Colorado Evaluation & Action Lab
UNIVERSITY OF DENVER

A strategic research partner for government agencies and a bridge to the research community

Expanding Multisystemic Therapy to Underserved Regions

Final Results of a Pay for Success Project

REPORT HIGHLIGHTS:

- The MST Expansion Program extended Multi-Systemic Therapy for youth ages 12-17 to regions of the state where the service was previously unavailable.
- Treatment under the MST Expansion Program was associated with a 15.6% reduction in youth out-of-home placements in a group setting, which resulted in a success payment of $41,769.
- Despite having been implemented with fidelity, treatment through the MST Expansion Program was not associated with a reduction in youth entering secure detention with new charges as the reason detained.
- A criteria for being included in this evaluation was completion of a baseline risk assessment administered when youth entered secure detention prior to MST treatment.
- The COVID-19 pandemic substantially reduced the number of youth entering secure detention in 2020 and 2021, dramatically reducing the number of youth who could be included in the evaluation. The small sample size increases the likelihood that the findings reported here are due to random chance.

AUTHOR:

Kristin Klopfenstein, PhD
Director, Colorado Evaluation and Action Lab

For inquiries contact: Kristin Klopfenstein | kristin@coloradolab.org |
www.ColoradoLab.org

Report Number: 18-01A. Date: March 1, 2023

Acknowledgements

Thank you to Infinite Frontier Consulting for their tireless support in wrangling the data and in developing and executing a rational matching process. This evaluation was supported by the Annie E. Casey Foundation. We thank them for their support but acknowledge that the findings and conclusions presented in this report are those of the author alone, and do not necessarily reflect the opinions of the Foundation, the State of Colorado, or other project partners. Policy and budget recommendations are the opinions of the Colorado Lab authors and do not represent the budget or legislative agendas of state agencies, the Governor's Office, or other partners. Any requests for funding or statutory changes will be developed in collaboration with the Governor's Office and communicated to the legislature through the regular budget and legislative processes.

Data Sources

This study relies on outcome data from the Colorado Department of Human Services, Divisions of Youth Services and Child Welfare.

Suggested Citation

Klopfenstein, K. (March 2023). *Expanding Multisystemic Therapy to Underserved Regions*: *Final Results of a Pay for Success Project.* (Report No. 18-01A). Denver, CO: Colorado Evaluation and Action Lab at the University of Denver.

Introduction

Multi-Systemic Therapy (MST) is a community-based intervention targeting youth ages 12-17 who are involved or at high risk of involvement in the juvenile justice system. It is delivered with careful fidelity by trained teams of therapists over the course of 3-5 months with multiple home-visits per week. It has a proven track record of success at reducing youth recidivism.

This project extended MST to regions of the state where the service was previously unavailable. The Center for Effective Interventions (CEI) housed at the University of Denver's Graduate School of Social Work[i] provided implementation support as six new MST teams rolled out across the state from January 2019 through July 2020. Teams operated through designated community-based mental health agencies.

Pay for Success

> **Pay for Success (PFS) is an innovative contracting model that ties funding for social programs to evidence-based programming and positive outcomes.**
>
> **Despite having been implemented with fidelity, treatment through the MST Expansion Program was not associated with a reduction in youth entering secure detention with new charges as the reason detained.**
>
> **The MST Expansion Program was associated with a 15.6% reduction in youth out-of-home placements in a group setting, resulting in a success payment of $41,769.**

The PFS model is an innovative approach to financing evidence-based programs that shifts risk from traditional funders—typically a government entity—to private investors who provide the up-front capital. Key outcomes, or "success measures," are agreed upon prior to the start of a rigorous independent evaluation. Only if the evaluation shows that the program meets these outcomes does the government funder repay the initial investment.

In 2015, the Colorado General Assembly passed House Bill (HB) 15-1317, which authorized the Governor's Office of State Budgeting and Planning (OSPB) to enter into PFS agreements with lead contractors for the provision of program-eligible interventions (CRS 24-37-403). HB 18-1323, a Joint Budget Committee bill signed into law in April 2018, provided full funding to cover all direct payments and maximum possible success payments for these projects, through a series of annual transfers into OSPB's PFS Contracts Fund.

