Journal of Applied Juvenile Justice Services (2020)

Journal of Applied Juvenile Justice Services

Examining What Works for High Risk Youth Involved in the Juvenile Justice System: Comparing the Effect of the Community Connections (CC) Program to Intensive Supervision Probation (ISP) Over A Nine-Year Time Period

Erin M. Espinosa, Ph.D.
National Council on Crime and Delinquency

Dan Sass, Ph.D.
University of Texas at San Antonio

Johanna Creswell Báez, Ph.D.
Columbia University

Cassandra Harper, MPH
Southwest Key Programs

Using administrative data collected by an urban juvenile probation department between January 2007 and August 2016, this study compared the impact of the Community Connections (CC) program to a matched sample of youth who received intensive supervision probation (ISP) on six time-to-event variables (i.e., time to second program, detention, out-of-home placement, another offense, violation of court order, and days in program). The study included youth who were assigned to court-ordered post-adjudication community supervision and who were deemed to have a high risk of re-offending by the department's risk and needs assessment. CC and ISP youth were matched using propensity score matching that created a final sample of 381 youth in each program. When examining the program effect of CC vs. ISP on the time-to-event variables the findings were mixed. However, across both programs, the analysis revealed that youth who remained in the programs longer and youth with a successful program discharge generally experienced better long-term outcomes than their peers.

INTRODUCTION

Generally, juvenile crime rates have declined over the last decade. In 2013, rates of youth arrests for violent offenses was the lowest they had been in 30 years (Puzzanchera, 2013). The number of cases processed by juvenile courts has decreased for all offense types for the last 10 years, with over one million cases still processed by juvenile courts each year (Hockenberry & Puzzanchera, 2015). For youth involved with the juvenile justice system, research indicates system involvement negatively influences the development of their psychosocial maturity; youth

who are removed from their homes by the juvenile justice system have increased risks for physical violence, self-destructive behavior, and engaging in future delinquent activity (Duke, Pettingell, McMorris, & Borowsky, 2010). For youth who have returned from juvenile justice placements, studies have demonstrated re-arrest rates ranging from 40% (Taylor, Kemper, Loney, & Kistner, 2009) and 65% (Benda, Corwyn, & Toombs, 2001) to as high as 85% (Trulson, Marquart, Mullings, & Caeti, 2005).

Community Success for Youth Involved in the Juvenile Justice System

Juvenile justice reformers are focusing on alternatives such as deinstitutionalization, alternatives to detention and incarceration, and diversion from deeper system involvement for youth who are already involved in the justice system (Bishop & Decker 2006; Loeb, Waung, & Sheeran 2015; McAra & McVie 2007). While balancing accountability for delinquent behavior with rehabilitation and treatment options (Beck, Ramsey, Lipps, & Travis, 2006; Mackin et al., 2010), diversion programs serve as an alternative to traditional processing (Harris, Lockwood, Mengers, & Stoodley, 2011; Leve & Chamberlain, 2005). National estimates inidicate at least a quarter of the youth arrested across the United States end up in a diversion program (Puzzanchera & Kang, 2008). These programs vary in their design and approach (Hamilton, Sullivan,Veysey, & Grillo, 2007) depending on the type of program (Hoge, 2016; Mears et al., 2016), the risk level of the youth enrolled (Vincent et. al., 2012), and the intercept point in the juvenile justice system process (Cocozza et al., 2005).

Studies consistently show that youth involved with the juvenile justice system who are supported by community-based services have lower re-offense rates than youth who are not supported by community-based services (Colwell, Villarreal, & Espinosa, 2012 ; Cuellar, Markowitz, & Libby, 2004; Cuellar, McReynolds, & Wasserman, 2006; Wilson & Hoge, 2013). A meta-analysis examining the impact of juvenile justice diversion programs on youth with varying degrees of justice system involvement concluded, that although the diversion programs varied in model and scope, recidivism rates were significantly lower in 60 of the 73 studies reviewed for youth diverted than for youth who remained involved in the juvenile justice system (Wilson & Hoge, 2013). These youth also have lower rates of pretrial detention (Cuellar et al., 2006), fewer subsequent arrests (Cuellar et al., 2004), and, when tied to services that address their involvement in delinquent behavior, are less likely to be adjudicated or placed deeper within the juvenile justice system (Colwell et al., 2012).

Community Connections Program

The Community Connections program (CC) was established in 2005 by Southwest Key (SWK) Programs. The program evolved out of SWK's Family Keys Program (status offender diversion model) and was developed in partnership with a combination of the Departments of Social Services, Mental Health, and Juvenile Probation. Family Keys formed when service providers recognized youth were receiving duplicate services from multiple case managers within the three departments and subsequently established a single case manager as the central point of contact for youth. CC was formed when the SWK Family Keys providers noticed many youth and families needed more comprehensive set of services than what was offered in the traditional 30-day Family Keys program. The CC model supports adolescent males and females between the ages of 10 and 17 under post-adjudication probation supervision who have been determined to

have a moderate- to high-risk of re-offending by standardized risk and needs assessments. The program was designed to last from four to six months and includes six service levels, ranging from the highest level of need (Pre-Release Reunification) to the lowest level of need (Transition). As youth move through the service levels, youth learn and practice coping skills, families gain self-sufficiency in accessing community-based services and resources, and gain skills to manage and de-escalate family conflict. The CC model includes the following: a counseling component for youth and families to supplement any need for long-term therapy, group counseling sessions to develop youth coping skills, intensive case management and service coordination to support the complex needs of families, and monitoring services to track each youth's location at curfew.

The CC aims to build on each family's strengths to support their individualized areas of need. Case managers refer youth and families to services within the community as needed and help families develop the skills to access community-based resources on their own for when program services end. The program uses the New Freedoms/Phoenix Core Curriculum, which incorporates evidence-based practices including elements of Cognitive-Behavioral Therapy and Motivational interviewing.

Current Study

This study was conducted to examine how well youth in the CC program did in the program compared to similar youth participating in Intensive Supervision Probation (ISP). Administrative data was collected on all youth processed by an urban juvenile probation department (JPD) in a large southwestern state during the period from January 2007 to August 2016. Youth determined to have high risk and needs levels were referred by their supervising probation officers to both programs. All risk and needs determinations were made by probation officers trained in administering the department's risk and needs assessment.

