LEAD Program Evaluation:

Recidivism Report

Susan E. Collins
Heather S. Lonczak
Seema L. Clifasefi

Harm Reduction Research and Treatment Lab
University of Washington – Harborview Medical Center

March 27, 2015

This report was prepared by the University of Washington LEAD Evaluation Team with important contributions from the LEAD Evaluation Advisory Committee and others acknowledged on the back page.

Executive Summary

- **Background:** This report was written by the University of Washington LEAD Evaluation Team at the request of the LEAD Policy Coordinating Group and fulfills the first of three LEAD evaluation aims.

- **Purpose:** This report describes findings from a quantitative analysis comparing outcomes for LEAD participants versus "system-as-usual" control participants on shorter- and longer-term changes on recidivism outcomes, including arrests (i.e., being taken into custody by legal authority) and criminal charges (i.e., filing of a criminal case in court). Arrests and criminal charges were chosen as the recidivism outcomes because they likely reflect individual behavior more than convictions, which are more heavily impacted by criminal justice system variables external to the individual.

- **Findings:** Analyses indicated statistically significant recidivism improvement for the LEAD group compared to the control group on some shorter- and longer-term outcomes.

  - **Shorter-term outcomes** were assessed for the six months prior and subsequent to participants' entry into the evaluation.
    - Compared to the control group, the LEAD group had 60% lower odds (likelihood) of arrest during the six months subsequent to evaluation entry. The effect of LEAD on getting arrested during the 6-month follow-up was statistically significant (p = .03).
    - This finding reflected the fact that—comparing the six months prior and subsequent to entry into the evaluation—the proportion of control participants who were arrested increased by 51%, whereas the proportion of LEAD participants who were arrested plateaued (+6%).
    - Inclusion of warrant-related arrests could either a) inflate apparent recidivism by reflecting nonappearance for prior violations or b) accurately represent new criminal activity that triggered prior warrants to be served even if there was no booking on a new crime. Thus, we examined the arrest data both with and without warrant arrests. Analyses of exclusively nonwarrant-related arrests indicated no significant LEAD effects.
    - Further, there were no statistically significant LEAD effects on total charges or felony charges filed over this shorter-term period.

  - **Longer-term outcomes** were assessed during the entirety of the LEAD evaluation time frame, ranging from October 2009 through July 2014. Analyses took into account the fact that participants had been in the program for differing amounts of time by statistically controlling for this factor.
    - Compared to the control group, the LEAD group had 58% lower odds of at least one arrest subsequent to evaluation entry. The LEAD effect on arrests over time was statistically significant (p = .001).
    - This finding reflected the fact that the proportion of control participants who were arrested at least once subsequent to evaluation entry increased by 4%, whereas the proportion of LEAD participants who were arrested subsequent to evaluation entry decreased by 30%.
    - Analyses indicated that, compared to control participants, LEAD participants had 34% lower odds of being arrested at least once when warrant-related arrests were removed. This effect was marginally significant (p = .09).
    - Although there was no statistically significant effect for total charges, the LEAD group had 39% lower odds of being charged with a felony subsequent to evaluation entry compared to the control group. This effect was statistically significant (p = .03).
    - The proportion of LEAD participants charged with at least one felony decreased by 52% subsequent to evaluation entry. The proportion of control group participants receiving felony charges decreased by 18%.

- **Interpretation of findings:** These statistically significant reductions in arrests and felony charges for LEAD participants compared to control participants indicated positive effects of the LEAD program on recidivism.

- **Next Steps:** This report is the second in a series that will be prepared by the University of Washington LEAD Evaluation Team over the next two years. The next report, which we plan to release in late spring of 2015, will describe our evaluation of the effectiveness of the LEAD program compared to the system-as-usual control group on criminal and legal systems utilization and associated costs. Later reports will evaluate changes among LEAD participants on psychosocial, housing and quality-of-life outcomes.

Introduction to the LEAD Program

Background and Rationale for the Law Enforcement Assisted Diversion (LEAD) Program

Despite policing efforts, drug users and dealers frequently cycle through the criminal justice system in what is sometimes referred to as a "revolving door."[1] The traditional approach of incarceration and prosecution has not helped to deter this recidivism.[2] On the contrary, this approach may contribute to the cycle by limiting opportunities to reenter the workforce, which relegates repeat offenders to continue to work in illegal markets.[3] This approach also creates obstacles to obtaining housing, benefits, and drug treatment. There have thus been calls for innovative programs to engage these individuals so they may exit the revolving door.[1]

Description of the LEAD Program

This need for innovative programs to prevent recidivism inspired the focus of the LEAD program, a collaborative pre-booking, community-based diversion program. The LEAD program was established in 2011 as a means of diverting those suspected of low-level drug and prostitution criminal activity to case management and other supportive services instead of jail and prosecution. The primary aim of the LEAD program is to reduce criminal recidivism.[a] Secondary aims include reductions in criminal justice service utilization and associated costs as well as improvements for psychosocial, housing and quality-of-life outcomes. Because LEAD is the first known pre-booking diversion program of its kind in the United States, an evaluation is critically needed to inform key stakeholders, policy makers, and other interested parties of its impact. The evaluation of the LEAD program described in this report represents a response to this need.

For the purpose of the evaluation, the implementation phase of this project occurred from October 2011 through July 2014. The Seattle Police Department's (SPD) officer shifts for squads making referrals to LEAD were randomly divided into 'red- and greenlight' shifts. Offenders who were encountered during greenlight shifts in the LEAD catchment area (i.e., Belltown neighborhood) were screened for project eligibility by officers on duty and, provided they met inclusion criteria and completed the intake process, they were diverted to the LEAD program at point of arrest instead of undergoing standard jail booking and criminal prosecution. A smaller number of individuals were referred by officers as 'social contacts.' Social contacts were individuals who were eligible for the LEAD program due to known recent criminal activity, but were recruited by officers outside of a criminal incident during a greenlight shift within the original LEAD catchment area. Both arrest and social contact referrals to LEAD

[a] Note: Because the LEAD program was launched as a pilot without sufficient resources to engage all possible participants within the planned catchment area, this evaluation did not focus on community- or neighborhood-level impact on crime. It is, however, possible that an approach that changed individual behavior, if later taken to scale with full commitment from all operational partners, would have neighborhood- or community-level impact.

required that participants were suspected of narcotics or prostitution activity and met other program criteria (see Purpose and Methods section below for inclusion criteria).

Interested individuals were referred to a LEAD case manager to complete an intake assessment. This assessment entailed items evaluating participants' substance-use frequency and treatment, time spent in housing, quality of life, psychological symptoms, interpersonal relationships, and health status. After completing the intake process, participants received case management through Evergreen Treatment Services' (ETS) REACH homeless outreach program, which connected participants with existing resources in the community (e.g., legal advocacy, job training or placement, housing assistance, counseling). Additionally, case managers had access to funds to provide financial support for the fulfillment of participants' basic needs (e.g., motel stays, housing, food, clothing, treatment, and various additional items and services). Other key program features included coordination of prosecution strategy in any other pending criminal cases participants had in local courts and legal assistance with miscellaneous civil legal problems. Six months following their entry into the LEAD program, participants completed additional one-on-one interviews with their case managers.

Eligible individuals who were arrested 1) during redlight shifts or 2) in non-LEAD neighborhoods—areas adjacent to Belltown that were not a part of the LEAD program but were patrolled by the same officers—were processed through the criminal justice system as usual (e.g., jail booking, criminal charges). These participants served as the control group in the current evaluation. Arrests in non-LEAD neighborhoods were included in the control group to increase the pool of participants while avoiding skewing the composition of the control group as the number of amenable, qualifying control participants available in the original catchment area decreased over time. All participants were recruited by the same officers using the same criteria.

Overall Program Evaluation Aims

The overall program evaluation will assess the LEAD program in meeting the following objectives compared to individuals who experienced the criminal justice system as usual.

- *Specific aim 1* is to test the relative effectiveness of the LEAD program compared to a 'system-as-usual' control condition in reducing criminal recidivism (i.e., arrests and charges) from the 6 months prior and subsequent to program entry, and as sufficient data accumulate, extending this analysis to evaluate longer-term outcomes.

- *Specific aim 2* is to test the effectiveness of the LEAD program compared to the 'system-as-usual' control condition in reducing publicly funded criminal justice service utilization and associated costs (i.e., court, prosecutor, public defense, jail) from the 6 months prior and subsequent to program entry. As sufficient data accumulate, this analysis will be repeated using longer-term outcomes.

- *Specific aim 3* is to test within-subjects differences on self-reported psychosocial and housing variables (i.e., alcohol and other drug use frequency; time spent in housing; quality of life; psychological symptoms; health status; and interpersonal relationships with family, partners and other community members).

Following a preliminary, within-subjects analysis that was released in September 2014, the current report reviews the complete set of findings from specific aim 1. Reports documenting findings for specific aims 2 and 3 will be released in late spring 2015 and fall 2015, respectively.

Purpose and Methods

Purpose

The purpose of this report is to describe and interpret findings from the quantitative evaluation of shorter- and longer-term recidivism outcomes (i.e., arrests and criminal charges) for evaluation participants who have been assigned to LEAD or the 'system-as-usual' control condition.

Participants

This quantitative evaluation included 318 adults who were suspected of low-level drug or prostitution offenses. Based on whether law enforcement contact was made during a red- or greenlight shift and whether it occurred in the LEAD catchment area, participants were either assigned to the LEAD (*n* = 203) or control (i.e., booking as usual; *n* = 115) conditions. At the time of referral, 146 of the LEAD participants were under arrest, and 57 were suspected of qualifying criminal activity but were referred outside of an alleged criminal incident.