The MST Expansion Program was selected in 2018 through an open, competitive process by OSPB to receive implementation funds financed through a hybrid PFS approach. In this case, the implementation costs and initial risk were shared between the state of Colorado and the investors. Through that PFS contracting process, the key outcomes ("success measures") were defined, and it was agreed upon that based on the level of success the project demonstrated, OSPB will pay investors back the money they invested in the implementation of the program, plus a maximum of a 2% return.

[i] CEI became integrated into the Kempe Center for the Prevention and Treatment of Child Abuse and Neglect at the University of Colorado Anschutz Medical Campus starting in July, 2021.

Because MST is well established as an evidence-based program, partners selected an interim fidelity measure as the lead measure for success (Schoenwald et.al. 2003). The Therapist Adherence Measure - Revised (TAM-R) measures fidelity to MST and was administered to families by an independent call center in the first two weeks of treatment and monthly thereafter. The TAM-R contains 28 items that assess the primary caregiver's perception of the extent to which their therapist focuses on specific MST treatment elements. Each item is rated on a 1 (not at all) to 5 (very much) scale. Caregivers can ask to not respond to a particular item or indicate that it is not an issue. In this case, these responses are not calculated as part of the score. The adherence score is calculated by the number of items rated as adherent (i.e., a 5) divided by the number of items that can be scored. Thus, adherence scores can range from 0-1 with a score of .61 considered the threshold for fidelity. Adherence scores for a therapist are calculated by generating an average score for each family, then averaging across all families served by that therapist during a specified time period. For the purpose of success payments, fidelity was monitored at the team level by averaging the TAM-R scores of the four therapists on each team.

The Colorado Lab began reporting TAM-R scores approximately 1 year after each team started up and approximately annually thereafter. For the three interim success payments, all teams exceeded the fidelity threshold of 0.61 so maximum payments were distributed to investors. The final success measures around entering secure detention or out-of-home-placement, which determine the final success payment, are the subject of this evaluation and are described below. The total amount of success payments from the State for this project cannot exceed $1,336,600. This includes the success payments for qualifying TAM-R Scores, already paid out at the maximum of $668,300, plus one success payment for reductions in secure detention and one success payment for reductions in out-of-home placements, each of which cannot exceed $915,571.

Description of the Study

This report represents the contractual analysis required for the final success payment, as outlined in the PFS contract. The purpose was to estimate the impacts of the MST Expansion Program on youth recidivism, defined as entering secure detention with new charges as the reason detained, within 18 months of treatment assignment, and secondarily, on out-of-home placement in a group setting over the same time period.

> **Research Questions for the Final Success Payment:**
>
> 1. What is the impact of the MST Expansion Program on the probability that youth enter secure detention on new charges within 18 months of starting treatment relative to similarly situated youth who did not have access to MST?
>
> 2. What is the impact of the MST Expansion Program on the probability that youth enter an out-of-home placement in a group setting within 18 months of starting treatment relative to similarly situated youth who did not have access to MST?

To answer these questions, the study used nearest neighbor propensity score matching to identify a "business as usual" comparison group similar to the treatment group receiving MST. Matching variables included gender, ethnicity (African-America, Anglo-American, Hispanic/Latino, Other), Colorado Juvenile

Risk Assessment (CJRA) Prescreen Risk Level (low, medium, high), age when took CJRA, and date when they took the CJRA (particularly important given COVID).