Both ISP and CC programs were implemented within the JPD as programs for youth with high risk of re-offending as determined by the department's risk and needs assessment, and were used as interventions for youth under post-adjudication supervision. ISP is a nationally recognized supervision approach for both adult and juvenile justice and is typically viewed as a "tough" community alternative to incarceration. The core components associated with ISP includes small caseloads, intensive surveillance, and strict conditions for compliance with court orders (Taxman, 2002). Research on ISP programs has demonstrated mixed results, ranging from no effect (Hyatt & Barnes, 2017) to a decreased risk of re-offending (Bouchard & Wong, 2018). Youth were included in the current study if they received a post-adjudication disposition decision of formal probation and if their involvement in either ISP or CC began after January 1, 2007.

The primary aim of the study was to examine the program characteristics, risk factors, and demographic variables that translated to a successful discharge status for youth receiving the CC program compared to the JPD's ISP program. The primary research question for this study was: Do youth who participate in CC have longer times before receiving a violation of court order, detention, out-of-home placement, another offense, or another program as compared to youth in ISP? The secondary research question was: Do youth outcomes vary by age, ethnicity, felony

history, history of violence, gender, substance use history, or mental health needs when CC is compared to ISP?

Participants

Administrative data collected included 676 CC and 647 ISP youth. These groups were not comparable given their differing background characteristics. Consequently, propensity score matching (PSM) analyses were conducted to create comparable groups and to increase the statistical and internal validity of the study. More specifically, PSM was conducted to ensure any program group differences resulted from the program (CC or ISP) and not the youth's risk and background characteristics. After PSM (see *Propensity score matching* section), there were 381 youth in both CC and ISP, with both groups' demographic and risk characteristics being nearly identical as shown in Table 1.

Table 1 results show most of the youth were males and Hispanic. The average youth age was 15 years old, and there was no statistically significant difference between CC (M = 15.33, SD = 1.24) and ISP (M = 15.34, SD = 1.27) youth, F (1, 760) ≈ 0.00, $p = 0.9769$. Regardless of the group, the majority of youth had previous mental health needs, substance use issues, prior felonies, and previous violent histories prior to placement in their respective programs.

Though the amount of time youth were involved in their programs differed prior to PSM, youth remained in their programs for a relatively equal number of days after matching, with no statistically significant difference between CC (M = 115.95, SD = 51.11) and ISP (M = 117.19, SD = 60.28) youth after PSM, F (1, 760) = 0.09, $p = 0.7592$. It is worth noting that despite the means and estimated standard deviations being comparable across groups, the estimated standard deviations were large, suggesting significant variation in the time spent in the program. Despite there being differences before matching, the number of previous programs received were equal after PSM. The mean and estimated standard deviations were as follows for the two groups after matching: CC (M = 3.65, SD = 2.58) and ISP (M = 3.69, SD = 3.10), F (1, 760) = 0.04, $p = 0.8391$.

Measures

Based on the data collected by the JPD, variables were classified into the following categories: risk factors, discharge status, demographic variables, and outcome variables. Note the first three categories were predictor variables, with the outcome variables being the time-to-event variables. These variables are described below, and there was no missing data present.

Predictor variables

Risk factors. Four risk factors were evaluated in this study. *Mental health needs*, which is identified by the youth's juvenile probation officer, consists of three categories: "No, mental health needs," "Yes, mental health needs," and "Unknown." For the survival analysis models, "No, mental health needs" was used as the reference group. *Previous felony*, which is defined as whether or not youths had a previous felony on their records resulting in their receiving a program, was classified as either "Yes" or "No." For analyses, "No" was used as the reference group. *Previous violence*, which is defined as whether or not the youth had a previous violent offense resulting in their receiving a program from the JPD, was classified as either "Yes" or

"No," with "No" being used as the reference group in statistical models. *Substance use* is defined as whether youths had reported substance use, such as marijuana or alcohol, on their record. This variable was classified as either "Yes" or "No," with "No" being used as the reference group. The *number of previous programs* from January 2007 until they started CC or ISP was also used as a covariate, as it is possible their prior involvement in less-intensive programs may provide an added benefit in how they respond to their current program under evaluation. Having received several previous less-intensive programs also provides an indication of previous problems within the juvenile justice system, suggesting these youths may be at a higher risk of repeat criminal offense overall.

Table 1. Categorical demographic variable information for CC and ISP youth post matching.

	CC (n = 381) *n (%)*	ISP (n = 381) *n (%)*	Test statistic (df)	p-value
Gender			χ^2 (1) = 0.08	0.7772
Male	311 (82%)	314 (82%)		
Female	70 (18%)	67 (18%)		
Race/ethnicity			χ^2 (2) = 0.03	0.9857
White	31 (8%)	32 (8%)		
Black	105 (28%)	106 (28%)		
Hispanic	245 (64%)	243 (64%)		
Mental health needs			χ^2 (2) = 3.30	0.1921
Yes	257 (67%)	271 (71%)		
No	78 (20%)	79 (21%)		
Unknown	46 (12%)	31 (8%)		
Substance use			χ^2 (1) = 0.01	0.9269
Yes	308 (81%)	307 (81%)		
No	73 (19%)	74 (19%)		
Previous felony			χ^2 (1) = 0.25	0.8843
Yes	207 (54%)	209 (55%)		
No	174 (46%)	172 (45%)		
Previous violence			χ^2 (1) = 0.25	0.6141
Yes	284 (75%)	290 (76%)		
No	97 (25%)	91 (24%)		
Discharge status			χ^2 (1) = 1.24	0.2658
Unsuccessful	142 (37%)	157 (41%)		
Successful	239 (63%)	224 (59%)		

Program characteristics. The *discharge status* variable indicates whether or not the youth had a successful or unsuccessful discharge, with "Unsuccessful discharge" being used as the reference group. One limitation of this variable is that program success was defined by juvenile probation

staff rather than by program CC staff; thus, the reason and justification for each release is unknown. Moreover, subjective perceptions likely vary across probation officers, thus influencing the reliability of the variable. The *days in program* variable was utilized as both a predictor and outcome variable. When employed as a predictor, this variable was used to determine whether the number of days in the program resulted in better, or possibly worse, program outcomes (see *Outcome variables* below) – this is known as a dose effect variable. As an outcome variable, the other predictor variables were used to predict the number of days the youth stayed in the program.

