The Oklahoma Title IV-E Waiver Demonstration Project: 2015-2019

Final Evaluation Report

David E. Bard, Ph.D.

Erin Maher, Ph.D.

Debra Hecht, Ph.D.

William H. Beasley, Ph.D.

Lana O. Beasley, Ph.D.

Michael Hunter, Ph.D.

Christopher E. Aston, Ph.D.

Andrew Peters, M.S.

Yutian Thompson, Ph.D.

Geneva Marshall, M.H.R.

Som Bohora, M.S.

Maleeha Shahid, M.B.B.S, M.P.H.

Amanda Ferraro, M.S.

Adon Rosen, B.S.

UNIVERSITY OF OKLAHOMA HEALTH SCIENCES CENTER, COLLEGE OF MEDICINE, DEPARTMENT OF PEDIATRICS,

BIOMEDICAL AND BEHAVIORAL METHODOLOGY CORE

AND

SECTION OF DEVELOPMENTAL AND BEHAVIORAL PEDIATRICS, CENTER ON CHILD ABUSE AND NEGLECT

March 31, 2020

Contents

3

Table of Figures

Table of Tables

The Oklahoma Title IV-E Waiver Demonstration Project: 2015-2019

Executive Summary

Background

In 2014, Oklahoma Department of Health Services was awarded a Title IV-E Waiver to spend child welfare dollars on the front end of the service continuum and address the increase in the number of children coming into care. The Intensive Safety Services (ISS) intervention was developed to complement the existing infrastructure of evidence-based home-based services (SafeCare) throughout the state. The goals of this Waiver demonstration project are to increase the safety and well-being of children in their own homes, who would otherwise be placed in foster care.

Eligibility for ISS is determined through a predictive risk model (PREMISS) for families with a child 12 years or younger. The model predicts, based on historical data, children who are 'moderate' risk and would otherwise be removed from the home were this intervention not in place. The purpose of ISS is to allow these families to remain safely together by assisting with sustainable behavior changes in caregivers to eliminate or reduce the reoccurrence of child abuse or neglect. A safety monitor is identified and caregivers' needs are assessed. ISS workers engage caregivers to motivate behavior change leading to abuse or neglect and connecting them to needed services (e.g., mental health, substance abuse, and domestic violence) in order to prevent out-of-home placements and improve child and family well-being.

ISS serves families at imminent risk of removal with a more immediate and intensive response. ISS begins within a week of the referral and eligibility and is a six-week program conducted by MA-level therapists who visit the home at least three times a week for 8-10 hours a week. During the 3rd or 4th week, SafeCare is introduced and continues for 4-6 months as a "stepdown" for the case after ISS is completed. Additionally, the caseloads are small (2-3 families), and improvements can be observed through weekly staffing sheets. ISS was rolled out incrementally in 5 regions of the state, Region 3 (Oklahoma County) commencing in July 2015, and region 4 (Southeast Oklahoma, the last implementing region) commencing in April 2018.

The Evaluation

The evaluation findings presented in this report are based on data from July 22nd, 2015 through October 2nd, 2019. The ideal outcome of the evaluation is to demonstrate the effectiveness of this approach without increasing overall spending for the state. The evaluation is conducted by investigators from the University of Oklahoma Health Sciences Center, Department of Pediatrics, Center for Child Abuse and Neglect and the Department of Sociology.

The mixed methods longitudinal study was designed as a randomized stepped wedge design. Once eligibility is determined, families are assigned to ISS or Services as Usual, though some families deemed eligible for ISS do not receive it due to other factors. Thus, the evaluation contains and compares outcomes across three study groups--ISS received, ISS Not Workable, and Services as Usual--using three approaches: a) As Treated, b) Intent-to-Treat, and c) Treatment on the Treated. Data is collected at baseline, six weeks (step down to home-based services), and six months (end of home-based services). Measures include administrative data (out-of-home placements and re-reports) and standardized measures of safety threats, protective capacities, working alliance, depression, domestic violence, and

child behavior. The qualitative component of the study compares the perspectives of ISS' implementation and effectiveness between the CPS workers, FCS workers, and ISS workers.

Evaluation Findings

The preliminary quantitative results provide great insight into the effectiveness of ISS. ISS resulted in fewer children entering out-of-home care compared to Service as Usual (SAU). Using the Intent-to-Treat approach (explained in "Outcomes Study" section), children eligible for ISS (some of whom received it and some who did not) were less likely to be removed (57%) compared to SAU families (63%). ISS did not reduce the likelihood of additional referrals or removals to CPS compared to SAU. However, the overall numbers of re-referrals and removals across all study groups was low.