The study included MST teams working out of Colorado Springs, Denver, Grand Junction, Greeley, and Pueblo who served youth residing in 10 of Colorado's 22 judicial districts (1,2,3,4,8,17,18,19,20,21). Health Solutions (serving JD10 and JD3) and North Range Behavioral Health (serving JD19) launched MST Expansion teams in January 2019. In July 2019, Hilltop (serving JD21) and Savio House (serving JD17, which also offers services to youth in adjacent counties) launched teams as well. In July 2020, Four Feathers (serving JD4) launched its first team and Health Solutions launched a second team. The first treatment youth included in this analysis started receiving services in February 2019 with outcomes tracked through August 2020, while the last youth started receiving services in April 2021 with outcomes tracked through October 2022.

Starting in March 2020, COVID had a dramatic impact on the work of expansion teams, especially those launching in July of that year. MST was developed and intended to be delivered as an in-person intervention for families and has only been validated with in-person service delivery. Consistent with this approach, all training and support for therapists was developed and intended to be delivered in person. However, the majority of this project, including program delivery as well as the follow-up time for families, occurred during the pandemic. The vast majority of the supervisors and therapists were new to MST at the start of this project and were trained remotely and started serving families remotely without the benefit of having ever delivered the program in person.

As a response to COVID, Governor Polis issued an executive order in March 2020 and lasting through June 2021 that reduced the number of youth detention beds available statewide. Significant juvenile justice reforms during the course of the study also decreased the detention population to all-time lows with the introduction of more permanent detention bed caps and enhanced efforts to divert youth away from detention. While the expansion teams were able to continue to enroll participants, the proportion of evaluation-eligible participants declined substantially below expectations. This is because the evaluation required completion of the CJRA as a baseline measure of risk for youth to be included in the study, and youth only completed the CJRA prior to entering secure detention. Without a baseline measure demonstrating the equivalence of treatment and comparison groups, matching studies generally fail to replicate the results of randomized controlled trials, the gold standard for estimating causal impact. Thus, having a baseline risk measure for both treatment and comparison youth was non-negotiable for this study, even though requiring it led to a substantial reduction in sample size. Multiple alternative approaches to obtaining measures of baseline risk were considered by the Pay for Success Operating Committee. Despite heroic efforts to find an alternative risk assessment or proxy measure of risk, nothing was as consistently available for both treatment and comparison youth as the CJRA.

Evaluation-eligible youth were primarily identified for MST services through a partnership between the MST provider, the community mental health center, and local Colorado Youth Detention Continuum (CYDC) SB 94 coordinators. Juvenile diversion and probation stakeholders served as additional sources of support for the prioritization of referrals. In practice, however, most of the youth referred for services by the MST Expansion Program came from child welfare. During the study, MST was formally added to the federal Family First Prevention Services Clearinghouse as a "well supported" intervention, presenting a new opportunity for states to draw down federal funding for deployment of the service to child welfare-involved families at risk of experiencing out-of-home placement. This federal support enhanced awareness of MST among child welfare agencies and generated new referrals for teams. While this allowed expansion

teams to maximize caseloads and expand services benefiting families, these referrals were not eligible for the evaluation due to the lack of a baseline risk assessment.

TAM-R Score Success Payments

The TAM-R Score success payments were based on fidelity of implementation of the MST model. Throughout the project, MST Expansion Program teams reported success with remote delivery, and the TAM-R scores showed that program delivery met adherence and fidelity benchmarks. The maximum amount of TAM-R Score success payments were made on three occasions (see Table 1) because each team of therapists met the minimum threshold TAM-R Score of 0.61. The remaining two success payments are based on youth outcomes. Since the maximum amount was paid out under the TAM-R Score success payments, $668,300 has the potential to be distributed under the final two outcome-based success payments.

Table 1: Interim Success Payments for Fidelity of MST Program Implementation

Period under Review#	TAM-R Scores Reviewed Through	State Payment Deadline	Success Payment Amount
First 2 MST teams to launch, CY2019	December 31, 2019	April 1, 2020	$305,190
First 4 MST teams to launch, CY2020	December 31, 2020	April 1, 2021	$305,190
All 6 MST teams funded through MST Project, CY2021	September 30, 2021	January 1, 2022	$57,920

Reduction in Secure Detention Success Payment

The Reduction in Secure Detention success payment is based on the estimated impact of MST treatment on the secure detention outcome for treatment youth relative to comparison youth divided by the mean rate of secure detention for the comparison group. This calculation yields the percentage difference in the probability of returning to secure detention for a treatment youth due to participating in MST. See the methods section for more detail. Per the Pay for Success contract, the outcome was assessed without regard to statistical significance.