All LEAD participants were those suspected of recent violations of the uniform controlled substances act (VUCSA) and/or prostitution offenses who were deemed eligible for the program by SPD officers. SPD considered individuals ineligible if they met any of the following criteria:

- The amount of drugs involved exceeded 3 grams, except where an individual was arrested for delivery of or possession with intent to deliver marijuana or possession, delivery or possession with intent to deliver prescription controlled substances (pills).
- The individual did not appear amenable to diversion.
- The suspected drug activity involved delivery or possession with intent to deliver (PWI), and there was reason to believe the suspect was dealing for profit above a subsistence income.
- The individual appeared to exploit minors or others in a drug dealing enterprise.
- The individual was suspected of promoting prostitution.
- The individual had a disqualifying criminal history as follows:
    - Without time limitation: Any conviction for murder 1 or 2, arson 1 or 2, robbery 1, assault 1, kidnapping, Violation of the Uniform Firearms Act (VUFA) 1, any sex offense, or attempt of any of these crimes.
    - Within the past 10 years: Any conviction for a domestic violence offense, robbery 2, assault 2 or 3, burglary 1 or 2, or VUFA 2.
    - The individual was already involved in King County Drug Diversion Court or Mental Health Court. This exclusion criterion served to ensure the

LEAD program was not combined with other models of intervention and case management.

The control group included only individuals arrested by LEAD-referring officers who would have been considered eligible for referral to LEAD had the arrest occurred during a greenlight shift in a LEAD catchment area. Individuals who would not have met LEAD referral criteria were not included in the control group. There was no penalty to officers for excluding individuals from the evaluation based on the inclusion/exclusion criteria. Officers completed forms for each arrest documenting these decisions.

Measures

The evaluation team obtained all necessary IRB exemptions and data sharing agreements from the appropriate entities. Next, with the assistance and guidance of the LEAD Policy Coordinating Group and the LEAD Evaluation Advisory Committee, the evaluation team obtained demographic and program data from the LEAD case management team and from the SPD LEAD records. Data on criminal recidivism (i.e., arrests, charges) were extracted by the King County Prosecuting Attorney's office from the FBI's National Crime Information Center (NCIC) and were given to the evaluation team for analysis. For the purpose of this evaluation, new arrests refer to having been taken into police custody for a crime committed during the LEAD program evaluation time frame (i.e., 10/1/2009 through 7/31/2014). New arrests did not include parole or probation violations or failure to comply offenses pursuant to prior violations, which were removed for these analyses (5.1%; n = 188). New charges were criminal charges—including felonies—that occurred during the LEAD evaluation time frame noted above. During their intake interviews, LEAD participants signed consent forms allowing the release of their administrative data.

Data Analysis Plan

Overview. The goal of this evaluation was to test LEAD effects on recidivism outcomes (i.e., arrests and charges) over both the shorter term (i.e., six months prior and subsequent to program involvement) and the longer term (i.e., encompassing two years prior to the LEAD start date through 7/31/14). This two-tiered data analysis plan was used to assess both shorter- and longer-term LEAD effects. Given their relative statistical rarity, recidivism counts were converted to dichotomous (yes/no) outcomes, excluding any arrest that occurred the day participants entered the evaluation. Dichotomizing recidivism outcomes is standard in analyzing effects of criminal justice programs in Washington State.[4] Because longer-term analyses involved unequal windows of time for participants starting at different points during the program implementation, we statistically controlled for this factor in each of the longer-term models.

Types of arrest included*.* The primary goal of these analyses was to assess changes in recidivism (i.e., new law violations) within the evaluation time frame. We therefore excluded arrests due to prior violations as noted above. Warrant arrests pursuant to incidents occurring after study entry, however, were considered differently because their inclusion could work in two different ways. On the one hand, arrest of control participants due to warrants from the arrest on the would-be LEAD referral date could have a reverberating effect that would overstate new criminal involvement. On the other hand, warrant arrests could reflect new criminal activity that triggered warrants to be served without an arrest for a new offense. Because it is unclear whether warrant arrests are independent of new criminal activity, we conducted two sets of arrest analyses—one including and one excluding warrant arrests—to allow us to understand the range of the possible LEAD effects.

Group allocation. Randomized controlled trials represent the gold standard in evaluation. A cluster randomization schema[5] was originally proposed for the LEAD evaluation, such that individuals arrested during specified greenlight shifts in the original catchment area would be randomized to receive LEAD, and individuals arrested during redlight shifts in the original catchment area would be randomized to the system-as-usual control condition.

LEAD, however, was implemented in a real-world setting. Thus, changes to the originally proposed evaluation design were made to ensure LEAD's success on the ground. First, having a pathway for social contacts (i.e., individuals who were encountered on a greenlight shift within the original catchment area, were suspected by officers of recent drug or prostitution activity, had been arrested for these offenses in the past, and met the same inclusion criteria) to enter into the LEAD program was deemed necessary from a policy and policing standpoint. Because they were all subject to the same inclusion criteria, LEAD participants recruited via social contacts and arrest diversion were very likely drawn from the same population (see analyses comparing these groups below). Second, after the evaluation began, operational partners recognized that there was a limited number of potential participants in the originally planned catchment area. Over time, most of these individuals were approached for program involvement leaving a dwindling number of individuals available for the comparison group. Thus, to accommodate the need for an adequate and comparable control group, redlight areas (in addition to redlight shifts) were added to the evaluation. This ensured adequate representation of amenable and qualifying participants in the control condition to make up for the initial catchment area's relatively small population.

After careful consideration, a nonrandomized controlled design was employed for the evaluation of LEAD to accommodate these deliberate and important program implementation features. According to federal standards, nonrandomized controlled designs are consistent with the early intervention development and evaluation exemplified by the LEAD program.[6] Further, high-quality nonrandomized controlled evaluations that account for potential confounds show similar effect sizes and widely correspond to outcomes of randomized controlled trials.[7] In fact, the current University of Washington evaluation team used a nonrandomized controlled design in a prior, well-regarded evaluation of the 1811 Eastlake Housing First program in Seattle.[8-14] In that evaluation, it was decided that real-world considerations would contraindicate a randomized controlled design, because it was deemed impractical and unethical to withhold essential social services (i.e., housing) from individuals in the community.[13]

Despite its appropriateness for the current evaluation, a nonrandomized controlled design can result in intervention and control group imbalances and statistical biases (e.g., selection bias).[15,16] We therefore employed both methodological and statistical approaches to avoid these problems. First, LEAD officers received focused instructions and training to ensure participants recruited to all groups were representative of the same population. Second, all control and LEAD participants had to meet the same set of inclusion criteria. The fulfilment of these criteria was systematically documented in participant files. Third, the same officers were involved in recruitment of both LEAD and control participants. Finally, we employed a statistical approach called propensity score weighting to balance the intervention and control groups, which increases confidence in the causal impact of the intervention effect.[16]

Propensity score weights. We used generalized boosted regression to estimate propensity scores for all eligible participants (N = 318). This type of regression employs an automated, data-adaptive algorithm that fits several models by way of a regression tree and then merges the predictions of these various models. The advantage of generalized boosted regression is that it is computationally fast to fit; handles various types of data distributions; and takes into account interaction terms. In addition, it is invariant to one-to-one transformations of the independent variables; thus, the raw, log, and exponentiated variants lead to the same propensity score adjustments.[17]

Next, we created two weighting variables: one for estimating the average treatment effect (ATE) and one for estimating the average treatment effect for treated participants (ATT).[16] ATE may be considered to be a between-subjects' difference or the average effect of moving an untreated population to a treated population.[18] Alternatively, treatment effects may be considered at the individual or within-subjects level. The ATT may be considered to be the average effect of treatment for those who receive the treatment—in this case LEAD.[18] Both types of propensity scores are relevant for the current analysis because, if considered effective, LEAD a) would be applied widely to the larger population of drug and sex work offenders (reflected in ATE) and b) is a highly tailored, individual-level intervention whose effects on treated participants, which are reflected in ATT effects, would be important to track as well. Both propensity score weights were thus used in analyses and reported on in the results section.

Propensity score analyses comprised three steps. First, we generated the propensity scores using generalized boosted regression. Where p is the propensity score, the ATE is $1/p$ for LEAD participants and $1/(1-p)$ for control participants. ATT is equal to 1 for treated participants, and $p/(1-p)$ for control participants. Second, we used ATE and ATT weights to conduct balance checks, which comprised a series of ordinary least squares, logistic and multinomial logistic regressions testing whether propensity scores improved the balance between the control and LEAD groups. Finally, we used the ATT and ATE as sampling weights in the primary analyses.

Primary analyses. Using SPSS 19 and Stata 13, descriptive analyses were conducted to describe the sample. Population-averaged generalized estimating equations (GEEs)[19] were used in primary analyses. GEEs model marginal effects and may be used to accommodate alternative distributions (e.g., binomial) and correlated data (e.g., data collected on the same participant over time). In this evaluation, GEEs were used to test the relative effects on recidivism outcomes of: a) *time* (0=baseline, 1=follow-up), which controlled for overall, longitudinal effects that could reflect regression to the mean; b) *intervention group* (0=control, 1=LEAD); and c) the two-way *time x intervention group* interaction. The interaction shows the effect of the LEAD intervention on longitudinal recidivism outcomes. Additionally, we controlled for time in the evaluation as a time-varying covariate (i.e., years prior and subsequent to evaluation entry).