Using the Intent-to-Treat approach, there was a small, but statistically significant difference between ISS and SAU groups in terms of reducing safety threats, with ISS having a slightly lower average number of safety threats at six months post-referral. Additionally, ISS families have a greater improvement in protective capacities than SAU. In terms of differences in child well-being and behavior, we find measuring Conflict Scales, all subscales are improving (decreasing), suggesting the program is having a good effect in decreasing levels of conflict in client families. These findings show that parenting concerns reported on Child Well-Being Scale decreased from baseline to 6 months. Using the child functioning outcomes on the Oklahoma behavioral health screener (OK-TASCC), there was not a significant difference in outcomes between ISS and SAU. Also, though we don't have comparison data between ISS and SAU, about 67% of ISS parents scored at-risk for clinical depression at baseline and this dropped to 30% at stepdown and 8% by six months.

Qualitative findings suggest that workers believe ISS helps families through education, engagement, and service enrollment and that the program displays good interactions with the families. Additionally, they see the more intense and expedited ISS services to families as helpful and that ISS builds rapport with families through information supports and has more insight into families. On the other hand, CPS workers feel ISS is too short, and not all families that could benefit from it receive the treatment. Many worry that ISS is too short and starts too quickly. FCS and CPS workers have been extremely pleased with the responsiveness and competence of contracted ISS teams, and appreciate the additional option of keeping children in the home with more intensive services.

The sustainability of ISS depends on the fiscal feasibility and measuring this was a requirement of the demonstration project. The cost analysis of ISS shows that in a short amount of time—three years, the more expensive and intensive intervention model gets closer to being no more expensive than not using it, which reflects the better outcomes that it achieves, offsetting the additional service delivery costs.

While some implementation and evaluation challenges have been experienced, such as barriers in obtaining eligibility data prior to a child safety meeting, lack of universal adherence to the ISS protocol, and some missing data and documentation, overall the evaluation has provided strong evidence that ISS is successfully reducing the number of children birth to twelve who enter foster care while not resulting in more experiences of abuse or neglect or costing the state more money. Some weaker evidence suggests that the intervention is also leading to improved well-being of the children and parents served by ISS.

The Oklahoma Title IV-E Waiver Demonstration Project: 2015-2019

Introduction and Overview

The following report reviews cumulative evaluation activities and findings from the Oklahoma Title IV-E Waiver Demonstration project between July 22, 2015 and October 2nd, 2019. This reporting period marks the 4.25 year operation point in Region 3, Oklahoma County, the 2.75 year mark of activity in Region 1, and the 2.25 year mark for Regions 2 and 5, and the 1.5 year mark for Region 4. The report first summarizes the activities performed, and then reviews the findings. Findings are summarized by regions and for the state as a whole.

Background, Context, and Purpose of the Waiver Demonstration

Prior to the start of this Waiver Demonstration, the State of Oklahoma experienced an increase in the number of children entering foster care that far surpassed the number of children exiting care, increasing the overall number of children in care annually. At the end of SFY 2012, there were 9,132 children in care 0 - 18 years old. By the end of SFY 2014, that number had risen to 11,562, more than a 26% increase. In contrast, the number of children exiting care each year, only increased by 4.1% from SFY 2012 to SFY 2014.

In 2014, the Oklahoma Department of Human Services (DHS) applied for a Title IV-E waiver demonstration project to address the large increase in the number of children in care. This demonstration project was designed to focus on accomplishing the following broadly defined goal:

Increase the safety and well-being of infants, children, youth, and families in their homes and communities, including tribal communities, by supporting families to prevent child abuse and neglect and entry and re-entry into foster care.

Demonstration Interventions and Associated Components

The waiver demonstration project allowed the Oklahoma Department of Human Services (OKDHS) to develop and implement the Intensive Safety Services intervention (ISS). ISS is delivered in the home to families who would, previously, have experienced a child removal. The intent was that this immediate, focused service will allow identified families to safely remain together, and thereby, address the drastic increase in the number of children in care. This Waiver project, is thus, closely aligned with the goals of the Family First Prevention Services Act.

Intensive Safety Services (ISS) utilizes Master's level providers who deliver frequent services in the home for 4 to 6 weeks. DHS contracts with providers for these ISS workers, and cases are assigned to DHS Family Centered Services (FCS) staff who work together on the case. The service array available through ISS allows both sets of workers to support families in the home whose children are at higher risk of removal. The goal is to support the family to allow the child to remain safely in the home.

