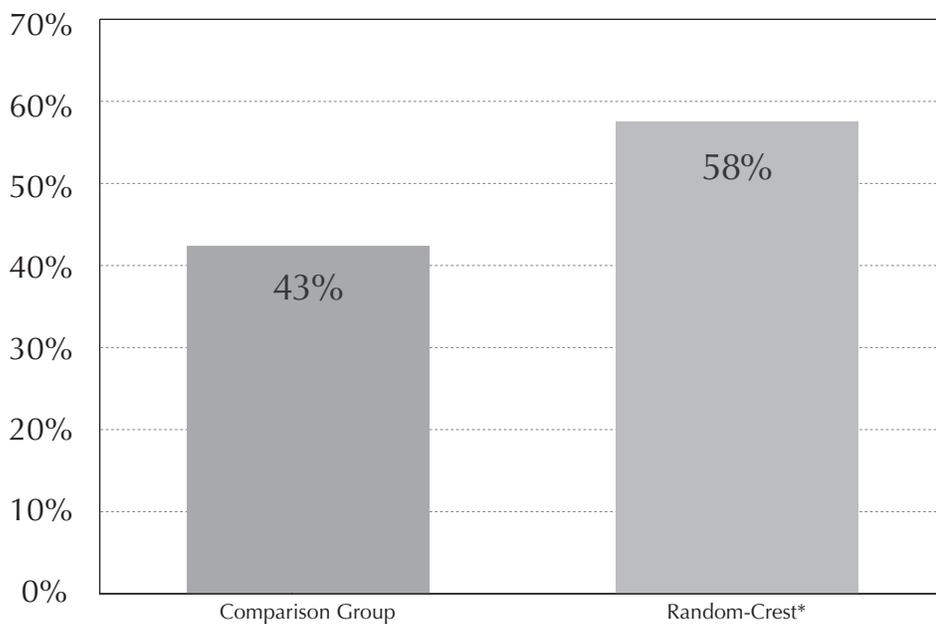
*Significantly different from COMPARISON group, p<.05

Note: Predicted probabilities (shown as percents) of arrest-free by group controlling for mean scores on age, number of prior arrests, times in prison, number of illegal drugs used frequency of drug use prior to prison, gender, race, and prior treatment.

for each experimental group, again adjusted for the other independent variables. The covariates significant in this model are age (older clients have less drug use), previous times in prison (fewer times in prison the less drug use), and previous drug history (the more baseline drug use, the more follow-up drug use).

Again it is apparent that the treatment group is doing much better than the comparison group

one year after work release. The COMPARISON cases are averaging illegal drug use once a week or more often while the RANDOM-CREST group is averaging once a month use. In logistic regression analyses not reported here, the treatment group is significantly more likely to have used no drugs; in an OLS regression analysis among those who have used any illegal drugs, the treatment group uses less often.

Issues of Client Selection

A paper reporting these research results would be a useful contribution and likely accepted in peer review journals. The experimental contrast with the randomly selected groups produces significant and meaningful effects in the predicted direction and strongly supports the efficacy of a transitional TC for drug-involved work release clients. However, clients entering correctional TC treatment rarely get there by a random selection process (chaotic, yes; but random, no).

There were three other relevant offender treatment groups existing during and after the random sample selection that were not part of the experimental manipulation, but from whom baseline and follow-up data were collected. The first group is NON-RANDOM CREST—those assigned to CREST by various criminal justice practitioners. Persons in this group were not randomly assigned, but were placed in CREST by a judge, prison counselor, or prison review board. Most NON-RANDOM CREST clients were recruited after the random selection process stopped and treatment was taken over by the State. In addition, two groups who had been in the in-prison therapeutic community in Delaware, the KEY, are being followed as part of this study: 1) the KEY group releasees from the in-prison TC who did not go to CREST because they were released before CREST was operational or who "maxed out" their sentence and did not have to go to work release; and 2) the KEY-CREST group—all of those clients who graduated from the KEY and then went on to CREST for work release treatment. Although not randomly selected, each of these groups did include all clients coming from the KEY who were being classified for release. More important, each of these groups provides an important contrast with the "experimental" groups. Table 2 lays out some of the salient distinctions among the five groups.