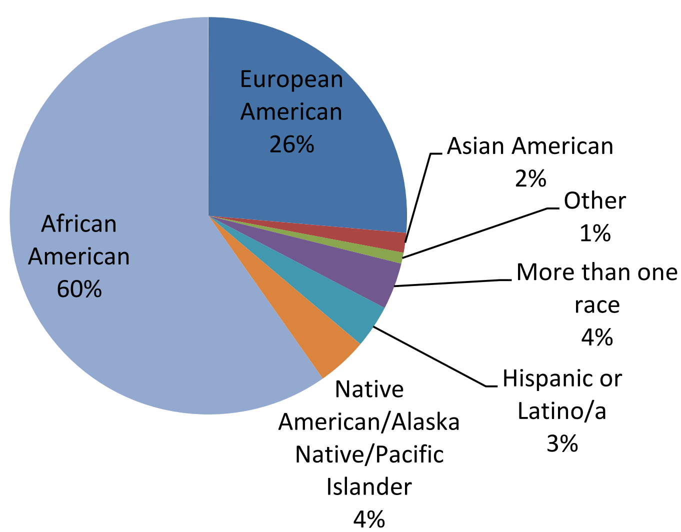
Because recidivism outcomes were dichotomous, we specified Bernoulli distributions with the logit link. We assumed an exchangeable correlation structure to accommodate repeated measures on one individual, which served as the sole clustering variable.[20] To enhance model interpretability, resulting effect sizes were exponentiated and reported as odds ratios (*OR*s), where *OR*s < 1 indicate an inverse association, *OR*s = 1 indicate no association, and *OR*s > 1 indicate a positive association. Alphas were set to $p = .05$, indicating statistically significant results, and $p = .10$, indicating marginally significant results. Confidence intervals were set to 95%.

Results

Overall Sample Description

Participants in this evaluation (*N* = 318) had an average age of 40.17 (*SD* = 11.85) years and were predominantly male (34.28% female; *n* = 109). The racial and ethnic diversity of the overall sample is shown in Figure 1.

Figure 1. Ethnic/racial backgrounds of participants



In the six months prior to evaluation entry, participants had accrued a total of 206 arrests and 151 charges, of which 17% (*n* = 26) were felony charges. Expanding out to all incidents since the start of the evaluation time frame (10/1/09) through the current evaluation window (7/31/14), participants had accrued 1,415 arrests and 994 charges, of which 21% (*n* = 213) were felony charges.

Group Differences at Baseline

Arrest diversion versus social contact participants who received LEAD. Of the baseline demographic and recidivism (i.e., criminal history) variables (including prior criminal history), participant age was the only variable that evinced a statistically significant difference between the arrest diversion (*M* = 40.35, *SD* = 11.09) and social contact (*M* = 45.24, *SD* = 10.65) groups (*p* = .006; other *ps* > .12). Given the lack of observed differences and the fact the two groups were recruited using the same inclusion criteria by the same officers, it was concluded that these two groups were very likely drawn from the same population. The arrest diversion and social contact groups were therefore collapsed and analyzed as a single LEAD group.

LEAD versus control group. Wilcoxon rank-sum and Pearson chi-square tests indicated significant group differences on demographic variables at baseline (see Table 1 for descriptive statistics) between LEAD and control participants. Further, 11 participants died during the 5-year evaluation, including 9 LEAD participants (4.43%) and 2 (1.74%) control participants. This group difference was not statistically significant, $X^2(1, N = 318) = 1.60$, $p = .21$. It should be noted that LEAD participants' deaths were systematically documented, whereas control participants' deaths were not. These individuals were included in all analyses, and death was used in propensity scores and subsequent weighted analyses. There were no significant group differences on baseline recidivism (i.e., criminal history) (*p*s > .09).

Table 1. Baseline demographic and participation data by group

Demographic Variables	LEAD Group *n* = 203 Mean(*SD*)/%(*n*)	Control Group *n* = 115 Mean(*SD*)/%(*n*)	z/X^2	*p*-value
Age	41.72 (11.16)	37.44 (12.57)	-3.03	.003
Gender	39% (79) female	26% (30) female	5.36	.021
Race/ethnicity			19.43	.003
American Indian/Alaska Native/Pacific Islander	6% (13)	0% (0)		
Asian American	<1% (1)	3% (4)		
Black/ African American	55% (112)	68% (78)		
European American	27% (55)	25% (29)		
Hispanic/Latino/a	5% (10)	1% (1)		
More than one race	4% (9)	3% (3)		
Other	1% (3)	0% (0)		
Death	4% (9)	2% (2)	1.60	.21
Overall years in evaluation	1.54 (.63)	1.78 (.52)	3.66	<.001

Note: Percentages may not total 100% due to rounding.

Pre- and Postevaluation Descriptive Statistics of Recidivism Outcomes by Group

Descriptive statistics for raw, unadjusted recidivism outcomes were calculated for LEAD and control groups prior and subsequent to entry into the evaluation (see Table 2).

Table 2. Recidivism outcome measures by group

Recidivism measures	**LEAD participants** *Mean* (*SD*)		**Control participants** *Mean* (*SD*)	
	Pre	Post	Pre	Post
Shorter-term (6 mo) measures				
Arrests	.55(.94)	.68(1.28)	.82(1.37)	1.04(1.24)
Nonwarrant arrests	.33(.71)	.48(.93)	.48(.91)	.59(1.03)
Total charges	.44(1.12)	.45(.93)	.53(1.09)	.59(1.36)
Felony charges	.07(.28)	.13(.45)	.10(.32)	.18(.54)
Longer-term measures				
Arrests/year	1.42(1.49)	1.11(1.69)	1.39(1.70)	1.71(1.75)
Nonwarrant arrests/year	.81(.93)	.86(1.42)	.86(1.14)	1.03(1.46)
Total charges/year	.99(1.52)	.73(1.31)	.95(1.25)	1.01(1.47)
Felony charges/year	.21(.35)	.20(.61)	.22(.33)	.27(.50)

Note: This table features raw values. Because recidivism outcomes were statistically rare events, however, these were dichotomized for primary outcomes.

Propensity Score Balance Check

We conducted a check of the group balance after the ATE and ATT weights were applied. Table 3 below shows the balance check results. Nonsignificant values indicate propensity scores successfully balanced the LEAD and control groups for these variables. Findings indicated that both ATE and ATT performed moderately well in balancing the groups; thus, we report findings for both ATE and ATT in this report.

Table 3. Group balance check following application of propensity score weights

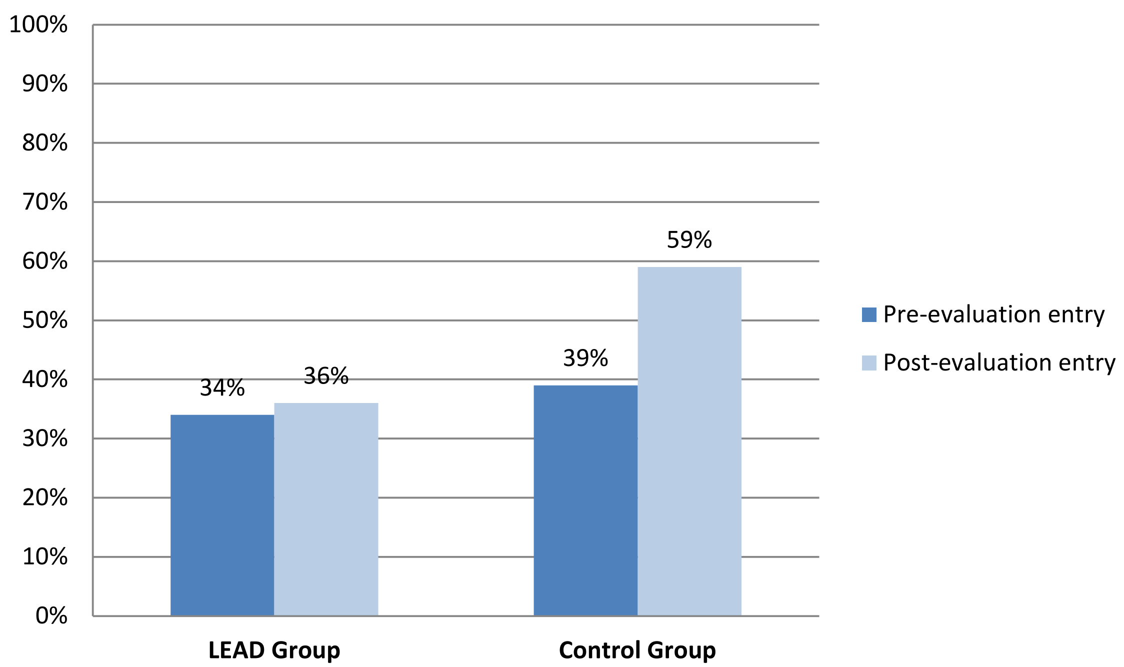
Covariates	**Significance level of treatment imbalance (*p*-value)**	
	ATE	ATT
Age	.03*	.11
Gender	.07	.13
Race/ethnicity (dummy group: European American)		
African American	.31	.37
Other race/ethnicity	.07	.05
Died	.21	.20
Overall years in evaluation	.002*	.003*
Total arrests prior to evaluation entry	.66	.37

Note: * $p < .05$. See Tables 1, 3 for mean values for the imbalanced variables prior to propensity score generation.