The philosophy of FCS, stated below, is consistent with that of ISS:

- Focus on the child's safety, preserve and strengthen the protective capacities of the caregiver and maintain the child safely in his or her home, whenever possible;
- Focus on the entire family rather than individuals;
- Provide comprehensive services that engage the family and target the family's therapeutic, supportive, and concrete needs.

The DHS FCS caseworker visits the family weekly, while the contracted ISS worker is in the home three to five times per week. To aid in the assessment of individual family needs, ISS workers are expected to use both individual and family assessment and analysis tools including, but not limited to, the CAGE, the Child Well Being Scales (CWBS), the Center for Epidemiologic Studies Depression Scale (CESD), and the Child Behavioral Health Screener. Based on need as determined by this battery of assessments, ISS workers are able to provide direct services that include: Motivational Interviewing (MI), Cognitive Behavioral Therapy (CBT), and Healthy Relationships. Additionally, the ISS worker is responsible for connecting the family to appropriate community resources and assisting the family with any barriers to accessing these resources.

In all ISS cases, both the DHS FCS caseworker and the contracted ISS worker assess the concrete needs (generally economic) of the family. Special funding for concrete needs is included in the budgets for the ISS contracts. At the appropriate time, usually within 6 weeks, and again, based on the needs of each family, Comprehensive Home Based Services (CHBS), the current family preservation service program within DHS, is provided to the family as a step down from ISS to less intensive services. CHBS is appropriate for families where the children are at a lower risk of removal, thus serving as the service option after ISS has worked successfully with families.

CHBS utilizes the SafeCare model. SafeCare is an evidenced-based curriculum for families who are at-risk for or have been reported for child maltreatment. Contracted workers provide services in the home approximately once a week for a period of up to six months. The introduction of CHBS prior to the ISS termination assures no break in service to the family, and a longer duration of service than most intensive family preservation programs.

ISS provides the following service array, to address, at a minimum:

- Parental depression – The provider conducts Cognitive Behavioral Therapy with parent(s) impacted by depression and refers them to appropriate services.
- Substance abuse and/or dependence – The provider links the parent to appropriate services in the area, using Motivational Interviewing to help ready the parent.
- Domestic violence – The provider assists the family in developing appropriate safety plans for the child and the victim of domestic violence. The provider delivers the Healthy Relationships curriculum and/or links the parent to appropriate services in the area.
- Parenting – Subsequent to family stabilization, parenting skills are assessed and referrals to additional services such as Parent Child Interaction Therapy (PCIT) are made.
- Safety of the home environment — physical hazards of the home environment are assessed and, if needed, appropriate steps are taken to address any concerns.
- Concrete needs – Supports are provided when deemed necessary.
- Parenting, safety, and health

Addressing these risk factors should positively improve the overall safety and well-being of the children in the home.

Implementation of the demonstration project began in Region 3, Oklahoma County, and services have rolled out incrementally to additional DHS regions. As of Fall 2018, ISS was active in all five regions served by DHS.

Target Population

This demonstration project works with families with at least one child between ages 0 - 12 who are at risk of having a child removed from the home. Risk of removal is determined by the use of a predictive

risk model described later in this section. At risk families are often struggling with substance abuse, domestic violence, parental depression, and fulfilling concrete needs. Based on available historical data, parental substance abuse was the second most cited reason for removal of a child and was a contributing factor in a significant number of substantiated investigations (e.g., 46% in 2013). Domestic violence was also a contributing factor in a large number of substantiated investigations (e.g., 26% in 2013). In addition, higher rates of parental depressive symptoms were associated with higher rates of repeated reports of maltreatment. DHS also chose to focus on this population because 84% of all substantiated maltreatment reports and 87-88% of all removals from SFY 2013 through SFY 2014 involved children ages 0 - 12 (see Table 1**Error! Reference source not found.**). Table 2 provides the number of substantiated child maltreatment reports starting a few years before the Waiver project began and through 2017.

Under contract with DHS, the Biomedical and Behavioral Methodology Core (BBMC) of the Department of Pediatrics, University of Oklahoma Health Sciences Center (OUHSC), developed a predictive risk model, PREMISS, to assist in making recommendations for ISS referral. To build the model, statewide data on removals from July 1, 2011 through February 1, 2014 were mined, and predictive analytics were used to forecast expected length of removal. Under the assumption that shorter lengths of removal indicate lower home safety risks, the model outputs a unidimensional risk continuum score that is now used to identify ISS eligible families. Actual ISS eligibility is determined on a per-region basis by setting cutoffs along the PREMISS risk continuum that forecast eligibility counts to match each region's anticipated service capacity. Further details about the construction, purpose, and use of the PREMISS model are described later in this report in section "ISS Selection and Services Flow-Chart".