Real world clients in TCs come from several sources. There are "walk-ins" seeking help who are screened and evaluated by staff to determine TC suitability. Prison-based TCs typically "recruit" candidates from the general prison population, followed by screening and evaluation by staff. Finally, many TCs accept or, more likely, are required to take court referrals. Judges will sentence an individual to a prison term, with a portion of the sentence suspended if the person completes the program. Additionally, some clients are referred to KEY or CREST as a result of a parole violation. In none of these scenarios are clients recruited through random assignment.

In fact, the RANDOM-CREST clients, those randomly assigned from a pool of work release eligibles with a history of drug abuse, could be more problematic than the non-random treat-

ment groups. Random assignment had a number of consequences. Some clients were not particularly excited about the prospect of entering CREST, but voluntarily accepted the assignment because they felt that turning it down might delay their move from prison to work release. Most adapted but some did not, and a few tried to poison the treatment environment. Many of these clients would not have gone to CREST without the random selection process.

There were also problems with staff attitude because they were constantly faced with a vocal minority of recalcitrant clients not "clinically" selected into treatment. A statement indicative of the treatment staff's mistrust/confusion about the research was, "Oh, so you don't intentionally send us the most difficult people!" This came after the random assignment process was explained to CREST staff. In reality, all clients had met criteria of past drug abuse and had volunteered for CREST, though perhaps not with a "motivation for treatment." However, TC staff had not assessed and selected the clients, so they found it easy to blame the research process for the "recalcitrant" clients.

The important point here is that, because of the random assignment, the project ended up evaluating a TC treatment arrangement that would not likely exist in reality. The purpose of random assignment is to develop equivalent groups so that valid and reliable comparisons of outcome can be made. But, random assignment made the client mix of the RANDOM-CREST group different from that in "real world" TCs. As noted by Stahler et al. (1993:672) in a random assignment study of homeless crack users to different treatment modalities, "... the randomization process may have interfered with the integrity and internal validity of the design by increasing attrition."

In fact, many studies comparing treatment conditions suggest that client samples based on random designs are different from those selected through traditional recruitment strategies, and randomization may actually change a program. In fact, the research is likely examining an artificial treatment initiative (De Leon 1979; Dennis

FIGURE 2
Frequency of Drug Use 12 Months After Leaving Work Release, Randomized Model
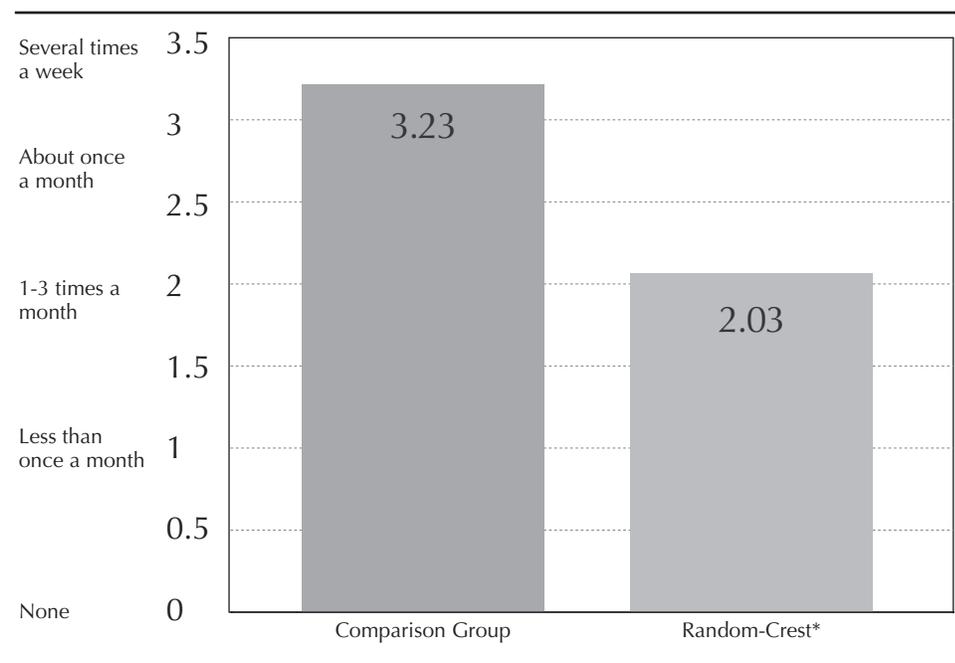
*Significantly different from COMPARISON group, p<.05

Note: Predicted scores on frequency of drug use scale controlling for mean scores on age, number of prior arrests, times in prison, number of illegal drugs used frequency of drug use prior to prison, gender, race, and prior treatment.