Primary Analyses

Shorter-term recidivism analyses. The average treatment effect (ATE) model, which tested overall group effects, was significant, Wald $X^2(3, N = 318) = 19.18$, $p < .001$. The ATE indicated that, compared to control participants, LEAD participants had 60% lower odds of having at least one arrest subsequent to program entry. Specifically, the time x intervention group interaction effect was significant indicating a LEAD effect over time ($OR = .49$, robust $SE = .16$, $p < .03$). The ATT model, which indicated the treatment effect for LEAD participants alone, was also significant, Wald $X^2(3, N = 318) = 16.10$, $p = .001$. The time x intervention group interaction was likewise significant ($OR = .50$, robust $SE = .17$, $p = .04$), and indicated 57% lower odds of arrest subsequent to LEAD involvement. See Figure 2 below for the percentage of participants arrested in each group both six months prior and subsequent to evaluation entry. See Appendix A for full output and Appendix B for effect size calculations reported in this Primary Analysis section.

Figure 2. Percent of participants arrested 6 months prior and subsequent to evaluation entry



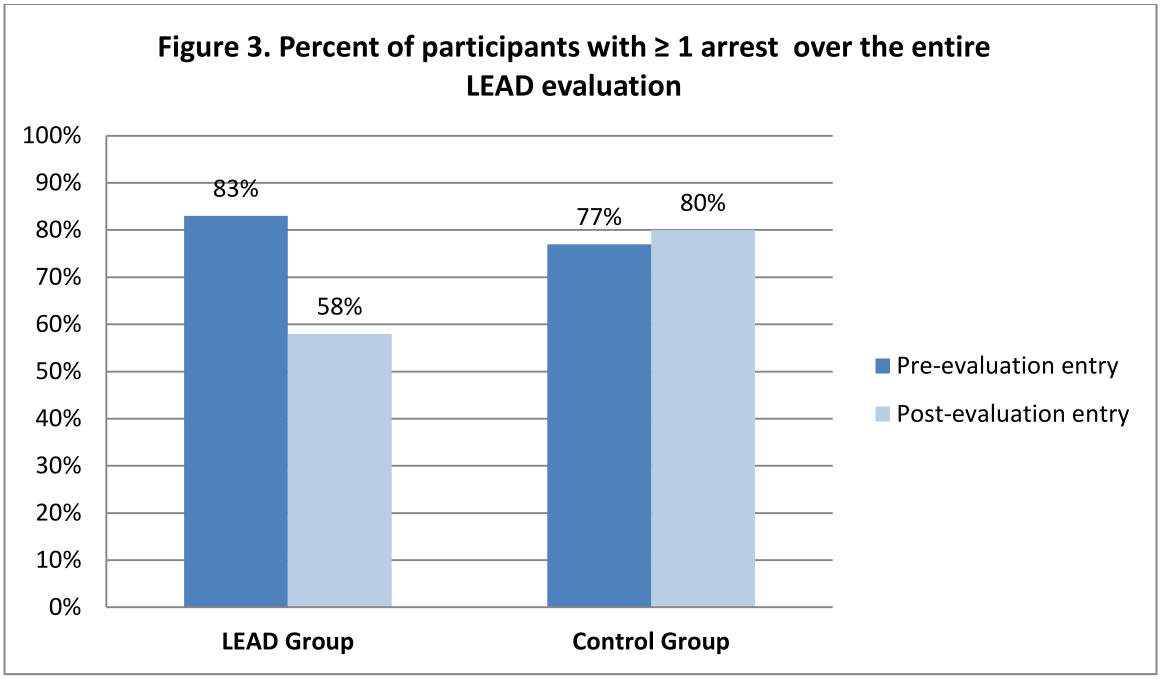
When we considered only nonwarrant arrests, however, these group differences were no longer statistically significant (model *p*s > .11; see Table 4). Further, there were no statistically significant differences between the LEAD and control groups on total charges or felony charges for the 6-month analyses (model *p*s > .28). See Table 4 for percentage of participants with arrests, total charges and felony charges both six months prior and subsequent to evaluation entry.

Table 4. Short-term changes in recidivism (6 months pre- to 6 months postevaluation entry)

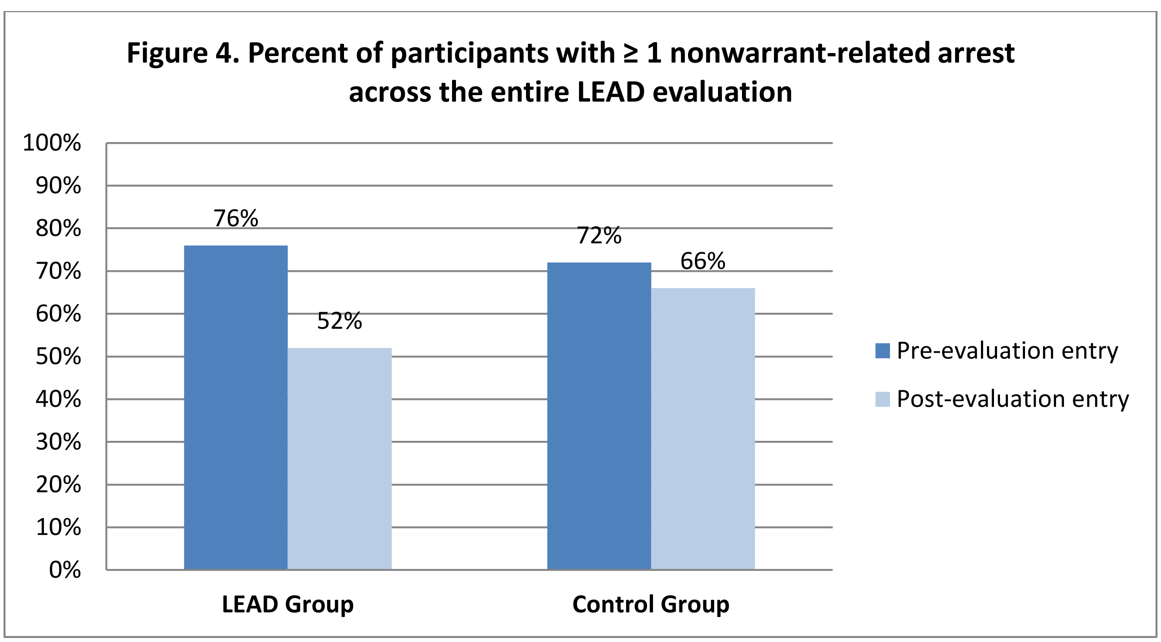
Recidivism measures	LEAD participants		Control participants	
	Pre	Post	Pre	Post
≥ one arrest*	34%	36%	39%	59%
≥ one nonwarrant arrest	24%	30%	29%	37%
≥ one charge	23%	28%	31%	26%
≥ one felony charge	7%	10%	9%	14%

Note: These values are unadjusted. * = significant group difference favoring the LEAD group ($p < .05$). Other group differences were not statistically significant.

Longer-term recidivism analyses. After evaluating short-term LEAD outcomes, we expanded the evaluation time frame to encompass two years prior to the initial LEAD program start date (10/1/2009) to our evaluation close date (7/31/2014). The average treatment effect (ATE) model, which tested overall group effects, was significant, Wald $X^2(4, N = 318) = 55.09$, $p < .001$. The time x intervention group interaction showed a significant LEAD effect over time ($OR = .30$, robust $SE = .11$, $p = .001$). This finding indicated that, compared to control participants, LEAD participants had 58% lower odds of being arrested at least once subsequent to program entry. The ATT model, which indicated the treatment effect for the LEAD participants alone, was significant, Wald $X^2(4, N = 318) = 53.66$, $p < .001$. Results indicated 56% lower odds of being arrested at least once subsequent to LEAD involvement, which was reflected in the significant time x intervention group interaction effect ($OR = .29$, robust $SE = .11$, $p = .001$). See Figure 3 for the percentage of participants arrested at least once in each group prior and subsequent to evaluation entry.

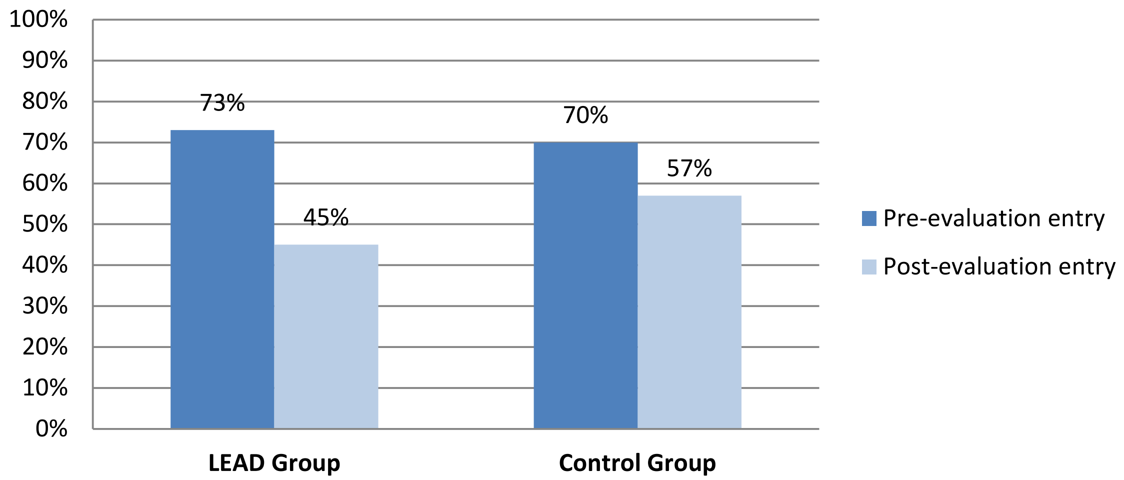


After warrant arrests were removed, the ATE, Wald $X^2(4, N = 317) = 42.16$, $p < .001$, and ATT, Wald $X^2(4, N = 317) = 42.26$, $p < .001$, models were significant. The ATE model indicated that the odds of at least one nonwarrant-related arrest among LEAD participants were 34% lower than those of control participants. The ATE interaction effect was marginally statistically significant ($OR = .58$, robust $SE = .18$, $p = .09$); however, the ATT interaction effect was not ($p = .11$). See Figure 4 for percentage of participants who were arrested for nonwarrant-related reasons.