The success payment was determined according to the business rules outlined in Table 2. MST was associated with an increase in the probability of returning to secure detention-- the opposite of the expected direction. The reduction in secure detention in the sample was -30%, which falls under the category “Less than 10%” in Table 2 below. Therefore, the funders are not eligible for a Reduction in Secure Detention Success Payment. This leaves $668,300 remaining to potentially be paid out under the Reduction in Out-of-Home Placement Success Payment discussed in the next section.

Table 2: Criteria for Reduction in Secure Detention Success Payment

Percentage Threshold	Payment Per Percentage Point at or above Minimum Threshold
Less than 10%	$0
10% through 28%	$35,174
29% through 35%	(19 x $35,174) + ($35,324 per percentage point above 28% to maximum of 35%)
Greater than 35%	Max payment $915,571

Reduction in Out-of-Home Placement Success Payment

The Reduction in Out-of-Home Placement success payment is based on the estimated impact of MST treatment on out-of-home placement in a group setting for treatment youth relative to comparison youth divided by the mean rate of out-of-home placement for the comparison group. This calculation yields the percentage difference in the probability of out-of-home placement in a group setting for treatment youth associated with participating in MST. See the methods section for more detail. Per the Pay for Success contract, the outcome was assessed without regard to statistical significance.

The success payment was determined according to the business rules outlined in Table 3. The percentage reduction in the probability of entering an out-of-home placement in a group setting was calculated as 15.6%. Therefore, the Reduction in Out-of-Home Placement Success Payment is $41,769.

> The MST Expansion Program yielded a 15.6% reduction in the rate at which youth were placed out of home in a group setting relative to the comparison group.

Table 3: Criteria for Reduction in Out-of-Home Placement Success Payment

Percentage Threshold	Payment Per Percentage Point at or above Minimum Threshold
Less than 15%	$0
15% through 30%	$41,769
31% through 40%	(16 x $41,769) + ($24,727 per percentage point above 30% to maximum of 40%)
Greater than 40%	Max payment $915,571

Methods

Statistical Model

The impacts of the MST Expansion Program as articulated in research questions 1 and 2 were estimated using linear regression. Due to the smaller than expected sample size and the high quality of the matching, we minimized the number of covariates included in the model. Thus, the final model estimated was:

$$y = \beta_0 + \beta_1 MST + \beta_2 Age + \beta_3 Female + \beta_4 CJRA\_High + \beta_5 CJRA\_Mod + \epsilon$$

where
RQ1: y = 1 if enters secure detention with new charges as reason detained during follow up period
RQ2: y = 1 if out-of-home placement in a group setting during follow up period
Both RQs:
MST = 1 if assigned to MST treatment, 0 otherwise
Age = age in years
$Female$ =1 if female, 0 if male
$CJRA\_High$ = 1 if CJRA score indicates high risk (relative to low risk)
$CJRA\_Mod$ = 1 if CJRA score indicates moderate risk (relative to low risk)
ϵ = error term

Outcome Measures

Entering Secure Detention with New Charges as the Reason Detained

Entering secure detention with new charges as the reason detained was obtained from the detention extracts in the Colorado Department of Human Services Division of Youth Services. The reason detained in the detention extract was required to be either "pre-adjudicated" or "probation violation-new charges." When the reason detained was missing, collateral information from the Trails data base was used to determine the reason detained. If there was a new offense within 7 days of the detention, the detention was considered to be due to new charges.