1994; Scarpitti, Inciardi & Martin 1994; De Leon et al. 1995). This change is evident from the time of client selection and assignment, and these changes may amplify and reify during the course of the research process.

This points to a conceptual problem with the random model in practice. The model assumes that the error of mismatch in random assignment is also randomly distributed—an error which should not bias any of the assigned conditions (modality, program or intervention) toward higher participation or attrition. For example, the initial attrition rate among those mismatched to a treatment program (e.g., CREST) should be proportional or equivalent to those mismatched by assignment to a no-treatment control group. The evidence from the CREST study suggests that this is not the case.

Even if random selection produces reasonable equivalency of individual differences at the start of the study, participation or attrition in the assigned categories or programs may not be equivalent in their engagement of the assignee, to say nothing about their subsequent influence upon the client. Stahler and colleagues (1993) noted that treatment dropouts often came from the category of clients who felt their assigned program did not meet their treatment or personal needs. In the Delaware study the possible mismatch effect of unmotivated TC clients who are

TABLE 2
Planned Characteristics of Research Groups in the Delaware TC Continuum for Subjects About to be Released From Prison

	Work Release Classified	Past Heavy Drug Use	Random Selection	In-prison TC Graduate	Assigned to work release TC
COMPARISON	yes	yes	yes	no	no
KEY	yes	yes	no	yes	no
RANDOM-CREST	yes	yes	yes	no	yes
NON-RANDOM CREST	yes	yes	no	no	yes
KEY-CREST	yes	yes	no	yes	yes

assigned to CREST (or even select CREST because they think it will help get them out of prison) may lead to no better or worse an outcome than will be found among the "no treatment" group. Thus, the assignment process may be random, but the influence upon the assignee may not be. Since fewer than 60 percent of assignees complete CREST, there is potential for an attrition effect.

To demonstrate the differences in findings and additional information to be gained from a quasi-experimental design (multiple groups resulting from the day-to-day running of the TCs), we repeat the above analyses including not only the random assignment groups but also including naturally occurring treatment groups.

An Example of a "Mixed Mode" Outcome Analysis

As noted earlier, there are five research groups used in these analyses: 1) COMPARISON—those who were placed in the conventional work release setting and received neither prison-based nor community-based TC treatment; 2) KEY, those who received their primary treatment at The KEY but no secondary/tertiary treatment; 3) RANDOM-CREST, those who received their primary and secondary treatment at CREST after being randomly assigned to the program from a pool of work release eligibles; 4) NON-RANDOM-CREST, those who received their primary and secondary treatment at CREST after being sent to the program by normal criminal justice procedures, and 5) KEY-CREST—those who received their primary treatment at The KEY and their secondary/tertiary treatment at CREST.

There are differences in the composition of the research groups: assignment to the COMPARISON or RANDOM-CREST groups was determined by lot; the COMPARISON and both CREST groups include men and women, while the KEY group does not; the KEY and KEY-CREST groups were KEY "graduates" (suggesting some treatment motivation); both CREST groups included all those who started the program, regardless of how much of the program they completed; and finally, the KEY-only group included clients who graduated before CREST was established. Table 3 shows baseline variables for all five groups included in the quasi-experimental analyses.

The RANDOM-CREST and COMPARISON groups remain very similar. There are, however, significant differences with the other groups. The two KEY groups contain more African-Americans. All of KEY and many of KEY-CREST respondents are male. Everyone from the KEY has had previous treatment. The major differences of interest for the present study are between the RANDOM-CREST and NON-RANDOM-CREST groups. The RANDOM-CREST group scored worse on prior drug use. The NON-RANDOM-CREST group scored worse on criminal history, but the difference was not significant. The NON-RANDOM group was also significantly older, by an average of 2.3 years at baseline. Perhaps the most significant difference is in prior treatment. Seventy-nine percent of the RANDOM group reported prior drug treatment, while only 56 percent of the NON-RANDOM group did so. This may reflect decisions on the part of criminal justice practitioners to route those with no past treatment into the TCs.