TABLE 1 - RISING NUMBERS OF CHILDREN IN CARE

SFY	No. children entering care	Ages 0-12	Percentage increase	Female	Male
2012	5,642	87%	N/A	48%	52%
2013	5,859	88%	3.9%	50%	50%
2014	6,275	88%	7.1 %	47%	53%

TABLE 2 NUMBER OF CHILDREN SUBSTANTIATED REPORTS BY YEAR

SFY	No. Children in Substantiated Reports	Ages 0-12	Female	Male
2012	9,627	86.8%	51.4%	48.6%
2013	11,553	86.7%	51.2%	48.8%
2014	13,183	85.9%	50.7%	49.3%
2015	14,449	86.5%	51%	49%
2016	14,308	86.7%	50.3%	49.7%
2017	14,457	86.4%	51.9%	48.1%

Evaluation Plan

Theory of Change

As previously discussed, prior to the beginning of the current demonstration project, the number of children entering Oklahoma's foster care system was consistently increasing, leaving OKDHS overburdened, with too few workers, foster families, or financial resources to be able to sustain the increase in service demand. The increase largely affected families with children between the ages of 0 to 12, where rates of influx have been increasing faster than rates of efflux for a number of years, and the difference in these rates had been on a rapid incline.

DHS and other service providers agreed that a major factor contributing to the increasing rate of children entering care is the lack of preventive services available to address high-risk antecedents of child maltreatment such as parental depression, substance abuse, and domestic violence. Notably, like several other states, the percentage of children removed as a result of substance abuse rose from 36% in SFY 2009 to 49% in SFY 2014, and the number of substantiated investigations in which substance abuse (45% - 48%) or domestic violence (25% - 27%) was a contributing factor to the abuse or neglect has increased each year from SFY 2012 to SFY 2014. In a review conducted by Casey Family Programs, in partnership with The Child Welfare Policy and Practice Group (CWG), an estimated 26% of removal cases in recent years could have been prevented were adequate preventive services available. Yet, even if these services existed, resources at the start of this Demonstration project limited any major expansion of preventive efforts for families at risk of child removal. Given past high-profile child fatalities among families known to DHS and the subsequent Pinnacle Plan reforms (resulting from a class action law suit), it was also widely acknowledged that hot-line workers were becoming less likely to screen out cases and Child Protection Services (CPS) workers were less likely to render a safe-home decision. However, consensus opposes the view that these cultural shifts in system response, while contributing to the growing demand for care, are unhealthy to the system as a whole. Changes in the Hotline referral process, for example, have increased the CPS workload volume in some regions, but these reforms have also established greater statewide consistency in the standards set for investigation and assessment.

Other areas challenging early and accurate assessment of child safety include competing DHS initiatives to expedite case processing, new legislative mandates to investigate particular cases, large numbers of new and inexperienced caseworkers, and inability to quickly engage uncooperative family members in early assessments and case planning. So, while reforms at the investigative stage of case processing were deemed beyond the scope of this demonstration project, it is anticipated that ISS will compensate for some of the known deficiencies in the early assessment of safety by enabling more frequent opportunities for family engagement, a more thorough and ongoing assessment of family strengths and needs, and direct oversight of family compliance with service recommendations. Moreover, separate concurrent initiatives, like the enhanced Child Safety Meetings (CSM) and the Sooners Sentinel Sites Project (SSSP), are aiming to address these and other front-end investigative concerns. Combined, these initiatives and the onset of ISS intend to instill confidence throughout the CPS workforce in the ability of the system to successfully retain more children safely within their own homes.

While taking custody of a child maltreatment victim may often seem the "safer path," the Agency and its partners recognize the host of additional problems this pathway can create. Obviously, the costs of placing and maintaining a child in care are expensive, and as more children enter the system, state and federal dollars for funding these services become extremely limited and the quality of care naturally

suffers. But beyond fiscal tenability and quality of care concerns, removal of a child from a caregiver also registers as additional trauma to the entire family, and many experience long-term difficulties adjusting and coping with this transition. To make matters worse for the children involved, this trauma typically gets repeated as rapid changes in placements are the norm. These children must cope not only with separation from a parent, sibling, and extended family, but also with a cavalcade of environmental changes brought on by new homes, new schools, new peers, and new neighborhoods. For all these reasons, preventable removals inevitably lead to unwarranted distress to the community, the parents, and the children.