Demographic variables. To understand whether the outcome variables differed based on youth demographics, the following demographic variables were included in the model as predictor variables: race/ethnicity (White, Black, and Hispanic), gender (Male and Female), and age at the beginning of the program. For these categorical variables, White (race/ethnicity) and male (gender) youth were used as the reference groups.
Time-to-event variables/Outcome Variables

Six time-to-event variables, measured in days, were examined to evaluate the long-term impact of the CC program on youth, while also using the aforementioned predictor variables to explain differences in survival rates. The first outcome variable, *Time to detention,* assessed the number of days until the youth received a detention, specifically if the youth had been arrested and placed in a detention center before trial or sentencing after their CC program ended. As with all outcome variables examined here, the data were censored if the youth had not received an event (in this case a detention) as of August 31, 2016. *Time to placement* was used to determine the number of days until the youth received a placement, defined as an out-of-home program option. These could be either secure or non-secure placements and could include treatment programs as well as correctional models. The *Time to next offense* outcome variable measures the number of days until the youth committed another (known) new or repeat offense. *Time to a violation of court order* is the number of days until the youth was charged with violating their court orders. *Time to second program* measures the number of days until a youth is enrolled in another program after CC or ISP. Lastly, *days in program* measures the number of days a youth spent in their program before being released. Recall this variable was also used as a predictor variable when predicting the other time-to-event variables.

Propensity score matching

To create statistically equivalent groups via PSM, nearest neighbor matching was conducted using R. Given that exact matching (i.e., youth being identical on all variables) would have reduced the sample size significantly, this study elected to use a caliper to assess the degree of difference between the groups being matched. While several calipers were explored, this study deemed a caliper of .11 was best in order to maximize the sample size and minimize group differences on the covariates of interest.

To create equal CC vs. ISP groups at pre-treatment, this study selected the set of variables that were required to be equivalent across groups for valid inferences. The following variables were included: age, gender, race, mental health needs, previous violence, previous substance use,

previous felony, number of previous programs/services, discharge status, and number of days in the program.

Statistical analyses for survival analyses

Model building process. Non-parametric Cox survival analysis models were estimated to predict the time-to-event conditional on the predictor variables in the model. All models were graphically examined using PROC LIFEREG and built using PROC PHREG within SAS 9.4 while following a five-step process. Step 1, all predictor variables, along with their two-way interactions, were included in the model to determine those variables/terms that significantly predicted the time-to-event outcome variable. This was done using a stepwise method with an entry and stay value of $p = 0.20$ and $p = 0.05$, respectfully. After the best stepwise model was selected, predictor variables associated with any statistically significant interaction term were brought back into the model in Step 2, if not already included. In Step 3, the model fit (i.e., R^2 statistic, C-statistics/area of the curve, AUC) was assessed to evaluate the overall model quality and determine whether complex interaction terms could be removed from the model without significantly harming model quality, prediction, or interpretability. This was done to reduce model complexity. Next for Step 4, we tested whether the model assumptions were met (such as proportional hazards assumption) and made model modifications that included time varying covariates, if model assumptions were violated. Lastly for Step 5, the final overall model quality was reassessed and interactions between variables were explored in detail to better understand the results.

Terminology. To ease the presentation and discussion of the results, commonly used survival analysis terms are presented here briefly. When discussing time-to-event variables, this refers to the number of days until that event of interest occurs (e.g., number of days between program start and detention or placement). An observation is said to be censored if the event does not occur before August 31, 2016 and, therefore, the time until that event occurs, if ever, is unknown. Given censoring depended on when the youth entered the program (i.e., a youth entering the program in June of 2007 is less likely to be censored than a youth entering the program in June of 2016) the *days since program ended* variable was included in the model, although it was not of practical interest.

The hazard function, f(t), describes the risk of the event occurring at time t conditional on the youth's survival up to that time t, S(t). The hazard ratio (HR) assesses how often a hazard event occurs in one group (e.g., CC) compared to how often it occurs in another group (e.g., ISP) over time. With categorical variables, it compares the hazard rate of the comparison group to the hazard rate of the reference group. For continuous predictor variables, the HR measures the change in hazards for a one-unit change in the predictor variable. Using the days in program as an example variable, it evaluates the change in hazards for being in the program 20 versus 21 days. For this reason, the hazards were compared at different time intervals (i.e., 30, 60, 90, 120, & 150 days) in this study to make the results easier to interpret.

It is important to note HRs can vary from zero to infinity, with a value of one indicating the hazard rate is the same for both groups. It is also important to recognize an HR of .80 does not have the same impact as an HR of 1.20. For comparability and interpretability purposes, one can

take the reciprocal (1/HR) to change the direction of the HR. For example, a HR of .50 (female vs. male) is equal to an HR of two (male vs. female, 2 = 1/.50). It is also worth mentioning any 95% confidence interval (CI) that does not include one is significant at the .05 level.

To assess how well the model predicted the outcome variables, the area of the curve (AUC) or C-statistic was used to provide the predictive accuracy at specific times - in this case, days. An AUC or C-statistic above .80 tends to provide evidence of a good predictive model; however, it is important to recall that prediction accuracy is time-dependent. For this study, the model accuracy tended to be poorer immediately after the program's end, likely due to the large number of changes in the lives of the youths and the adjustment back to their homes and neighborhoods after incarceration, which improves and stabilizes over time.

Results Roadmap

For each outcome variable, the results are presented in three major sections. The first section (labeled *Model performance*) focuses on the performance or predictive accuracy of the overall model while also discussing the level of censoring. More specifically, this section focuses on whether those variables in the model adequately predict survival rates and whether model prediction varied over the life of the study.

The second section (labeled *Program-related effects*) evaluates the program effect variables of program, discharge status, and days in the program, while also including potential interactions with other background variables. The program variable (CC vs. ISP) is of greatest interest, as this study's primary goal was to assess differences in survival rates between CC and ISP. Discharge status was also of interest because it evaluated the impact of youth success in completing the program, which should serve as a marker of later successes and positive outcomes. The number of days in the program was also examined in this section to determine whether youth who stayed in the program longer – and thus have a larger dose effect – display better outcomes, as seen by longer times until an event occurs.

The third section (labeled *Background variable effects)* focuses on those background variables that significantly predicted the time to event variables. While these background variables cannot be manipulated to alter program outcomes, which are the time to event variables, it is extremely useful to know which youth background characteristics predict success and positive outcomes long-term. In turn, knowing these background characteristics could result in more targeted enrollment of youth and program modifications to provide better services. For example, if youth with a history of mental health needs are at greater risk of receiving later court order violations, perhaps programs should be altered to better serve these youths in close coordination with court staff.