The basic hypothesis is still that drug-involved offenders receiving treatment in a TC will be more likely to remain arrest free and be less drug-involved 12 months after work release than those who have not had treatment. And, again, logistic regression is used for the arrest-free analysis, while OLS regression is utilized for the drug use analysis. For each dependent variable, we present full models for all 5 groups that were followed. In all analyses the data are examined in the full regression model using a dummy classification for group, with COMPARISON the excluded category.

Figure 3 shows the predicted probabilities of arrest-free within each group one year after work release. The black bar again represents the results of the COMPARISON group and the white bar the RANDOM CREST GROUP. The results are similar to those shown in Figure 1 for these 2 groups, but it is also clear that more is happening. By utilizing all available data, the gray bars show the stair-step result of each additional phase of treatment. Again, the other significant independent variables in the model are age and previous arrest and prison history.

The analyses reveal that transitional treatment in work release seems more effective than in-prison treatment alone in preventing new arrests. Those who get both prison and transitional treatment (KEY-CREST) are the group that does the best.

The difference between the RANDOM and NON-RANDOM CREST groups is of note. While the randomly assigned group did significantly better than the comparison group, the group assigned to CREST by criminal justice practitioners using their own eligibility criteria did even better than the RANDOM-CREST group. Keeping in mind that the NON-RANDOM group scored worse than the RANDOM group on prior criminal history measures; this finding may indicate that the system does an even better job of selecting clients for treatment than random assignment.

A final regression model analogous to that in Figure 2 above but including the five comparison groups is shown in Figure 4. The same 3 covariates (age, times in prison, and previous drug history) are significant here as well. Also, here again, the effects of treatment are seen in the reduced

TABLE 3
Baseline Sample Characteristics by Group: Delaware Therapeutic Community Continuum

	Total	Comparison	Key	Random Crest	Non-Random Crest	Key/ Crest
N	997	248	40	182	320	207
Age	30.4	29.8	31.7	29.2	31.5	30.6
Number of Arrests	10.5	9.6	11.3	9.7	10.8	11.6
Times in Prison	3.1	3.2	3.2	2.9	3.2	3.1
Illegal Drugs Used	5.2	5.7	5.9	5.8	4.7	4.7
Drug use before prison	4.1	3.9	5.1	4.7	3.7	4.3
Male percent	79	81	100	77	79	76
White percent	24	30	15	25	28	20
Prior Treatment percent	74	75	100	79	56	89
Arrest-free at 12 Months percent	59	43	48	56	65	72

FIGURE 3

Percent Arrest Free 12 Months After Leaving Work Release,
Full Group Model
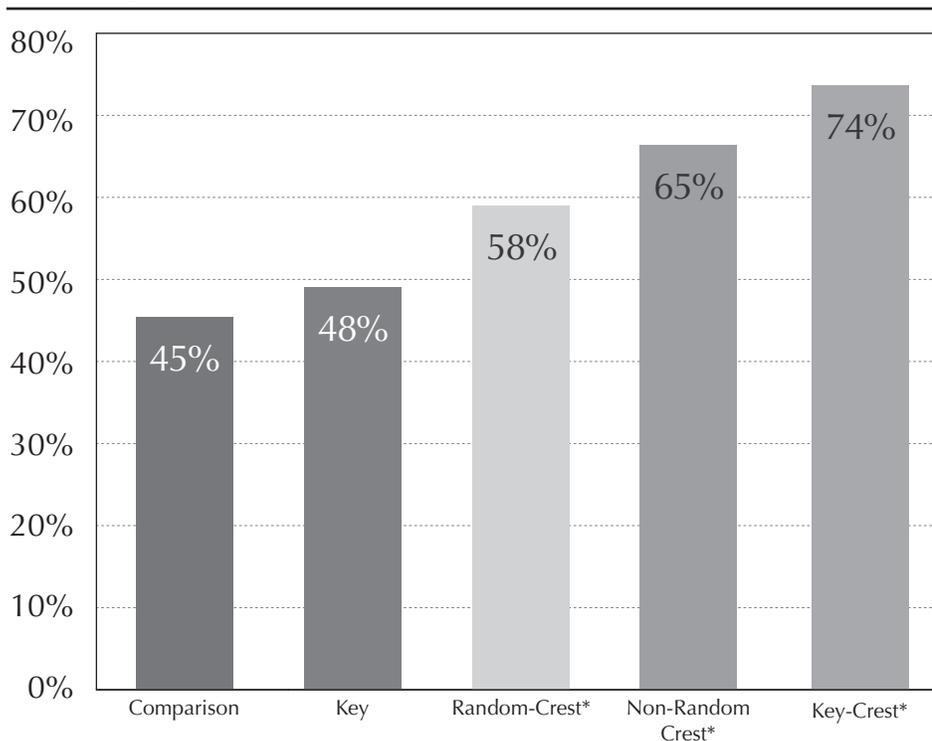
*Significantly different from COMPARISON group, p<.05