Entering Out-of-Home Placement in a Group Setting

Entering an Out-of-Home Placement in a Group Setting was obtained from the child welfare data system. The following were considered group settings: Group Center Care, Psychiatric Care (which includes Psychiatric Residential Treatment Facilities), Qualified Residential Treatment Program, or Residential Child Care Facility. The vast majority of youth who entered an out-of-home placement in a group setting during the follow-up period were in Residential Child Care Facilities. The next most frequent setting was Group Center Care.

Criteria for Success Payments

Pay for Success Outcomes related to RQ1 & RQ2 will be based on the estimated impact of the intervention, regardless of statistical significance. The impact is defined as the percentage change from baseline. Since linear regression coefficients reflect changes in percentage point terms irrespective of baseline, they must be transformed to reflect percentage change from baseline. We accomplish this by dividing the impact estimate from the linear regression by the comparison group outcome mean for both the "Percentage Difference in Probability of Secure Detention" and the "Percentage Difference in Probability of Out-of-Home Placement" according to the equation below.

$$\beta_1 / \mu_{Control}$$

where

β_1 = impact estimate of MST treatment
$\mu_{Control}$ = mean rate of detention (RQ1) or OOH placement (RQ2) for the comparison group

Description of the Sample

The study population includes youth ages 12-16 who have a valid CJRA pre-screen score from within the previous six months prior to referral and who meet the criteria for MST. While the MST program accepts 17-year-olds, for the evaluation youth needed to be less than 17 years of age at the time of MST enrollment to allow for adequate follow-up time in the juvenile system to which we had data access. The CJRA pre-screen is required when a youth enters secure detention, therefore, the population will have entered secure detention within the previous six-months.

MST programming inclusion criteria:
Youth must exhibit at least one of the following characteristics to be eligible for MST. Providers, with oversight from CEI, ensured that youth are appropriately selected.

- Youth behavioral characteristics such as:
  - Violent behavior
  - Crimes against person, property, and/or drug related crimes
  - Status offenses (e.g., curfew, underage drinking)
- Legal characteristics such as:
  - Current out out-of-home placement or at-risk of out-of-home placement
  - Multiple or first-time convictions
  - Multiple police contacts
- School involvement characteristics such as:
  - Expelled or dropped out of formal education or multiple suspensions for problem behavior
  - Poor relationships with prosocial peers and/or school staff
  - Academic problems with risk of expulsion or failure
- Community Peer relationship characteristics such as:
  - Gang membership or strong affiliation
  - Low affiliation with prosocial peers
  - High affiliation with antisocial peers

MST programming exclusion criteria:

- Youth living independently or in long-term residential settings
- Youth in need of crisis psychiatric hospitalization/stabilization or youth who are actively suicidal, homicidal, or psychotic
- Youth for whom a sex offense is the primary reason for referral
- Youth with autism (other therapies, treatments and approaches may be a better fit)

Matching

All youth referred for treatment who completed the initial intake to determine eligibility and were assigned to a therapist were included in the treatment group even if they did not complete a full MST treatment dose. Youth referred for treatment with an MST expansion team were matched to youth who were not referred for MST but met the inclusion criteria for referral, were comparable in terms of demographics and CJRA pre-screen risk score, and resided in a comparable judicial district as described below.

Because youth can be detained by judicial districts distinct from their homes and county of residence, youth were selected for inclusion in the comparison group based on the following process:

1. All secure detention records were reviewed to find all youth who may have ever been living in one of the target counties served by the different expansion teams at the time of a detention admission.
2. All records were reviewed to find all youth who may have ever been detained by a target JD.
3. The final dataset from which to select appropriate comparison youth for each MST provider, includes:
    a. Youth who have ever resided in the target area at the time of a detention, regardless of detention JD.
    b. Youth who have ever been detained by the target JD and are most likely geographically eligible for services.

Youth were excluded from the comparison dataset based on the following:

1. If it was determined that the youth, despite being detained by a target JD, had never lived in an area served by the MST Expansion Program. This is to prevent inclusion of comparison youth who are from regions of the state that are not comparable to those served by the MST Expansion Program.
2. If the youth has already participated in MST. This is determined in three ways, repeated at each of the matching procedures: 1) reviewing services paid for through CYDC 2) reviewing clients served by other agencies providing MST services and 3) reviewing all Pay for Success client enrollments.