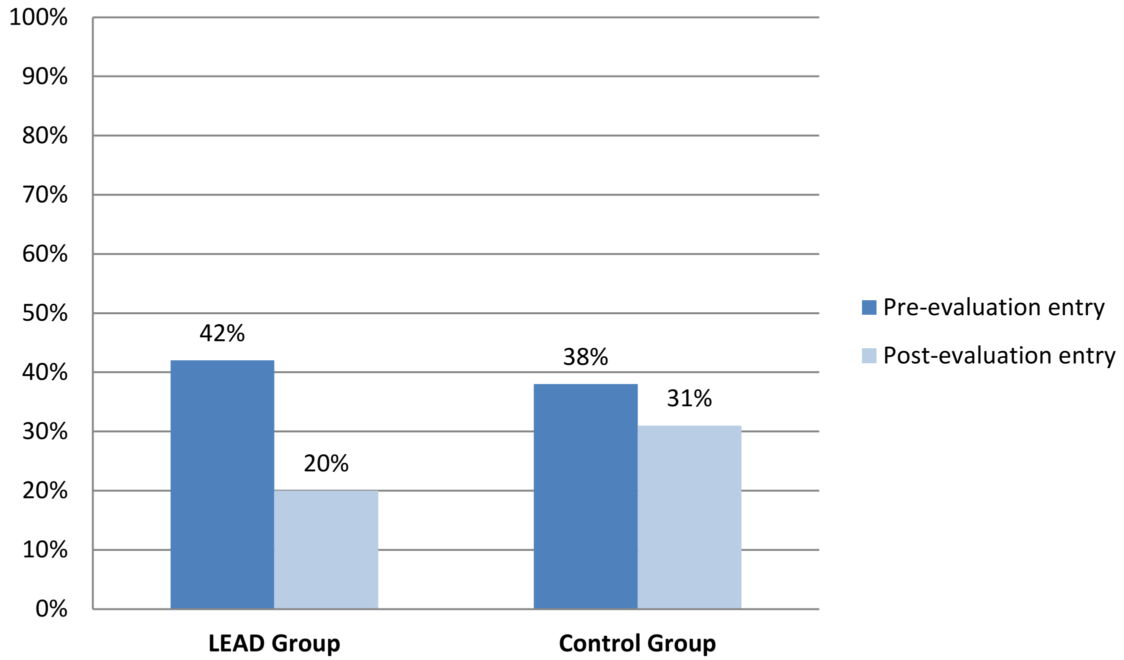
Criminal charge models were statistically significant (*p*s < .001); however, the time x intervention group interactions were not (*p*s > .18). That said, descriptive statistics indicated that the group differences were in the desired direction (see Figure 5).

Figure 5. Percent of participants charged at least once across the entire LEAD evaluation



When we considered group differences for felony charges, the ATE model was significant, Wald $X^2(4, N = 318) = 33.47$, $p < .001$. The time x intervention group interaction effect indicated a significant LEAD effect over time (*OR* = .49, robust *SE* = .16, *p* = .03). This finding indicated that, compared to control participants, LEAD participants had 39% lower odds of being charged with at least one felony subsequent to program entry. The ATT model, which indicated the treatment effect for the LEAD participants specifically, was significant, Wald $X^2(4, N = 318) = 34.85$, $p < .001$. Results indicated 36% lower odds of being charged with a felony subsequent to LEAD involvement, and this was reflected in a significant time x intervention group interaction (*OR* = .47, robust *SE* = .16, *p* = .02). See Figure 6 below for the percentage of participants charged with at least one felony in each group prior and subsequent to evaluation entry.

Figure 6. Percent of participants charged with at least one felony across the entire LEAD evaluation



Discussion

The LEAD program is reaching a diverse population that has experienced the street-to-jail-to-street revolving door. Findings indicated that LEAD is associated with positive effects for some shorter- and longer-term recidivism outcomes.

Arrest Outcomes

When looking at shorter-term, six-month arrest outcomes, there was a significant LEAD effect, which reflected the fact that the number of LEAD participants being arrested leveled off, whereas the number of control participants arrested increased. This shorter-term effect for arrests did not hold when warrant arrests were removed. Over the longer term, however, these effects were more pronounced. When the time frame was expanded to include recidivism since the start of data collection (10/1/09) until last summer (7/31/14), significantly fewer LEAD participants were arrested after they started LEAD, and there was a marginally significant effect for nonwarrant-related arrests, compared to control participants.

Taken together, arrest findings indicate positive LEAD effects on recidivism that are likely due to features of the LEAD program. All LEAD participants receive case management, which supports fulfilment of basic needs, including housing stability, job attainment and enrollment in drug and alcohol treatment. Further, LEAD participants' case managers coordinate with prosecutors to ensure nondiverted cases are managed to support and not compromise LEAD intervention plans.

It is, however, important to discuss other potential explanations for these findings. First, increases in the control group's odds of arrest following evaluation entry across all analyses are worth discussing. It is important to bear in mind that the Seattle West Precinct was subject to policy changes during the LEAD evaluation time period, which could have affected both the LEAD and control groups' rates of arrest. It is therefore possible that more focused enforcement—and not necessarily increased criminal activity—was responsible for increases in the prevalence of arrests in the control group. These larger, systemic changes, however, would not account for the LEAD group's drop in arrest prevalence, which would have been expected to reflect the same environmental conditions as the control group.

Another potential explanation for these findings is that officers could have made intentional decisions to avoid arresting LEAD participants. Upon further consideration, however, this explanation is not highly probable. Only approximately 40 of 1,300 SPD officers were involved in the LEAD program. Further, few—if any—officers outside of the LEAD squads were aware of individuals' group assignment. There were neither department-wide communications/trainings about the program nor system flags visible to officers that would signal LEAD participation. Thus, we are confident the observed LEAD effect in reducing arrest is not primarily due to intentional differences in decision-making by SPD officers.

Charge Outcomes

Over the 6-month follow-up, LEAD participants did not show statistically significant differences in odds of being charged with a crime or being charged with a felony crime. When considered over the longer term, however, LEAD participants had significantly lower odds of being charged with a felony.

It should be noted that felonies were included for completeness in considering differentiated indices of recidivism. In contrast to arrests, however, this indicator could have been affected by the decisions of LEAD stakeholders, particularly the Trial Unit Chief for the King County Prosecutor. As an unblinded operational partner, the prosecutor's office could take into account LEAD participation and progress in the program when deciding whether and when to file felony charges. Thus, the lower odds of felony charges among LEAD participants compared to control participants could have been precipitated by differential decision-making in the prosecutor’s office. As charges may be less purely indicative of changes in recidivism than arrest prevalence, these findings will likely play a more important role in the system utilization analysis that will be addressed in the next report.

Understanding These Findings in the Context of Existing Evaluations

The present findings are particularly meaningful when placed in the context of the existing literature on interventions targeting recidivism. For example, nationwide meta-analyses and systematic reviews have shown that some programs targeting recidivism, including mental health court, drug court and tailored psychosocial interventions, are superior to mainstream criminal justice processing across various outcomes.[21-23] Closer to home, a recent Washington State Institute for Public Policy (WSIPP) evaluation found that existing evidence- and research-based approaches focusing on tailoring supervision to offender’s relative risk level, motivation and needs had a small but significant collective effect (d = -.23) and reduced recidivism by about 14 percentage points compared to traditional supervision.[24] It is notable that the current evaluation indicated LEAD had an even larger effect size (d = -.33) and reduced recidivism by about 22 percentage points compared to the system as usual, which, in King County where this evaluation was conducted, includes various therapeutic courts. This evaluation therefore provides compelling support for LEAD—an innovative approach to reducing criminal recidivism—as a viable alternative to existing criminal justice system approaches.

Limitations

This evaluation’s limitations should be noted. First, large administrative datasets often feature missing data and clerical errors. That being said, we have no reason to believe such errors asymmetrically affected LEAD participants versus control participants.

Second, given real-world implementation realities, the originally planned randomization schema was relaxed, and a nonrandomized controlled design was employed in its place. To increase confidence in the causal impact of LEAD versus the system-as-usual control condition, both methodological and statistical approaches were used to balance the control and LEAD groups. For example, LEAD officers were trained on the application of the inclusion/exclusion criteria, and they made a systematic effort to identify qualifying LEAD, control and social contact participants using the same criteria. Further, there was no penalty to officers for excluding individuals from the evaluation based on the inclusion/exclusion criteria. LEAD squads were also consistent over the course of the evaluation for both control and LEAD groups; thus, the same officers were responsible for assessing all participants' inclusion/exclusion criteria over the course of the evaluation. Finally, we reduced the influence of potential selection bias using propensity score weighting, which is a statistical technique designed to ensure greater balance across groups and thereby decrease bias due to potentially confounding variables. The propensity scores balanced the groups on variables aside from years included in the evaluation. Thus, we controlled for this factor separately in primary outcome analyses.