RESULTS

Time to second program

Model performance. Table 2 includes the significant predictors of time to receiving a second intervention program. The overall model's likelihood ratio (LR) χ^2 was statistically significant, (df = 13) = 246.98, $p < .0001$, $R^2 = .28$, AUC = .69. In terms of censoring, the percentage was

comparable between CC (44%, *n* = 168) and ISP (43%, *n* = 164) youth. Those youth classified as censored had not received another program after being released from their current programs as of August 2016. From a program outcomes perspective, higher rates of censoring are a positive finding, as it indicates these youth had not received a second program. However, it is important to interpret these results with caution, as they do not consider when the program ended, and a program that ends later is more likely to be censored.

Table 2. Parameter estimates and test statistics for time to second program model.

Predictor	NRG1	NRG2	χ^2
Model 1: Time to second program			
Program	CC		2.86
Discharge status	Successful		12.62
Days in Program	116.57		20.65
Race	Black		0.63
Race	Hispanic		0.58
Gender	Female		8.33
Substance use	Yes		5.96
Mental health needs	Unknown		5.06
Mental health needs	Yes		5.94
Gender*Substance use	Female	Yes	6.72
Age	15.33		23.05
Age*Time			29.00
Days since program end	2028.90		13.02

Note. Mental health needs, discharge status, days in program, days since program ended, and non-reference group are denoted by MHN, PO, DIP, DSPE, and NRG, respectfully. NRG2 is the additional reference group when an interaction is modeled.

Program-related effects. Although marginally significant, the program variable was not a significant (HR = 1.18, *p* = 0.0908) predictor of time to a second program and the hazard ratio was relatively small, indicating there was really no difference in the time to second program outcome between youth in CC and youth in ISP. However, discharge status was a significant predictor of time to second program (HR = 0.69, *p* = 0.0004), suggesting youth in both programs who were successfully discharged, had a 69% reduction in risk compared to youth who were classified as unsuccessful at discharge. The number of days in the program was also a significant predictor of time to second program. Follow-up analyses revealed that the hazard rates increased as the number of days in the program increased to 30 (14.3% reduction), 60 (30.5% reduction), 90 (49.2% reduction), 120 (70.4% reduction), and 150 (94.9% reduction) days. Stated more simply, the likelihood that a youth would discharge successfully increased the longer they were enrolled in either program.

Background variable effects. An investigation of the background predictor variables in Table 2 revealed several significant findings. For example, examining the gender-by-substance use interaction uncovered a difference in hazard rates between those with and without substance use for both males (HR = 1.37, 95% CI = 1.06 – 1.75) and females (HR = 0.20, 95% CI = 0.03 - 0.66). However, these results suggest, regardless of treatment program, males with a substance use history had a hazard rate 1.37 times greater than males without a substance use history, whereas the opposite held true for females (based on the hazard ratio of 0.20 for females with a substance use history). Given this significant gender-by-substance use interaction, the substance use status and gender direct effects should not be interpreted.

The mental health needs variable was also a significant predictor of time to a youth's referral to a second program. Using "no mental health needs" as the reference group, the results revealed youth with mental health needs (regardless of program) were at a higher risk (HR = 1.38) of experiencing a referral to a second program. There was also a significant difference in hazard rates between racial/ethnic groups when examining the overall joint tests, χ^2 (df = 2) = 7.41, p = 0.0246. Although this result is not evident in Table 2, follow-up analyses revealed a significant difference in hazard rates between Black and Hispanic youth (HR = 0.73, p = 0.0068), with Black youth having a hazard rate 1.37 (HR = 1/0.73 = 1.37) times that of Hispanic youth. Finally, younger youth were less likely than older youth to receive a second program based on the HR below one. Although this finding was common across many of the survival analysis models, the reason for this result remains unclear, and therefore it was not given considerable attention in the later models.

Time to detention

Model performance. The overall model significantly predicted a youth's time to detention, df = 13) = 364.74, $p < .0001$, $R^2 = .38$, AUC = .70, with 44% of the CC youth (n = 168) and 42% of the ISP youth (n = 160) being censored.

Program-related effects. These results show the program variable (p = 0.0007), not the discharge status variable (p = 0.1441), was a significant predictor of time to detention (Table 3). Both of these findings need to be interpreted with caution given the significant program by discharge status interaction (p = 0.0006), which suggests these two variables interact to influence a youth's time to detention. Further analyses revealed no significant difference between CC and ISP when the discharge status was classified as successful (HR = 1.22, 95% CI = 0.94 to 1.59), but there was a significant difference in hazard rates when the discharge status was labeled unsuccessful (HR = 0.62, 95% CI = 0.47 to 0.82). This suggests CC youth had a 62% reduction in risk of an unsuccessful discharge status compared to similarly at-risk youth in ISP. However, the discharge status by time interaction implies the hazards were not proportional over the duration of this study, so these results can only be interpreted as the average effect and cannot be considered equal at each time point in the study.

Table 3. Parameter estimates and test statistics for time to detention model.

Predictor	NRG1	NRG2	χ^2	p-value	HR
Model 2: Time to detention					
Program	CC		11.53	0.0007	
Discharge status	Successful		2.13	0.1440	
Days in program	116.57		5.35	0.0207	1.00
Race	Black		1.72	0.1892	1.32
Race	Hispanic		6.75	0.0094	1.67
Gender	Female		0.12	0.7256	
Violence	Yes		0.04	0.8397	
Program*Discharge status	CC	Successful	11.92	0.0006	
Gender*Violence	Female	Yes	4.50	0.0339	
Discharge status*Time			20.95	<.0001	1.00
Age	15.33		44.17	<.0001	0.76
Days since program ended	2028.90		22.84	<.0001	1.00

Note. Mental health needs, discharge status, days in program, days since program ended, and non-reference group are denoted by MHN, PO, DIP, DSPE, and NRG, respectfully. NRG2 is the additional reference group when an interaction is modeled.

The number of days in program was also a significant predictor of time to detention, indicating youth experienced better their rates of success the longer they remained in the program. Specifically, the reduction in hazard rates for detention increased as the number of days in the program increased to 30 (6.6% reduction), 60 (13.5% reduction), 90 (21.1% reduction), 120 (29.0% reduction), and 150 (37.4% reduction) days.