Baseline Equivalence of Demographic Measures

Youth were exact-matched on gender, ethnicity, and baseline CJRA risk level, while differences in treatment and comparison youth's age when they took CJRA and the date when they took the CJRA were minimized. Youth often have multiple CJRAs, so for treatment youth, age at CJRA was determined using the CJRA that most closely preceded their enrollment in MST service. For comparison youth, age at CJRA was the youth's age at the time of the CJRA taken within 6 months of their MST counterpart's CJRA. To address potential developmental differences, we selected youth who were as close in age as possible when they took the CJRA. We also minimized the time between CJRA administration dates for treatment and comparison youth. This was important given the varying context with COVID.

Whenever possible, treatment youth were matched to comparison youth in the same judicial district (64 percent of youth). When that wasn't possible, they were matched to comparison youth from a neighboring judicial district (30 percent of youth). Barring that, they were matched with a comparison youth from a judicial district where the treatment youth had some historic connection—having ever lived in or been detained by a county in that district (6 percent of youth).

The sample for whom outcomes were measured at the 18 month follow up included 80 treatment and 80 comparison youth. For reference, the target sample size at the outset of the study was 816 youth, 408 treatment and 408 comparison. The sample demographics for variables where exact matching occurred is presented in Table 4.

Table 4. Demographic Characteristics of Sample for Variables Where Exact Matched

Exact Matched#	n = 80 treatment, 80 identical comparison
Gender	
Male	84%
Female	16%
Ethnicity	
African-American	8%
Anglo-American	50%
Hispanic/Latino	41%
Other	1%
Baseline CJRA Risk Level	
Low	24%
Moderate	40%
High	36%

As continuous variables, baseline equivalence for age at CJRA and date of CJRA were calculated using the Hedge's g. The average age at CJRA of treatment youth was 15.36 years and of comparison youth was 15.38 years, with a pooled standard deviation of just over one year. The effect size for this difference was -0.03. The effect size for the difference in the date when the CJRA was taken, measured in days, was 0.02. Effect sizes of less than the absolute value of 0.25 are typically determined to be "equivalent," so the quality of our matching was exceptionally good.

Results After 18-Month Follow Up

Research Question 1: Entering Secure Detention with New Charges as Reason Detained

During the 18 month follow up, 20 of 80 comparison youth (25%) and 26 of 80 treatment youth (32.5%) entered secure detention with new charges as the reason detained. After statistically controlling for age, gender, and baseline risk, youth who did not receive MST services were still less likely to return to secure detention with new charges as the reason detained. However, as shown in Table 5, this difference is not statistically significant. The small sample size increases the likelihood that the findings reported here are due to random chance rather than due to a true lack of program efficacy.

Table 5. Results of Secure Detention Regression Model

Variable	Coefficient	Standard Error	p-Value
Assigned to MST	0.076	0.071	0.289
Age at CJRA	0.017	0.034	0.612
Female (rel to male)	-0.180	0.100	0.064*
CJRA High Risk (rel to low risk)	0.205	0.095	0.032**
CJRA Moderate Risk (rel to low risk)	0.175	0.093	0.061*
Constant	-0.132	0.540	0.807

Note: N=160, 80 treatment and 80 comparison; * = p < 0.1, ** = p < 0.05, *** = p < 0.01

Research Question 2: Out-of-Home Placements in a Group Setting

During the 18 month follow up, 20 of 80 comparison youth (25%) and 17 of 80 treatment youth (21.25%) entered out-of-home placement in a group setting. After statistically controlling for age, gender, and baseline risk, youth who received MST services were less likely to enter an out-of-home placement in a group setting. For purposes of the success payment, this amounts to a 15.6% decrease in placements in a group setting and a payment of $41,769. Though this is not a consideration for the success payment, this difference is not statistically significant given the small sample size. As for the detention outcome, the small sample size increases the likelihood that the findings reported here are due to random chance.