DHS and its partners are committed to solving this problem, and the ongoing demonstration project shows significant progress toward the mission of keeping more children safely in their homes. As stated previously, DHS has established and implemented a new Intensive Safety Services (ISS) intervention. This intervention utilizes Master's level ISS workers who partner with FCS workers and provide, as needed, Motivational Interviewing (an empirically supported technique for counseling substance abuse clients), Healthy Relationships (an evidenced informed education module for addressing domestic violence issues), and Cognitive Behavioral Therapy (an evidence-based treatment for depression) in the home, three to five times per week for four to six weeks. The ISS worker also links families with the needed community services for continued support and treatment, and within six weeks, most families are also referred to CHBS to receive a lower intensity program for up to six additional months. By combining ISS and CHBS, DHS now has the ability to keep more children safely in their homes and effectively reduce the number of children coming into care.

When effective, this new service is anticipated to have the intended impact through the following outcome chain:

Intervention: DHS implements Intensive Safety Services (ISS)
So That
DHS offers in-home services that are high in frequency and intensity
So That
ISS workers can respond within 24 hours to support the family
And
Complete assessments to determine the needs, strengths and natural supports of the family
So That
ISS workers develop the treatment plan with the family, youth and DHS to determine goals
And
Goals are prioritized and it is determined which goals must be met prior to step down to CHBS
So That
ISS worker determines what services are provided in-home by ISS worker (e.g., Motivational
Interviewing, Healthy Relationships, Cognitive Behavioral Therapy) and links the family to needed
services in the community for continued treatment
And
ISS worker assists with removing barriers to accessing community services
So That
Family is participating and engaged in services
So That
Parents have increased skills and protective capacities that they can sustain
And
Safety threats are managed
And
Goals are met so family can step down to CHBS
So That
Children remain safely with their families
And
With the combined ISS and CHBS, children are safe from future abuse and neglect and the social and
emotional well-being of children and families is improved

Logic Model

The goal of ISS is keeping families safely together in the home. The benefits of ISS will vary depending upon the interventions used to address the families' specific needs. Depending on the needs of the family and the targeted interventions used, expected results may include a reduction in parental depression; reduction in parental substance abuse; and/or an increase in the caregiver's communication and decision-making skills, thus, reducing the possibility of future domestic violence incidents and relationship conflict. The long-term benefit of ISS is the sustainable behavior changes in caregivers to eliminate or reduce the reoccurrence of abuse or neglect. The ultimate goal is the reduction of children being removed from the home and entering care as well as reduction of the total number of children in care. Figure 1 portrays the logic model for ISS.

ISS = Intensive Safety Services, MI = Motivational Interviewing, CBT = Cognitive Behavioral Therapy, CSM = Child Safety Meetings, SSSP = Sooners Sentinel Sites Project, OK-TASCC = Oklahoma Trauma Assessment and Service Center Collaborative, tx = treatment, SA = substance abuse, DV = domestic violence

DHS developed ISS to target identified risk factors and address those factors through comprehensive services evidence-based and evidence-informed interventions. As shown in The goal of ISS is keeping families safely together in the home. The benefits of ISS will vary depending upon the interventions used to address the families' specific needs. Depending on the needs of the family and the targeted interventions used, expected results may include a reduction in parental depression; reduction in parental substance abuse; and/or an increase in the caregiver's communication and decision-making skills, thus, reducing the possibility of future domestic violence incidents and relationship conflict. The long-term benefit of ISS is the sustainable behavior changes in caregivers to eliminate or reduce the reoccurrence of abuse or neglect. The ultimate goal is the reduction of children being removed from the home and entering care as well as reduction of the total number of children in care. Figure 1 portrays the logic model for ISS.