Third, descriptive sample analyses indicated some significant baseline differences between LEAD and control groups. Specifically, the LEAD group comprised more older, female participants. However, since the groups were comparable in terms of recent criminal history, this difference does not seem likely to account for differences in post-entry recidivism. It is also worth noting that there was a higher proportion of African Americans in the control condition. Past arrest data suggest that drug arrests in the south end of the West Precinct were more likely to involve African-Americans than those in the Belltown neighborhood. The south end was, however, not included in the LEAD catchment area, and these participants were instead included in the control condition. Thus, the observed imbalance is more likely due to preexisting factors rather than officer behavior. Fortunately, this as well as all other baseline group demographic differences—accept the ATE for age--were successfully balanced by the propensity scores.

Conclusions and Future Directions

Findings indicated positive effects of the LEAD program on reducing criminal recidivism over shorter six-month and longer evaluation-wide time frames. Specifically, the odds of arrests and felony charges were lower among LEAD versus control participants. The limitations of the current evaluation were ameliorated using both methodological and statistical approaches, which increased our confidence that the LEAD effects were due to the program itself and not other potentially confounding factors.

This report represents the second in a series that are being prepared by the University of Washington LEAD Evaluation Team over the next two years. The next report, which we plan to release in late spring of 2015, will describe our evaluation of the effectiveness of the LEAD program compared to the system-as-usual control group on criminal and legal systems utilization and associated costs. Later reports will evaluate changes among LEAD participants on psychosocial, housing and quality-of-life outcomes.

References

1. Warner TD, Kramer JH. Closing the revolving door? Substance abuse treatment as an alternative to traditional sentencing for drug-dependent offenders. *Criminal Justic and Behavior.* 2009;36:89-109.
2. Wormith JSO, M. Offender treatment and attrition and its relationship with risk, responsivity and recidivism. *Criminal Justice and Behavior.* 2002;29:447-471.
3. Fletcher DR. Offenders in the post-industrial labour market: Lubricating the revolving door? *People, Place and Policy.* 2013;1:80-89.
4. Aos S, Miller M, Drake E. *Evidence-based public policy options to reduce future prison construction, criminal justice costs and crime rates.* Olympia, WA: Washington State Institute for Public Policy; 2006.
5. Donner A, Klar N. *Design and analysis of cluster randomization trials in health research.* London: Arnold; 2000.
6. Rounsaville BJ, Carroll KM, Onken LS. A stage model of behavioural therapies research: Getting started and moving on from Stage I. *Clinical Psychology: Science and Practice.* 2001;8:133-142.
7. Benson K, Hartz AJ. A comparison of observational studies and randomized, clinical trials. *The New England Journal of Medicine.* 2000;342:1878-1886.
8. Clifasefi SL, Malone D, Collins SE. Associations between criminal history, housing first exposure and jail outcomes among chronically homeless individuals with alcohol problems. *International Journal of Drug Policy.* 2013;24:291-296.
9. Collins SE, Malone DK, Clifasefi SL. Housing retention in single-site Housing First for chronically homeless individuals with severe alcohol problems. *American Journal of Public Health.* 2013;103:S269-S274.
10. Collins SE, Malone DK, Clifasefi SL, et al. Project-based Housing First for chronically homeless individuals with alcohol problems: Within-subjects analyses of two-year alcohol-use trajectories. *American Journal of Public Health.* 2012;102:511-519.
11. Collins SE, Malone DK, Larimer ME. Motivation to change and treatment attendance as predictors of alcohol-use outcomes among project-based housing first residents. *Addictive Behaviors.* 2012;37:931-939.
12. Conner M, Warren R, Close S, Sparks P. Alcohol consumption and the theory of planned behavior: An examination of the cognitive mediation of past behavior. *Journal of Applied Social Psychology.* 1999;29:1676-1704.
13. Larimer ME, Malone DK, Garner MD, et al. Health Care and Public Service Use and Costs Before and After Provision of Housing for Chronically Homeless Persons With Severe Alcohol Problems. *Journal of the American Medical Association.* 2009;301:1349-1357.
14. Mackelprang JL, Collins SE, Clifasefi SL. Housing First is associated with reduced use of emergency medical services. *Prehospital Emergency Care.* 2014.
15. Kunz R, Vist G, Oxman AD. Randomisation to protect against selection bias in healthcare trials (Review). *Cochrane Database Systematic Review.* 2007;18:PMID: 17443633
16. Guo SY, Fraser MW. *Propensity score analysis: Statistical methods and applications 2nd Edition.* Los Angeles: SAGE Publications, Inc.; 2015.

17. McCaffrey DF, Ridgeway G, Morral AR. Propensity score estimation with boosted regression for evaluating causal effects in observational studies. *Psychological Methods.* 2004;9:403-425.
18. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research.* 2011;46:399-424.
19. Zeger SL, Liang K-Y. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics.* 1986;42:121-130.
20. Hardin JW, Hilbe JM. *Generalized estimating equations.* Boca Raton, FL: Chapman & Hall/CRC; 2003.
21. Scott DA, McGilloway S, Dempster M, Browne F, Donnelly M. Effectiveness of criminal justice liaison and diversion services for offenders with mental disorders: A review. *Psychiatric Services.* 2013;64:843-849.
22. Brown RT. Systematic review of the impact of adult drug-treatment courts. *Translational Research.* 2010;155:263-274.
23. Perry A, Coulton S, Glanville J, et al. Interventions for drug-using offenders in the courts, secure establishments and the community. *The Cochrane Library.* 2006;3:Art. No.: CD005193.
24. Drake E. *Inventory of evidence-based and research-based programs for adult corrections (Document No. 13-12-1901).* Olympia, WA: Washington State Institute for Public Policy; 2013.

ACKNOWLEDGMENTS

The authorship team for this report included Drs. Susan E. Collins, Heather S. Lonczak, and Seema L. Clifasefi. We would like to thank the current and former members of the LEAD Evaluation Advisory Committee, including Mark Baird, Mary Barbosa, Mark Cooke, Clifton Curry, Lisa Daugaard, Ian Goodhew, Ron Jackson, Jutta Joesch, Anita Khandelwal, Kris Nyrop, Christa Valles and Mike West, for their valuable contributions to this report. We would especially like to thank the King County Prosecutor's office for obtaining the administrative data for this report as well as the SPD Narcotics Unit and the REACH team for their help in obtaining the LEAD program data. We would like to thank Cynthia Lum, PhD, for her helpful comments on initial drafts. Finally, we thank our program staff, Sara Hoang, Gail Hoffman and Emily Taylor, for their additional administrative and data management contributions.

This work is supported by a grant from the Laura and John Arnold Foundation to Susan E. Collins and Seema L. Clifasefi.

LEAD services and project management have been funded by the Ford Foundation, the Open Society Foundations, the RiverStyx Foundation, the Vital Projects Fund, the Massena Foundation and the City of Seattle.

APPENDICES

Appendix A. Primary outcome analysis output
Appendix B. Effect size calculations for interpretation of the interaction effect for the LEAD group