Table 6. Results of Out-of-Home Placement Regression Model

Variable	Coefficient	Standard Error	p-Value
Assigned to MST	-0.040	0.064	0.547
Age at CJRA	-0.045	0.031	0.151
Female (rel to male)	0.028	0.088	0.748
CJRA High Risk (rel to low risk)	0.318	0.086	0.000***
CJRA Moderate Risk (rel to low risk)	0.191	0.084	0.025**
Constant	0.742	0.489	0.132

Note: N=160, 80 treatment and 80 comparison; * = p < 0.1, ** = p < 0.05, *** = p < 0.01

Discussion

The question remains why, especially given high fidelity of MST implementation as measured by TAM-R scores, youth served through the MST Expansion Program did not achieve better outcomes, particularly for secure detention. Although this question was not rigorously explored in the course of this evaluation, there are several possibilities supported by theory and conversations the MST support team at the Kempe Center had with therapists.

While MST Expansion Program teams reported success with remote delivery, and the TAM-R scores showed that program delivery met adherence and fidelity benchmarks, it is likely that the stressors and strains of the COVID pandemic, along with the challenges specific to remote delivery, impacted services. It may be that there are aspects of treatment fidelity that aren't readily captured on the TAM-R scores that are important for treatment outcomes. The following examples do not apply to all teams or circumstances but provide suggestive evidence and hypotheses that may be worth exploring in future research.

Adapting to New Technology: Therapists reported that it took over six months to get comfortable with telehealth delivery, with initial comfort levels being fairly low. Some families struggled with the technology required for participation in MST. Therapists would provide technical support the best they could, but technical difficulties reduced treatment time within a given session and may have distracted from the long-term skill development essential to the model in favor of the more immediate "putting out fires." Therapists particularly noted challenges for families where English was not their first language.

Challenges with Consistency: The inability of therapists to be physically present with families during much of the pandemic may have also led to missed opportunities to teach skills. Therapists noted more than the usual number of cancellations due to exposures and quarantines which may have led treatment to feel more disjointed. Therapist availability also impacted consistency of treatment. For example, on one team, everyone was diagnosed with, exposed to, or had a dependent with COVID at the same time. Consequently, families had no opportunity to receive consistent treatment during that period.

Social Distancing: Many of the youth enrolled in MST were being cared for by grandparents or family members with compromised immune systems. Therapists observed that when youth left the home, some caregivers were reluctant to allow them back in the home for fear of COVID exposure. Many of the prosocial activities that are a main focus of treatment-- and important for generalization of gains over time—also were not available during the study period due to COVID. During the protracted period of social distancing recommended by public health officials, there was a limited ability for natural supports to come into the home due to risk for exposure. This can be a critical component of treatment that was widely unavailable for a substantial period of the project.

Unpredictable Return to In-Person School: The pandemic also reduced the ability of therapists to address school-related concerns during the course of treatment. Return-to-school plans were often last minute and ever-changing, which prevented therapists from adequately supporting youth and their families with the transition back. The consequences of this may have been more youth re-engaging with antisocial peers, school attendance problems, and academic challenges.

From a purely statistical standpoint, the findings on youth outcomes should not be interpreted as evidence of MST's efficacy (or inefficacy). The target sample size at the outset of the study was 816, but the final sample included only 160 youth. The difference in the number of treatment and comparison youth who entered secure detention was just six individuals. The lack of statistical significance indicates that this difference was likely driven by random chance, or factors that were not measured in the model, rather than by inefficacy in the MST model. Similarly, the fact that three fewer youth entered an out-of-home placement cannot be attributed to the efficacy of MST.

Five of the six teams launched for this project are still going and have substantially returned to standard in-person MST implementation. The acute COVID response period taught these teams a lot in a short time under incredible pressure. It also provided researchers with interesting implementation questions to explore that will ultimately deepen our understanding of the conditions under which MST is most effective.