Note: Predicted probabilities (shown as percents) of arrest-free by group controlling for mean scores on age, number of prior arrests, times in prison, number of illegal drugs used frequency of drug use prior to prison, gender, race, and prior treatment.

frequency of drug use. Those getting transitional treatment do better than those with in-prison treatment alone, while those with both prison and transitional treatment do the best. Both RANDOM CREST and NON-RANDOM CREST groups show a significant reduction in drug use from the COMPARISON group, and the magnitude of the effect is quite similar. This should be noted in the context that the RANDOM CREST group was significantly more drug-involved than NON-RANDOM CREST. The fact that the more drug-involved and less criminally involved RANDOM CREST group did marginally better in reducing drug use and marginally worse in preventing recidivism than the NON-RANDOM CREST group is worth noting. If there was any implicit difference in selection criteria, it would involve an emphasis on drug use in the randomly selected group and an emphasis on criminal history in the system selected group in determining entry into CREST.

Discussion

Numerous outcome analyses from the Delaware project completed thus far have shown significant treatment effects for the TC continuum for periods ranging from 6 months up to 5 years (Mathias

1995; Inciardi et al. 1997, Martin et al. 1999, Inciardi et al. 2003). These analyses have also indirectly revealed something of the limits of randomization and the necessity for other kinds of controls in the analyses. In this paper, we made the comparisons explicit, looking first at the "experimental" groups and subsequently at the naturally occurring groups. Of particular interest is comparing results between the random and non-random treatment groups (RANDOM CREST and NON-RANDOM CREST). The results suggest that randomization alone may not show the best picture of a treatment program's success, and that clients selected by criminal justice practitioners based on addiction and criminal history criteria beyond work release eligibility perform better than those randomly assigned to CREST.

Client selection for the major "experimental contrast" of the RANDOM-CREST group in this study was based on a random draw from the pool of work release eligible inmates. Yet clinical assessments of readiness and suitability for TC treatment were not used for this group, as is the case in "real world" TCs. In many instances, the RANDOM-CREST admitted clients who, under more typical circumstances, would not have been considered appropriate for a TC.

In the real world of drug abuse treatment, program staff or criminal justice practitioners usually choose the clients they feel are ready for treatment and are appropriate for the particular modality. Random assignment in field settings does not allow for client selection. As a result, clients unready for treatment are assigned to a program, sometimes undermining the effects of treatment and contaminating the treatment environment. Clients who are ready for particular treatments may also be assigned to conditions that are not suitable for them, resulting in attrition or lack of benefits. Consequently, conclusions made about treatment conducted within the context of controlled research may not necessarily apply to treatment conducted with clinically selected and appropriate clients.