APPENDIX A. Primary outcome analysis output

Key for abbreviations used in this output

xtgee	darrest6_	t TxGroup txTx	[pweight=ATE]	i(id)	t(time)	family(bin)	link(logit)	corr(exch)	eform	robust
analysis type: generalized estimating equations	Outcome d=dichotomous arrest=outcome 6/all=length of follow-up	Predictors t=time TxGroup= treatment group (LEAD vs control) txTX = time x treatment group interaction	Propensity score weighting (ATT or ATE)	Case identifier (participant ID number)	Time point as a unique identifier within ID	Distribution type (binomial)	Link function (logit)	Correlation structure for panel data (exchangeable)	Requests exponentiated coefficients (*OR*s)	Robust standard errors to account for correlated data structure

```
. xtgee   darrest6_  t TxGroup txTx [pweight=ATE], i(id) t(time) family(bin) link(logit) corr(exc
> h) eform  robust

Iteration 1: tolerance = 6.874e-11

GEE population-averaged model                   Number of obs      =       636
Group variable:                         id      Number of groups   =       318
Link:                                logit      Obs per group: min =         2
Family:                           binomial                     avg =       2.0
Correlation:                  exchangeable                     max =         2
                                                Wald chi2(3)       =     19.18
Scale parameter:                         1      Prob > chi2        =    0.0003

                                     (Std. Err. adjusted for clustering on id)
------------------------------------------------------------------------------
             |               Robust
   darrest6_ | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           t |   2.226787   .5723846     3.11   0.002     1.345493    3.685325
     TxGroup |   .8137984   .2001108    -0.84   0.402     .5025841    1.317725
        txTx |     .49352   .1575403    -2.21   0.027     .2639893     .922621
       _cons |   .6124741   .1195511    -2.51   0.012     .4177712    .8979185
------------------------------------------------------------------------------

. xtgee   darrest6_  t TxGroup txTx [pweight=ATT], i(id) t(time) family(bin) link(logit) corr(exc
> h) eform  robust

Iteration 1: tolerance = 3.891e-11

GEE population-averaged model                   Number of obs      =       636
Group variable:                         id      Number of groups   =       318
Link:                                logit      Obs per group: min =         2
Family:                           binomial                     avg =       2.0
Correlation:                  exchangeable                     max =         2
                                                Wald chi2(3)       =     16.10
Scale parameter:                         1      Prob > chi2        =    0.0011

                                     (Std. Err. adjusted for clustering on id)
------------------------------------------------------------------------------
             |               Robust
   darrest6_ | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           t |   2.210097   .5983857     2.93   0.003     1.300013    3.757292
     TxGroup |   .8543784   .2154594    -0.62   0.533     .5211851    1.400582
        txTx |   .5044208   .1664102    -2.07   0.038     .2642279     .962958
       _cons |   .5895558   .1199757    -2.60   0.009     .3956432    .8785086
------------------------------------------------------------------------------
```

```
. xtgee   dcharge6_  t TxGroup txTx [pweight=ATE], i(id) t(time) family(bin) link(logit) corr(exc
> h) eform  robust

Iteration 1: tolerance = 1.147e-10

GEE population-averaged model                   Number of obs      =       636
Group variable:                         id      Number of groups   =       318
Link:                                logit      Obs per group: min =         2
Family:                           binomial                     avg =       2.0
Correlation:                  exchangeable                     max =         2
                                                Wald chi2(3)       =      3.30
Scale parameter:                         1      Prob > chi2        =    0.3473

                                     (Std. Err. adjusted for clustering on id)
------------------------------------------------------------------------------
             |               Robust
   dcharge6_ | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           t |   .7331703    .209313    -1.09   0.277     .4189818    1.282964
     TxGroup |   .6443475    .171368    -1.65   0.098      .382594    1.085181
        txTx |   1.769279   .6270393     1.61   0.107     .8833386    3.543769
       _cons |    .450501    .092713    -3.87   0.000     .3009668    .6743307
------------------------------------------------------------------------------

. xtgee   dcharge6_  t TxGroup txTx [pweight=ATT], i(id) t(time) family(bin) link(logit) corr(exc
> h) eform  robust

Iteration 1: tolerance = 1.400e-10

GEE population-averaged model                   Number of obs      =       636
Group variable:                         id      Number of groups   =       318
Link:                                logit      Obs per group: min =         2
Family:                           binomial                     avg =       2.0
Correlation:                  exchangeable                     max =         2
                                                Wald chi2(3)       =      3.26
Scale parameter:                         1      Prob > chi2        =    0.3533

                                     (Std. Err. adjusted for clustering on id)
------------------------------------------------------------------------------
             |               Robust
   dcharge6_ | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           t |    .701397   .2100087    -1.18   0.236     .3900331    1.261323
     TxGroup |   .6562352   .1790887    -1.54   0.123     .3843827    1.120354
        txTx |   1.853739   .6780948     1.69   0.092     .9050663    3.796792
       _cons |   .4464765   .0960485    -3.75   0.000     .2928758    .6806342
------------------------------------------------------------------------------


. xtgee   dfelony6_  t TxGroup txTx [pweight=ATE], i(id) t(time) family(bin) link(logit) corr(exc
> h) eform  robust

Iteration 1: tolerance = 7.939e-07

GEE population-averaged model                   Number of obs      =       636
Group variable:                         id      Number of groups   =       318
Link:                                logit      Obs per group: min =         2
Family:                           binomial                     avg =       2.0
Correlation:                  exchangeable                     max =         2
                                                Wald chi2(3)       =      3.80
Scale parameter:                         1      Prob > chi2        =    0.2841

                                     (Std. Err. adjusted for clustering on id)
------------------------------------------------------------------------------
             |               Robust
   dfelony6_ | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           t |   1.639358   .7094269     1.14   0.253     .7019709    3.828499
     TxGroup |   .8020622   .3501033    -0.51   0.613     .3409221    1.886952
        txTx |    .947415    .519472    -0.10   0.922     .3234614    2.774968
       _cons |   .0930288   .0312686    -7.07   0.000     .0481409    .1797714
------------------------------------------------------------------------------
```

```
. xtgee   dfelony6_  t TxGroup txTx [pweight=ATT], i(id) t(time) family(bin) link(logit) corr(exc
> h) eform  robust

Iteration 1: tolerance = 5.471e-07

GEE population-averaged model                   Number of obs      =       636
Group variable:                         id      Number of groups   =       318
Link:                                logit      Obs per group: min =         2
Family:                           binomial                     avg =       2.0
Correlation:                  exchangeable                     max =         2
                                                Wald chi2(3)       =      3.41
Scale parameter:                         1      Prob > chi2        =    0.3331

                                     (Std. Err. adjusted for clustering on id)
------------------------------------------------------------------------------
             |               Robust
   dfelony6_ | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           t |    1.59348   .7000028     1.06   0.289     .6736293    3.769403
     TxGroup |   .8112143   .3600199    -0.47   0.637     .3399141    1.935985
        txTx |   .9775409   .5447402    -0.04   0.967     .3279427    2.913881
       _cons |   .0913126   .0316316    -6.91   0.000      .046309    .1800511
------------------------------------------------------------------------------

. xtgee   dwarrest6_  t TxGroup txTx [pweight=ATE], i(id) t(time) family(bin) link(logit) corr(ex
> ch) eform  robust

Iteration 1: tolerance = 2.319e-09

GEE population-averaged model                   Number of obs      =       634
Group variable:                         id      Number of groups   =       317
Link:                                logit      Obs per group: min =         2
Family:                           binomial                     avg =       2.0
Correlation:                  exchangeable                     max =         2
                                                Wald chi2(3)       =      5.90
Scale parameter:                         1      Prob > chi2        =    0.1168

                                     (Std. Err. adjusted for clustering on id)
------------------------------------------------------------------------------
             |               Robust
  dwarrest6_ | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           t |   1.447434   .3934268     1.36   0.174     .8496379    2.465833
     TxGroup |   .7961789   .2141083    -0.85   0.397     .4700084    1.348701
        txTx |   .9553831   .3275748    -0.13   0.894     .4878921    1.870817
       _cons |   .3820835   .0807838    -4.55   0.000     .2524579    .5782658
------------------------------------------------------------------------------

. xtgee   dwarrest6_  t TxGroup txTx [pweight=ATT], i(id) t(time) family(bin) link(logit) corr(ex
> ch) eform  robust

Iteration 1: tolerance = 9.001e-10

GEE population-averaged model                   Number of obs      =       634
Group variable:                         id      Number of groups   =       317
Link:                                logit      Obs per group: min =         2
Family:                           binomial                     avg =       2.0
Correlation:                  exchangeable                     max =         2
                                                Wald chi2(3)       =      5.12
Scale parameter:                         1      Prob > chi2        =    0.1632

                                     (Std. Err. adjusted for clustering on id)
------------------------------------------------------------------------------
             |               Robust
  dwarrest6_ | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           t |   1.460824   .4200739     1.32   0.188     .8314324    2.566664
     TxGroup |   .8532817   .2361269    -0.57   0.566      .496068    1.467721
        txTx |   .9495852   .3359098    -0.15   0.884     .4747082    1.899508
       _cons |   .3629252   .0805065    -4.57   0.000     .2349622    .5605783
------------------------------------------------------------------------------
```

```
. xtgee   darrestall_  t TxGroup txTx evaltime [pweight=ATE], i(id) t(time) family(bin) link(logi
> t) corr(exch) eform  robust

Iteration 1: tolerance = .01567455
Iteration 2: tolerance = .00027194
Iteration 3: tolerance = 5.455e-06
Iteration 4: tolerance = 8.671e-08

GEE population-averaged model                   Number of obs      =       636
Group variable:                         id      Number of groups   =       318
Link:                                logit      Obs per group: min =         2
Family:                           binomial                     avg =       2.0
Correlation:                  exchangeable                     max =         2
                                                Wald chi2(4)       =     55.09
Scale parameter:                         1      Prob > chi2        =    0.0000

                                     (Std. Err. adjusted for clustering on id)
------------------------------------------------------------------------------
             |               Robust
 darrestall_ | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           t |   2.836746   1.032337     2.87   0.004     1.390127    5.788771
     TxGroup |   1.409593    .420773     1.15   0.250     .7852436    2.530365
        txTx |   .2983829   .1065201    -3.39   0.001     .1482185    .6006831
    evaltime |   1.902659   .2935394     4.17   0.000     1.406173    2.574442
       _cons |   .4395685   .2283035    -1.58   0.114     .1588286    1.216535
------------------------------------------------------------------------------

. xtgee   darrestall_  t TxGroup txTx evaltime [pweight=ATT], i(id) t(time) family(bin) link(logi
> t) corr(exch) eform  robust

Iteration 1: tolerance = .01447
Iteration 2: tolerance = .00018418
Iteration 3: tolerance = 3.288e-06
Iteration 4: tolerance = 4.140e-08

GEE population-averaged model                   Number of obs      =       636
Group variable:                         id      Number of groups   =       318
Link:                                logit      Obs per group: min =         2
Family:                           binomial                     avg =       2.0
Correlation:                  exchangeable                     max =         2
                                                Wald chi2(4)       =     53.66
Scale parameter:                         1      Prob > chi2        =    0.0000

                                     (Std. Err. adjusted for clustering on id)
------------------------------------------------------------------------------
             |               Robust
 darrestall_ | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           t |   2.777839    1.06185     2.67   0.008     1.313193    5.876049
     TxGroup |   1.503565   .4569244     1.34   0.180     .8287947    2.727704
        txTx |   .2920957   .1075516    -3.34   0.001     .1419407    .6010954
    evaltime |   1.867028   .2884276     4.04   0.000     1.379282    2.527253
       _cons |   .4444125   .2358074    -1.53   0.126     .1570849    1.257297
------------------------------------------------------------------------------
```

```
. xtgee   dwarrestall_  t TxGroup txTx evaltime [pweight=ATE], i(id) t(time) family(bin) link(log
> it) corr(exch) eform  robust

Iteration 1: tolerance = .0192158
Iteration 2: tolerance = .00031694
Iteration 3: tolerance = 5.390e-06
Iteration 4: tolerance = 8.497e-08

GEE population-averaged model                   Number of obs      =       634
Group variable:                         id      Number of groups   =       317
Link:                                logit      Obs per group: min =         2
Family:                           binomial                     avg =       2.0
Correlation:                  exchangeable                     max =         2
                                                Wald chi2(4)       =     42.16
Scale parameter:                         1      Prob > chi2        =    0.0000

                                     (Std. Err. adjusted for clustering on id)
------------------------------------------------------------------------------
             |               Robust
dwarrestall_ | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           t |   1.108003   .3593668     0.32   0.752     .5867652    2.092269
     TxGroup |   1.135148     .30573     0.47   0.638     .6695728    1.924451
        txTx |   .5838587   .1828716    -1.72   0.086     .3160102    1.078734
    evaltime |   1.417918   .1899503     2.61   0.009     1.090487    1.843663
       _cons |   .8728559   .4125679    -0.29   0.774     .3456288    2.204323
------------------------------------------------------------------------------

. xtgee   dwarrestall_  t TxGroup txTx evaltime [pweight=ATT], i(id) t(time) family(bin) link(log
> it) corr(exch) eform  robust

Iteration 1: tolerance = .01876881
Iteration 2: tolerance = .0002751
Iteration 3: tolerance = 4.419e-06
Iteration 4: tolerance = 6.268e-08

GEE population-averaged model                   Number of obs      =       634
Group variable:                         id      Number of groups   =       317
Link:                                logit      Obs per group: min =         2
Family:                           binomial                     avg =       2.0
Correlation:                  exchangeable                     max =         2
                                                Wald chi2(4)       =     42.26
Scale parameter:                         1      Prob > chi2        =    0.0000

                                     (Std. Err. adjusted for clustering on id)
------------------------------------------------------------------------------
             |               Robust
dwarrestall_ | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           t |   1.064584   .3518934     0.19   0.850     .5569539    2.034889
     TxGroup |   1.173192   .3233712     0.58   0.562      .683517    2.013673
        txTx |   .5935156   .1908217    -1.62   0.105     .3160542    1.114558
    evaltime |   1.410193   .1872166     2.59   0.010      1.08711    1.829295
       _cons |   .8725695   .4129598    -0.29   0.773     .3451064    2.206211
------------------------------------------------------------------------------
```

```
. xtgee   dchargeall_  t TxGroup txTx evaltime [pweight=ATE], i(id) t(time) family(bin) link(logi
> t) corr(exch) eform  robust

Iteration 1: tolerance = .01251121
Iteration 2: tolerance = .00006101
Iteration 3: tolerance = 6.108e-07

GEE population-averaged model                   Number of obs      =       636
Group variable:                         id      Number of groups   =       318
Link:                                logit      Obs per group: min =         2
Family:                           binomial                     avg =       2.0
Correlation:                  exchangeable                     max =         2
                                                Wald chi2(4)       =     46.27
Scale parameter:                         1      Prob > chi2        =    0.0000

                                     (Std. Err. adjusted for clustering on id)
------------------------------------------------------------------------------
             |               Robust
 dchargeall_ | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           t |   .8663963   .2738005    -0.45   0.650     .4663564     1.60959
     TxGroup |   1.099226   .2940063     0.35   0.724     .6507567    1.856757
        txTx |    .644395   .2174559    -1.30   0.193      .332589    1.248523
    evaltime |   1.410499   .1990524     2.44   0.015     1.069669    1.859928
       _cons |   .8013325   .3760103    -0.47   0.637     .3194496    2.010126
------------------------------------------------------------------------------

. xtgee   dchargeall_  t TxGroup txTx evaltime [pweight=ATT], i(id) t(time) family(bin) link(logi
> t) corr(exch) eform  robust

Iteration 1: tolerance = .01285182
Iteration 2: tolerance = .00005905
Iteration 3: tolerance = 6.400e-07

GEE population-averaged model                   Number of obs      =       636
Group variable:                         id      Number of groups   =       318
Link:                                logit      Obs per group: min =         2
Family:                           binomial                     avg =       2.0
Correlation:                  exchangeable                     max =         2
                                                Wald chi2(4)       =     47.91
Scale parameter:                         1      Prob > chi2        =    0.0000

                                     (Std. Err. adjusted for clustering on id)
------------------------------------------------------------------------------
             |               Robust
 dchargeall_ | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           t |    .861725   .2828659    -0.45   0.650      .452853     1.63976
     TxGroup |   1.122255   .3069712     0.42   0.673      .656541    1.918321
        txTx |   .6357422   .2190762    -1.31   0.189     .3235622     1.24912
    evaltime |   1.416724     .19809     2.49   0.013      1.07713    1.863385
       _cons |   .7879315   .3721243    -0.50   0.614      .312236    1.988355
------------------------------------------------------------------------------
```

```
. xtgee   dfelonyall_  t TxGroup txTx evaltime [pweight=ATE], i(id) t(time) family(bin) link(logi
> t) corr(exch) eform  robust

Iteration 1: tolerance = .01610324
Iteration 2: tolerance = .00008353
Iteration 3: tolerance = 4.640e-06
Iteration 4: tolerance = 2.301e-08

GEE population-averaged model                   Number of obs      =       636
Group variable:                         id      Number of groups   =       318
Link:                                logit      Obs per group: min =         2
Family:                           binomial                     avg =       2.0
Correlation:                  exchangeable                     max =         2
                                                Wald chi2(4)       =     33.47
Scale parameter:                         1      Prob > chi2        =    0.0000

                                     (Std. Err. adjusted for clustering on id)
------------------------------------------------------------------------------
             |               Robust
 dfelonyall_ | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           t |   .9341366   .3235802    -0.20   0.844     .4737601    1.841884
     TxGroup |   1.239366   .3003752     0.89   0.376     .7707268    1.992959
        txTx |   .4888799   .1591162    -2.20   0.028     .2583216     .925217
    evaltime |   1.186283   .1660111     1.22   0.222     .9017152    1.560657
       _cons |   .3347915   .1656909    -2.21   0.027     .1269136    .8831626
------------------------------------------------------------------------------

. xtgee   dfelonyall_  t TxGroup txTx evaltime [pweight=ATT], i(id) t(time) family(bin) link(logi
> t) corr(exch) eform  robust

Iteration 1: tolerance = .0174253
Iteration 2: tolerance = .00009575
Iteration 3: tolerance = 6.315e-06
Iteration 4: tolerance = 3.247e-08

GEE population-averaged model                   Number of obs      =       636
Group variable:                         id      Number of groups   =       318
Link:                                logit      Obs per group: min =         2
Family:                           binomial                     avg =       2.0
Correlation:                  exchangeable                     max =         2
                                                Wald chi2(4)       =     34.85
Scale parameter:                         1      Prob > chi2        =    0.0000

                                     (Std. Err. adjusted for clustering on id)
------------------------------------------------------------------------------
             |               Robust
 dfelonyall_ | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           t |   .9617581   .3433658    -0.11   0.913     .4777171    1.936248
     TxGroup |   1.347295   .3332451     1.21   0.228      .829704    2.187772
        txTx |   .4716183   .1556055    -2.28   0.023     .2470277    .9004005
    evaltime |   1.195887   .1678032     1.27   0.202     .9083476    1.574447
       _cons |   .3030095   .1514837    -2.39   0.017     .1137403    .8072312
------------------------------------------------------------------------------
```