Background variable effects. Although neither gender nor violence had a direct effect on the time to detention hazards, the gender-by-violence interaction was significant ($p = .0339$). Follow-up analyses revealed no difference between males with and without a history of violence (HR = 0.97, 95% CI = 0.75 to 1.25), but there was a significant difference in hazards between females with and without a history of violence (HR = 1.86, 95% CI = 1.08 to 3.19). Based on these results, females with a history of violence were 1.86 times more likely to have a shorter time to detention than females without a history of violence. As seen in Table 3, while White and Black youth did not differ in their hazard rates for time to detention, Hispanic youth had a hazard rate 1.32 times that of White youth.

Time to placement

Model performance. The model variables in Table 4 were significant predictors of time to placement, (df = 13) = 484.49, $p < .0001$, $R^2 = .47$, AUC = .77, with about half of the data being censored for CC (46%, $n = 230$) and ISP (46%, $n = 176$) youth. This indicates that about half of the youth had not yet received a second placement after their program ended.

Table 4. Parameter estimates and test statistics for time to placement model.

Predictor	NRG1	NRG2
	Model 3: Time to placement	
Program	CC	
Discharge status	Successful	
Days in program	116.57	
Gender	Female	
Violence	Yes	
Mental health needs	Unknown	
Mental health needs	Yes	
Program*Violence	CC	Yes
Program*Discharge status	CC	Successful
Day in Program*Time		
Age	15.33	
Age*Time		
Days since program ended	2028.89	
Predictor	NRG1	NRG2

Note. Mental health needs, discharge status, days in program, days since program ended, and non-reference group are denoted by MHN, PO, DIP, DSPE, and NRG, respectfully. NRG2 is the additional reference group when an interaction is modeled.

Program-related effects. All three of the program effect variables (i.e., program, discharge status, and days in the program) were significant predictors of time to placement. However, as seen in Table 4, all of these variables also possessed significant interaction terms with other variables in the model, making it hard to understand the direct effect of each variable. Starting with the program-by-discharge status interaction, these results revealed that while CC always had better hazard rates, the difference in hazard rates was greatest when the discharge status was classified as unsuccessful. Specifically, the hazard rate (HR = 1.67, 95% CI = 1.21 to 2.31) for ISP youth was 1.67 times greater than that of CC youth when the discharge status was classified as successful, and 3.90 times greater (HR = 3.90, 95% CI = 2.87 to 5.33) when youth were categorized as unsuccessful.

The program-by-violence interaction showed CC youth performed better than ISP youth, but this program effect differed based on the youth's history of violence. The hazard ratios showed the hazard rate for ISP youth (HR = 2.28, 95% CI = 1.78 to 2.92) was 2.28 times that of CC youth when there was no history of violence, though this difference increased to 4.74 times the risk for ISP youth (HR = 4.74, 95% CI = 2.97 to 7.70) compared to CC youth when there was a history of violence.

As stated above and indicated in Table 4, the days in program-by-time interaction suggests the proportional increase of days in program is not equal over time. This finding is evident when examining the hazard ratios at various time intervals, as the reduction in hazard rates increased (again, non-proportionally) as the number of days in the program increased to 30 (34.2%

reduction), 60 (79.9% reduction), 90 (141.5% reduction), 120 (223.6% reduction), and 150 (334.8% reduction) days.

Background variable effects. Two background variables not included in the interaction terms were significant predictors of time to placement. Results revealed females had a lower hazard rate than males (HR = 0.66), with the hazard rate for males being 1.52 times the hazard rate for females (HR = 1/0.66 = 1.52). Mental health needs were also a significant predictor of time to placement; using no mental health needs as the reference group, the results in Table 4 indicates youth with mental health needs were at a higher risk (HR = 2.01) of experiencing a subsequent out-of-home placement than youth without mental health needs. However, there was no significant difference in hazard rates between youth with unknown mental health needs and those with no mental health needs (HR = 0.85, $p = 0.6043$).

Time to next offense

Model performance. The model predictors in Table 5 provided evidence of a good prediction of time to next offense, (df = 12) = 199.12, $p < .0001$, $R^2 = .23$, AUC = .67. For this outcome, the percent of censoring was only about one third for both CC youth (36%, $n = 137$) and ISP youth (33%, $n = 127$), a somewhat disheartening finding as it shows most youth committed another offense after their program ended.

Table 5. Parameter estimates and test statistics for time to next offense model.

Predictor	NRG1	NRG2	χ^2	p-value	HR
Model 4: Time to next offense					
Program	CC		1.56	0.2111	1.12
Discharge status	Successful		2.75	0.0974	
Days in program	116.57		12.10	0.0005	1.00
Race	Black		0.24	0.6220	1.10
Race	Hispanic		4.75	0.0292	1.47
Gender	Female		17.25	<.0001	0.57
Mental health needs	Unknown		2.44	0.1181	
Mental health needs	Yes		1.10	0.2950	
Discharge status*Mental health needs	Successful	Unknown	5.65	0.0174	
Discharge status*Mental health needs	Successful	Yes	0.37	0.5420	
Age	15.33		101.20	<.0001	0.71
Days since program end	2028.89		7.48	0.0062	1.00

Note. Mental health needs, discharge status, days in program, days since program ended, and non-reference group are denoted by MHN, PO, DIP, DSPE, and NRG, respectfully. NRG2 is the additional reference group when an interaction is modeled.

Program-related effects. The results in Table 5 indicate there was no difference between CC and ISP in time to next offense (HR = 1.12, p = 0.2111), nor was there a significant difference between discharge groups ($p = 0.0974$) at the .05 significance level. The significant discharge status by mental health needs interaction suggests the difference between discharge status groups was a function of the youth's mental health needs. Follow-up analyses revealed no difference

between those youth classified as successful and those classified as unsuccessful when they did not have mental health needs (HR = 0.69, 95% CI = 0.45 to 1.08). However, among youth with mental health needs, there was a 25% reduction (HR = 0.80 or 1/.80 = 1.25, 95% CI = 0.65 to 0.99) in their hazard rate when their discharge status was successful. While somewhat puzzling, this same finding was obtained when their mental health needs were unknown, since youth with unknown mental health needs who completed the program had a 370% (HR = 0.27 or 1/0.27 = 3.70, 95% CI = 0.14 to 0.51) reduction in risk of reoffending.