De Leon et al. (1995) explicitly examine the dilemmas of conducting research on treatment effectiveness. Federal regulations, real world limitations on accomplishing random case selection, and even the simple knowledge that the program is under study combine to make the circumstances for judging treatment effectiveness elusive and difficult to isolate, describe and quantify. In 2002, new federal guidelines promulgated by the Office of Human Research Protection (OHRB) and interpreted by increasingly vigilant (and even paranoid) local Institutional Review Boards make true "no treatment" control groups unacceptable if there is even the slightest hint that the treatment will be effective (a Catch-22 for the true experiment). Less manipulated research designs may alleviate these problems, but raise new issues about not controlling for the effects of non-manipulated intervening variables. In this paper we used covariate controls. Other more complex controls allowing for more covariates and interactive effects can be accomplished with "propensity score" techniques (D'Agostino 1998). Possible solutions are less intrusive designs with larger samples, replication in different samples, greater emphasis on measuring non-treatment covariates, and assessing a variety of outcome measures—outcome measures that vary in topic (e.g., relapse, other health behaviors, recidivism, employment) and in degree of behavior (e.g., how many ER visits, how often use drugs). An example of using baseline covariates to statistically control for group differences was shown above. Such effectiveness studies, however, require a sufficient number of subjects, the ability to follow subjects over time, and the ability to measure the same variables in different programs and samples.

Overall, our research experience does not call for an abandonment of randomization in treatment research, but a recognition of its limitations. Randomization will not begin to completely "control" for the real differences that will

FIGURE 4

Frequency of Drug Use 12 Months After Leaving Work Release, Full Group Model
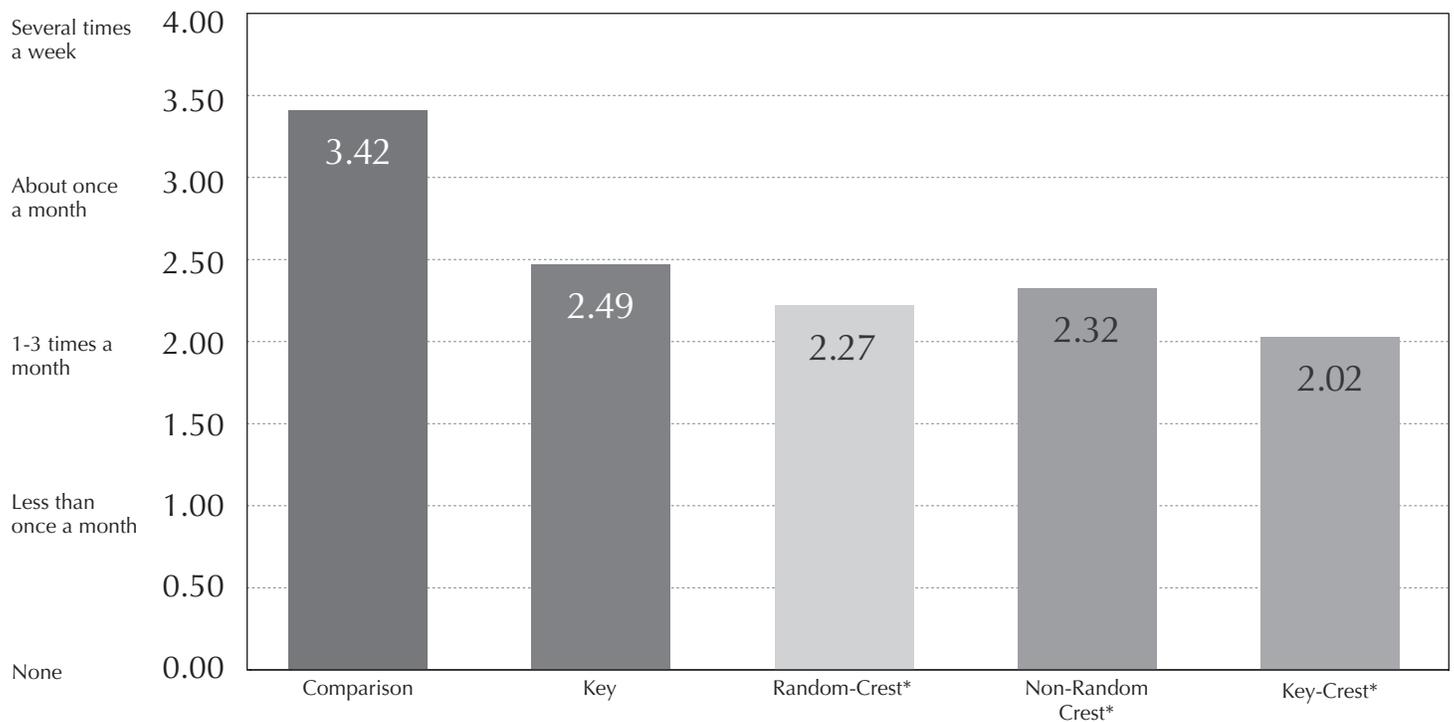
*Significantly different from COMPARISON group, p<.05