Appendix B. Effect size calculations for interpretation of the interaction effect for the LEAD group

Outcomes	**Intervention group *OR***	**Interaction *OR***	**OR incident at follow-up**	**Reduction/Increase**
arrest6 ATE	0.8137984	0.49352	0.40	-0.60
arrest6 ATT	0.8543784	0.5044208	0.43	-0.57
arrestall ATE	1.409593	0.2983829	0.42	-0.58
arrestall ATT	1.503565	0.2920957	0.44	-0.56
warrest6 ATE	0.7961789	0.9553831	0.76	-0.24
warrest6 ATT	0.8532817	0.9495852	0.81	-0.19
warrestall ATE	1.135148	0.5838587	0.66	-0.34
warrestall ATT	1.173192	0.5935156	0.70	-0.30
charge6 ATE	0.6443475	1.769279	1.14	0.14
charge6 ATT	0.6562352	1.853739	1.22	0.22
chargeall ATE	1.099226	0.644395	0.71	-0.29
chargeall ATT	1.122255	0.6357422	0.71	-0.29
felony6 ATE	0.8020622	0.947415	0.76	-0.24
felony6 ATE	0.8112143	0.9775409	0.79	-0.21
felonyall ATE	1.239366	0.4888799	0.61	-0.39
felonyall ATT	1.347295	0.4716183	0.64	-0.36

Notes: Outcomes followed by a “6” indicate shorter-term, six-month outcomes; whereas outcomes followed by “all” indicate longer-term, evaluation-wide outcomes. ATT = Average treatment effect for the LEAD participants. ATE = Average overall treatment effect. OR = Odds ratio.