The number of days in program also had a significant direct effect on time to next offense (HR ≈ 1.00, $p = 0.0005$), with those remaining in the program longer displaying a longer time until their next placement. Once again, the reduction in hazard rates for time to reoffending increased as the number of days in the program increased to 30 (9.5% reduction), 60 (19.9% reduction), 90 (31.2% reduction), 120 (43.9% reduction), and 150 (57.5% reduction) days.

Background variable effects. As seen in Table 5, both gender and race were significant predictors of time to next offense. These results indicate while there was no difference in hazard rates between Black and White youth (HR = 1.10, $p = 0.6220$), a difference did emerge between Hispanic and White youth (HR = 1.47, $p = 0.0292$), with Hispanic youth having a hazard rate 1.47 times that of White youth. It is worth noting that Black and Hispanic youth were also significantly different (HR = 0.75 or 1/0.75 = 1.33, $p = 0.0053$), with Hispanic youth having a hazard rate 1.33 times greater than Black youth. The results in Table 5 indicate males had a hazard rate 1.75 times (HR = 1/0.57 = 1.75) greater than that of females.

Time to violation of court order

Model performance. This model resulted in a significant prediction of time to violation of court order using the variables in Table 6 and resulting in a large AUC regardless of study time point, (df = 9) = 456.44, $p < .0001$, $R^2 = .47$, AUC = .84. However, about two thirds of the data were censored for CC (66%, $n = 252$) and ISP (65%, $n = 246$) youth, indicating most of the CC youth never experienced a violation of their court order.

Program-related effects. The program effect results in Table 6 revealed a significant program ($p = 0.0085$), program-by-time interaction ($p = 0.0041$), and discharge status ($p < 0.0001$) effect. However, the days in program effect ($p = 0.4207$) was not significant, thus implying the number of days in the program did not influence whether youth later had a violation of their court orders. The program variable effect indicated the hazard rate for CC youth was .64 times less than that of ISP youth (HR = 0.64 or 1/0.64 = 1.56). However, the significant program-by-time interaction implies the hazard rates were not proportional over time; thus, the hazard ratio of 0.64 is the average hazard ratio, but it is not the hazard ratio at every point in time. A visual inspection of the hazards conferred this non-proportionality with a larger difference in hazards between roughly day 250 and day 450. The hazard ratio of 0.36 (or HR = 1/0.36 = 2.78) for discharge status suggests youth with an unsuccessful discharge status had a hazard rate of violating court orders 2.78 times more than youth with a successful discharge status.

Background variable effects. Only two background variables, mental health needs and substance use, were significant predictors of time to violation of court order. For the mental health needs

variable, there was a significant difference between youth with both known (HR = 1.46, p = 0.0258) and unknown (HR = 0.61, p = 0.0288) mental health needs when compared to those with no mental health needs. Contrary to expectations, youth with no substance use history possessed a larger hazard compared to those with a substance use history (HR = 0.67 or 1/0.67 = 1.49).

Table 6. Parameter estimates and test statistics for time to violation of court order model.

Predictor	NRG1	NRG2	χ^2	p-value	HR
Model 5: Time to violation of court order					
Program	CC		6.92	0.0085	0.64
Discharge status	Successful		59.32	<.0001	0.36
Days in program			0.65	0.4207	1.00
Substance use	Yes		6.04	0.0140	0.67
Mental health needs	Unknown		4.78	0.0288	0.61
Mental health needs	Yes		4.97	0.0258	1.46
Program*Time			8.26	0.0041	1.00
Age	15.33		6.78	0.0092	0.88
Days since program ended	2028.89		262.65	<.0001	1.00

Note. Mental health needs, discharge status, days in program, days since program ended, and non-reference group are denoted by MHN, PO, DIP, DSPE, and NRG, respectfully. NRG2 is the additional reference group when an interaction is modeled.

Days in program

Model performance. This model had several significant predictors (Table 7) of days in program; however, the model overall did not predict very well based on the lower R^2 and AUC statistics, (df = 8) = 105.83, $p < .0001$, $R^2 = .13$, AUC = .61. Stated differently, the major concern with this model is its rather poor predictive accuracy and stability over time. For this model, no data were censored because all youth had an official program end date.

Table 7. Parameter estimates and test statistics for days in program model.

Predictor	NRG1	NRG2	χ^2
Model 6: Time to Days in program			
Program	CC		7.83
Discharge status	Successful		45.76
Race	Black		0.00
Race	Hispanic		1.98
Gender	Female		3.61
Age	15.33		15.60
Program*Time			14.32
Discharge status*Time			18.70

Note. Mental health needs, discharge status, days in program, days since program ended, and non-reference group are denoted by MHN, PO, DIP, DSPE, and NRG, respectfully. NRG2 is the additional reference group when an interaction is modeled.

Program-related effects. The significant program variable effect (HR = 0.61, $p = 0.0051$) indicated the hazard rates differed significantly across the programs. However, this finding should be interpreted with caution due to the significant program-by-time interaction ($p = 0.0002$). A visual inspection of the hazard rates reveals CC youth had better survival rates early on (up to about 150 days), whereas ISP youth perform slighter better after 150 days.

The discharge status variable also had mixed findings, as the direct effect of discharge status ($p < 0.0001$) and the interaction between discharge status and time ($p < 0.0001$) were significant. After examining the hazard rates, the results indicate youth classified as having a successful discharge status had better survival rates up to about 200 days in the program, with the magnitude of these differences decreasing and tapering off after about 200 days.

Background variable effects. As seen in Table 7, both gender and race were significant predictors of the days in program variable. While there was no difference in hazard rates between Black and White youth (HR = 1.00, $p = 0.9703$) or between Hispanic and White youth (HR = 0.83, $p = 0.1589$), there was a significant difference between Black and Hispanic youth (1.20, $p = 0.0287$), with Hispanic youth having a hazard rate 1.20 times greater than that of Black youth. The results also show males had a hazard rate 1.20 times greater than that of females (HR = 1/0.83 = 1.20, $p = .0575$), but this difference was not significant at the .05 level.