Note: Predicted scores on frequency of drug use scale controlling for mean scores on age, number of prior arrests, times in prison, number of illegal drugs used frequency of drug use prior to prison, gender, race, and prior treatment.

remain among the comparison groups, both at initial assignment and during the course of the "quasi-experiment." It may create circumstances not directly applicable to the real world of treatment. And sometimes, it may be important to compare effects among groups that have not or cannot be randomly assigned. As demonstrated earlier, reliance on randomization may obscure the need to measure many other factors related to individual differences and to differences in treatment program contact.

References

Apsler, R. (1991). Evaluating the cost-effectiveness of drug abuse treatment services. In W.S. Cartwright & J.M. Kaple (Eds.), *Economic costs, cost-effectiveness, financing, and community-based drug treatment* (NIDA Research Monograph 113, pp. 57-66). Rockville, MD: National Institute on Drug Abuse.

Butzin, C.A., Martin, S.S., Inciardi, J.A. (2002). Evaluating component effects of a prison-based treatment continuum. *Journal of Substance Abuse Treatment 22*, 63-69.

California Department of Drug and Alcohol Programs (1994). *Evaluating recovery services: The California Drug and Alcohol Treatment Assessment* (CALDATA). Sacramento, CA: California Department of Alcohol and Drug Problems.

Chaiken, M.R. (1989). *In prison programs for drug involved offenders.* Washington, D.C.: Office of Justice Programs, National Institute of Justice.

D'Agostino, Ralph B. (1998). Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Statistics in Medicine 17*, 2265-2281.

De Leon, G. (1979). *The illusion of control groups in therapeutic community* research. Paper presented at the Twelfth Annual Eagleville Conference, Eagleville, PA.

De Leon, G., Inciardi, J.A., & Martin, S.S. (1995). Residential drug treatment research: Are conventional control designs appropriate for assessing treatment effectiveness? *Journal of Psychoactive Drugs, 27*, 85-91.

Dennis, M. L. (1994). Design, management, and analysis of randomly controlled trials in community-based settings. In J.S. Wholey, H. Hatry, and K. Newcomer (Eds.), *Handbook of practical program evaluation* (San Francisco: Jossey-Bass).

Fourcier, M.W. (1991). Substance abuse, crime and prison-based treatment: Problems and prospects. *Sociological Practice Review 2*(2), 123-131.

French M.T., Salome H.J., & Carney M. (2002). Using the DATCAP and ASI to estimate the costs and benefits of residential addiction treatment in the State of Washington. *Social Science and Medicine* 55:2267-2282.

Gaes, G.G. (1998) Correctional Treatment. In Tonry, M. (Ed.) *The Handbook of Crime and Punishment.* (New York, Oxford University Press.)

Geas, G.G., Flanagan, T.J., Motiuk, L.L., & Stewart, L. (1999). Adult Correctional Treatment. In Tonry, M., and Petersilia (Eds.), *Crime and Justice: A Review of Research Volume 26.* Chicago, University of Chicago Press.

Harwood, H.J., Napolitano, D.M., Kristianses, P., & Collins, J.J. (1984) . *Economic costs to society of alcohol and drug abuse and mental illness.* Research Triangle Park, NC: Research Triangle Institute.

Harwood, H.J., Fountain, D., & Livermore, G. (1998). *The economic costs of alcohol and drug abuse in the United States, 1992* (98-4327). Rockville: National Institute on Drug Abuse and National Institute on Alcohol Abuse and Alcoholism.

Holland, S. (1978). Gateway houses: Effectiveness of treatment on criminal behavior. *International Journal of the Addictions 13*, 369-381.

Hubbard, R.L, Marsden, M.E., Rachal, J.C., Harwood, H.J., Cavanaugh, E.R., & Ginzburg, H.M. (1989). Drug abuse treatment: *A national study of effectiveness.* Chapel Hill, NC: University of North Carolina Press.

Inciardi, J.A. (1993). *Criminal justice.* Fort Worth: Harcourt Brace.