DISCUSSION

The study results provide several beneficial findings that could be used to aid practitioners and policymakers in the juvenile justice system. The first finding of this study is that discharge status always had a significant effect (either directly or moderating) on the time to event variables. From these findings, it is clear practitioners and organizations should make every effort to ensure youth are on track for a successful discharge and able to receive and/or participate in the programing within a juvenile probation department. This includes continually communicating with probation offices to identify youth who are at risk of an unsuccessful discharge so as to understand what factors and/or behaviors need to be avoided and supported in order for the youth to earn a successful discharge.

A second finding was the number of days in the program was a significant predictor of successful outcomes for all but one model (i.e., time to violation of court orders). These results imply the longer a youth remained in a program, the better the outcome. Clearly there is a fine line between keeping the youth in a program longer and ending the program before the youth receives the services they need to be successful after leaving the program. If the youth is kept in the program longer, this increases the cost to treat the youth. However, ending the program before the youth receives the needed services (e.g., mentoring, social support, meaningful interaction, decision-making skills) influences how successful they are after leaving the program.

When examining the program effect of CC vs. ISP across the five outcome variables (excluding Model 6, days in the program), the findings are mixed. After controlling for the other variables in the model, there was not a significant difference between CC and ISP in the time to second

program (Model 1) or the time to next offense (Model 4) variables. In other words, the program effect was not significant when considering the direct effect or any interaction effects with other variables. However, there was a significant direct effect of program on the time to violation of court order variable (Model 5), with CC youth on average having more days until they violate their court order compared to ISP youth.

Model 2 (Time to detention) and Model 3 (Time to placement) displayed more complex interpretation of their program effects, as the program effect depended on (or interacted with) other variables in the model. For the time to detention model, the program-by-discharge status interaction suggested there was no significant difference in time to next detention between CC and ISP youth when the discharge status was classified as successful, but CC youth had longer times until their next detention than ISP youth when their discharge status was considered unsuccessful.

For the time to placement model, the program effect was even more difficult to interpret because the program effect depended on their history of previous violence and discharge status. The program-by-discharge status interaction indicated that while CC always had better hazard rates when compared to youth in ISP, the difference in hazard rates was greater when the discharge status was classified as unsuccessful. The program-by-violence interaction again demonstrated CC youth consistently performed better than ISP youth, but this program effect was greater when the youth had a history of violence.

Although of secondary interest, this study did provide several interesting and useful model predictions using the background variables of gender, race, age, previous history of a felony, violence, mental health needs, and substance use. As seen in Models 1 through 5, the following variables or their interactions were included as significant predictors in 80% or more of the models: gender (80%), age (100%), and mental health needs (80%). Race (60%), history of a felony (0%), history of violence (40%), and substance use (40%) were less frequently included in the models, but still at notable rates. From these collective results, it is evident gender, age, and mental health needs were more often a significant predictor than a history of violence, a felony, or substance use. As a result, programs should dedicate more attention to these variables when developing models (e.g., youth with mental health needs may require more services or supports if they are to have similar levels of success as youth without mental health needs) and conduct research and program evaluations to better understand how these variables influence success.

From a general model performance perspective, our AUC statistics indicate the models (with the exception of the days in program model) frequently performed worse immediately following the completion of the program, and then continued to improve performance thereafter. These results suggest that predicting the model's success is more difficult immediately after the youth leaves the program. Consequently, other variables should be added to future models to improve model predictions immediately following program completion. For example, it is possible that other variables (such as social support of friends and family, dedication to turning their lives around, and environmental conditions) play a critical role in improving outcomes as time passes after their program is completed. Also, youth may have difficulty with transitions until routines and structures become familiar.

Although Model 6 (the days in program model) had the least predictive accuracy, discharge status was the best predictor of days in program. These results indicated youth with a successful program outcome were in the program longer (129 days on average) than youth with an unsuccessful program outcome (97 days on average). While the program effect was significant, the program-by-time effect renders this effect less intuitive as the program effect changes over time. From a descriptive perspective, there was also no mean difference between the programs due to the PSM. Instead, the demographic variables of race, gender, and age were significant predictors of the number of days in the program.

Interestingly, none of the risk variables (e.g., history of violence, felony, substance use, mental health needs) were significant predictors of the number of days the youth remained in their program. In some ways, this finding is rather concerning, given demographic variables cannot be altered and many of the risk variables were deemed good predictors of future success based on the results in Models 1 through 5.

Limitations

One significant limitation of this study is that researchers only had access to data after January 1, 2007. It is unknown what additional programs and services the youth might have received before that date, prior to their involvement in CC or ISP. This study does not consider any previous programs the youth might have been involved in prior to the study. CC or ISP were the first program the youth received after January 1, 2007, so this study is really a test of the program effect independent of previous youth programs. Since the outcome variables were time-to-event variables, any programs the youth might have received after going through CC or ISP is largely irrelevant. Consequently, if a second program was received after their time in CC or ISP, it would be captured in the *time to a second program* variable and would be an indication of failure, as in detention, another placement, another offense, or violated their court order, had already occurred and their time-to-event was already measured in this study. Instead, statistically equivalent groups were created using PSM to ensure the groups were nearly identical in terms of risk, background characteristics, and number of previous programs. Moreover, this study assumed any previous programs the youth might have received were inadequate in addressing their needs, since the youth eventually ended up in a subsequent intensive program - either CC or ISP.

Another significant limitation of this study is that there were no program characteristic variables included in the model, nor was there an accurate measure of program dosage. Specifically, our models did not consider the actual number of contact hours, the quality of the relationship between program staff and the youth, the length and quality of therapy, or the quality and amount of interaction between the youth and their caseworkers. Moreover, this study did not include any measures of treatment/program integrity; thus, the degree of variation between youth experiences in the programs is unknown and not accounted for in the models.

ABOUT THE AUTHORS

Erin M. Espinosa, PhD (Corresponding Author)
Senior Researcher
National Council on Crime and Delinquency
520 3rd Street, Suite 101
Oakland, CA 94607
EEspinosa@NCCDglobal.org

Dr. Espinosa is a senior researcher with the National Council on Crime and Delinquency. She has published numerous articles on the intersection between juvenile delinquency, mental health, and service delivery systems.

Dan Sass, PhD
Director, Statistical Consulting Center and Associate Professor
Department of Management Science and Statistics, University of Texas at San Antonio
One UTSA Circle, BB 4.06.08
San Antonio, TX 78249
Daniel.sass@utsa.edu

Dr. Sass the Director of the Statistical Consulting Center and Associate Professor in the Department of Management Science and Statistics at The University of Texas at San Antonio. In addition to his methodological/statistical research, he is also dedicated to applied collaborative research projects.