Inciardi, J.A. (Ed) (1993b) SAGE *Criminal justice annuals, Volume 27: Drug treatment and criminal justice.* Newbury Park-SAGE. Inciardi, J.A., Lockwood, D. & Martin, S.S. (1991). Therapeutic communities in corrections and work release: Some clinical and policy considerations. Paper presented at National Institute on Drug Abuse Technical Review Meeting on Therapeutic Community Treatment Research, May 16-17, Bethesda, MD.

Inciardi, J.A., Lockwood, D., & Martin, S.S. (1994). Therapeutic communities in corrections and work release: Some clinical and policy implications. In F.M. Tims, G. De Leon, and N. Jainchill (Eds.), *Therapeutic community: Advances in research and application* (NIDA Research Monograph 144, pp. 259-267). Rockville, MD: National Institute on Drug Abuse.

Inciardi, J.A., Martin, S.S., Butzin, C.A., Hooper, R.M., & Harrison, L.D. (1997). An effective model of prison-based treatment for drug-involved offenders. *Journal of Drug Issues 2*, 261-278.

Inciardi, J.A., Martin, S.S., Butzin, C.A. (2003) Five-year outcomes of therapeutic community treatment of drug involved offenders after release from prison. *Crime and Delinquency*, in press.

Leukefeld, C.G., & Tims, F.M. (Eds.) (1988a). Compulsory treatment of drug abuse: *Research and clinical practice* (NIDA Research Monograph 86). Rockville, MD: National Institute on Drug Abuse.

Martin, S.S., Butzin, C.A., & Inciardi, J.A. (1995). Assessment of a multistage therapeutic community for drug involved offenders. *Journal of Psychoactive Drugs 27*, 109-116.

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 to aftercare. *The Prison Journal. 79*, 294-320.

Mathias, R. (1995). Community treatment helps offenders stay drug and alcohol free. *NIDA Notes* (May/June). Rockville, MD: National Institute on Drug Abuse.

Office of Applied Studies (1998). Substance abuse and mental health services administration, Uniform Facility Data Set (UFDS) survey, October 1. URL: http://dasis3.samhsa.gov/99ufds/resource-559/t38.htm.

Prendergast, M. L., 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 Rice, D.P., Kelman, S., Miller, L.S. & Dunmeyer, S. (1990). The economic costs of alcohol and drug abuse and mental illness: 1985. Report submitted to the Office of Financing and Coverage Policy of the Alcohol, Drug Abuse, and Mental Health Administration. U.S. Department of Health and Human Services. San Francisco: Institute for Health and Aging.

Rockholz, P.B. (2000). Findings of a national survey on therapeutic communities for substance abusing offenders in state prison. *Association of State Correctional Administrators Newsletter,* Middletown CT.

Rouse, J.J. (1991). Evaluation research on prison based drug treatment programs and some policy implications. *The International Journal of the Addictions 26*, 29-44.

Scarpitti, F.R., Inciardi, J.A., & Martin, S.S. (1994). Assertive community treatment: Obstacles to implementation. In B.W. Fletcher, J.A. Inciardi and A.M. Horton, Jr. (Eds.), *Drug abuse treatment: The implementation of innovative approaches* (pp. 63-78). Westport, CT: Greenwood Press.

Simpson, D.D., Wexler, H.K., & Inciardi, J.A. (1999) Drug treatment outcomes for correctional settings, Part 1; Introduction. *The Prison Journal 79.*

Smith, R. (1977). *Drug programs in correctional institutions.* Washington, DC: Law Enforcement Assistance Administration.

Stahler, G.J., Cohen, E., Shipley, T.E., & Bartelt D. (1993). Why clients drop out of treatment: Ethnographic perspectives on treatment attrition among homeless male 'crack' cocaine users. *Contemporary Drug Problems 20*, 651-680.

Wexler, H.K. (1995). The success of therapeutic communities for substance abusers in American prisons. *Journal of Psychoactive Drugs 27*, 57-66.

Wexler, H.K., De Leon, G., Thomas, G., Kressel, D., & Peters, J. (1999). The Amity prison TC evaluation—reincarceration outcomes. *Criminal Justice and Behavior 26*, 147-167.