Johanna Creswell Báez PhD
Manager of Course Development and Adjunct Assistant Professor
School of Social Work, Columbia University
1255 Amsterdam Avenue
New York, NY 10027-5927
Jc2515@columbia.edu

Dr. Creswell Báez is the Manager of Course Development and an Adjunct Assistant Professor at Columbia University School of Social Work with the Online Campus. Johanna earned her masters from Columbia University School of Social Work in 2006. In 2015, she earned her PhD in clinical social work from Smith College in Northampton, Massachusetts.

Cassandra Harper, MPH
Advanced Evaluation Specialist
Southwest Key Programs

Ms. Harper was the Data Integration & Research Specialist at Southwest Key Programs.

REFERENCES

Beck, V., Ramsey, R., Lipps, T., & Travis, L. (2006). Juvenile diversion: An outcome study of the Hamilton County, Ohio, Unofficial Juvenile Community Courts. *Juvenile and Family Court Journal 57*(1), 1–10.

Benda, B. B., Corwyn, R. F., & Toombs, N. J. (2001). Recidivism among adolescent serious offenders: Prediction of entry into the correctional system for adults. *Criminal Justice and Behavior, 28*(5)*,* 588–613.

Bishop, D. M., & Decker, S. (2006). "Punishment and Control: Juvenile Justice Reform in the USA." In J. Junger-Tas and S. H. Decker (eds.). International Handbook of Juvenile Justice. New York: Springer, 3–35.

Bouchard, J., & Wong, J. (2018). Examining the effects of intensive supervision and aftercare programs for at-risk youth: A systematic review and meta-analysis. *International Journal of Offender Therapy and Comparative Criminology, 62*(6), 1509-1534.

Cocozza, J., Veysey, B., Chapin, D., Dembo, R., Walters, W., & Farina, S. (2005). Diversion from the juvenile justice system: The Miami-Dade Juvenile Assessment Center Post-Arrest Program. *Substance Use and Misuse 40*(7), 935–951.

Colwell, B., Villarreal, S., & Espinosa, E. (2012). Preliminary outcomes of a pre-adjudication diversion initiative for juvenile justice involved youth with mental health needs in Texas. *Criminal Justice and Behavior, 39(*4), 444–460

Cuellar, A., Markowitz, S., & Libby, A. (2004). Mental health and substance abuse treatment and juvenile crime. *Journal of Mental Health Policy and Economics, 7*(2), 59–68.

Cuellar, A., McReynolds, L., & Wasserman, G. (2006). A cure for crime: Can mental health treatment diversion reduce crime among youth? *Journal of Policy Analysis and Management, 25*(1), 197-214. doi: 10.1002/pam.20162.

Duke, N., Pettingell, A., McMorris, B., & Borowsky, W. (2010). Adolescent violence perpetration: Associations with multiple types of Adverse Childhood Experiences. *Pediatrics 125,* 778–787.

Hamilton, Z., Sullivan, C., Veysey, B., & Grillo, M. (2007). Diverting multi-problem youth from juvenile justice: Investigating the importance of community influence on placement and recidivism. *Behavioral Sciences and the Law, 25,* 137 – 158.

Harris, P., Lockwood, B., Mengers, L., & Stoodley, B. (2011). Measuring recidivism in juvenile corrections. *OJJDP Journal of Juvenile Justice 1*(1), 1–16.

Hockenberry, S., & Puzzanchera, C. (2015). *Juvenile court statistics 2013* (National Report Series Bulletin). Washington DC: US Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention.

Hoge, R. (2016). Application of precharge diversion programs. *Criminology & Public Policy 15*(3), 991–99.

Hyatt, J., & Barnes, G. (2017). An experimental evaluation of the impact of intensive supervision on the recidivism of high-risk probationers. *Crime & Delinquency 63*(1) 3 – 38.

Leve, L. & Chamberlain, P. (2005). Association with delinquent peers: Intervention effects for youth in the juvenile justice system. *Journal of Abnormal Child Psychology 33*(3), 339–47.

Loeb, Roger C., Waung, M. & Sheeran, M. (2015). Individual and familial variables for predicting successful completion of a juvenile justice diversion program. *Journal of Offender Rehabilitation 54*(3), 212–237.

Mackin, J., Lucas, L., Lambarth, C., Herrera, T., Walter, M., Carey, S., & Finigan, M. (2010). *Baltimore County Juvenile Drug Court Outcome and Cost Evaluation*. Portland, OR: NPC Research.

McAra, L. & McVie, S. (2007). "Youth Justice? The Impact of System Contact on Patterns of Desistance from Offending*." European Journal of Criminology 4*, 315–45.

Mears, D., Kuch, J., Lindsey, A., Siennick, S., Pesta, G., Greenwald, M., & Blomberg, T. (2016) Juvenile court and contemporary diversion: Helpful, harmful, or both? *Criminology and Public Policy 15(*3), 1–29.

Puzzanchera, C. M., & Kang, W. (2008). In Schwalbe, C. S., Gearing, R. E., Mackenzie, M. J., Brewer, K.B., & Chechanover, S. A meta-analysis of experimental studies of diversion programs for juvenile offenders, *Clinical Psychology Review, 32,* 22-33.

Taxman, F. S. (2002). Supervision: Exploring the dimensions of effectiveness. *Federal Probation, 66,* 14-27.

Taylor, J., Kemper, T. S., Loney, B. R., & Kistner, J. A. (2009). Recidivism in subgroups of severe male juvenile offenders. *Psychology, Crime and Law, 15*(5)*,* 395–408.

Trulson, C. R., Marquart, J. W., Mullings, J. L., & Caeti, T. J. (2005). In between adolescence and adulthood: Recidivism outcomes of a cohort of state delinquents. *Youth Violence and Juvenile Justice, 3*(4)*,* 355–387.

Vincent, G., Paiva-Salisbury, M., Guy, L., & Perrault, R. (2012). Impact of risk/needs assessment on juvenile probation officers' decision making: Importance of implementation. *Psychology, Public Policy, and Law, 18*(4), 549 – 576.

Wilson, H. A., & Hoge, R. D. (2013). Diverting our attention to what works: Evaluating the effectiveness of a youth diversion program. *Youth Violence and Juvenile Justice, 11(*4), 313-331.