Figure 1 above, these changes in the service system are expected to have positive direct effects on the quality and breadth of family assessments, the receipt of concrete supports, the number of families receiving services for depression, substance abuse, and interpersonal conflict, the quality and breadth of individualized service plans, and the number of families successfully engaged in the CHBS parenting intervention. These direct outputs should have immediate impact on the Agency's understanding of family risk and need, the quality and expediency of case decision-making, and the capacity to effectively and more expeditiously address and manage a broader array of safety threats. As families progress through ISS, utilize referrals to other services, and stepdown to CHBS, intermediate outcomes should

include improved parenting skills (e.g., effective discipline practices, better understanding of developmental needs, etc.), reductions in risk factors (e.g., depressive symptoms, substance abuse, and conflict), accelerated reductions in safety threats (as measured by the DHS Assessment of Child Safety, AOCS), increased protective capacities (as measured by the AOCS), and improved child well-being (as measured by the Child Behavioral Health Screener), and these improvements should drive the desired longer-term outcomes of child maltreatment reduction (as indicated by a lower number of recurrent CPS events among ISS exposed individuals relative to SAU-exposed individuals), reduced number of initial placements and re-entries into out-of-home care, and reduced financial costs on the Agency. In remaining sections of this document, details are provided for the evaluation components that will track the direct ISS outputs (Process Study), the safety, permanency, and well-being outcomes of the family (Outcomes Study), and the financial costs associated with attaining those outcomes (cost analysis).

Design Overview

Data Collection Methods

Evaluation of the demonstration project utilized multiple data sources (i.e., historical data, administrative data, DHS and ISS worker collected data, staffing and program fidelity data, parent report and child data) and both quantitative and qualitative data collection methods. These methods are used to inform a Process Study, an Outcomes Study, and the costs associated with the implementation of ISS.

Outcome Data

Administrative Data - Data on child abuse and neglect referrals, the initial Assessment of Child Safety (AOCS) administered by DHS and ISS workers, which assesses protective capacities and safety threats, and case characteristics and demographic information about the family and child, were extracted from the Oklahoma state SACWIS system, KIDS. These data were stored on secure SQL Server tables housed on the OUHSC campus.

Worker collected information: additional data were collected by ISS workers, FCS workers, and/or CHBS workers and included subsequent AOCS forms, Child Well-Being Scales, depression screeners (BDI and CESD), and drug and alcohol screeners. These data were collected at baseline, stepdown (at the conclusion of ISS timeframe – typically 6 weeks), and six-months. Additionally, at the stepdown and six-month time-points, satisfaction surveys were collected from clients, and Working Alliance Inventory questionnaires were collected from both workers and clients.

Process Data

ISS and FCS workers provided staffing and fidelity checklists weekly for the duration of ISS services. These data were collected and stored within the Research Electronic Data Capture (REDCap) system. Additional process measures were collected from administrative data, qualitative interviews, and supplemental quantitative surveys with DHS staff involved in either the SAU or ISS case flow and/or service delivery, ISS service contractors, and members of the Core Waiver Demonstration Team (comprised of DHS program administrators, finance directors, and quality assurance officials as well lead representatives from the program contractors, consultants, and evaluation).

Design

We approached the evaluation of the Demonstration project from the perspective of "local molar (pragmatic, atheoretical) causal validity" (p. 69, Campbell, 1986)[1]. Here "molar" refers to all project activities and interventions (see Logic Model above) which will be examined in totality and not in isolation. In other words, due to the nature of the design and lack of tight controls on implementation across regions, the internal validity will be focused on effectiveness of the local (region-specific) treatment package and not on any single "active ingredient" (Maxwell & Delaney, 2004).[2] Our evaluation questions are pragmatic and aim to assess, simply, whether the local treatment package has an impact on defined outcomes. The treatment package in this instance refers to the introduction of new ISS workers and services which will be paired with already existing services, like CHBS, provided through the Family Centered Services (FCS) program.

A randomized multilevel interrupted time-series (stepped-wedge) design with two experimental conditions, Services As Usual (SAU) vs. Intensive Safety Services (ISS) is being used in the study. This design is complex and deserves further explanation. The multilevel structure involves clients, nested within providers, who are nested within regions. Oklahoma DHS has divided the state into 5 Regions, and within each region, groups of counties comprise districts. The exception to this is Region 3, which is comprised solely of Oklahoma County and contains six sub-districts. In this implementation, we are manipulating the experimental conditions (SAU vs. ISS) at the district (or sub-district) level. Manipulating the experimental condition at the district level has several advantages, both practical (it does not require a district to randomize either its providers or its clients nor to initiate the intervention widely at the very onset, thus, reducing implementation burden), and methodological (it manages potential problems with cross-contamination and is less vulnerable to individual provider effects). This choice does have statistical power and analytic drawbacks (relative to a simple client-level randomization design), but these concerns are mitigated by the state-wide nature of the proposed implementation and the large number of districts available for study. The within-subjects and interrupted time-series aspect of the design refers to the fact that both of the conditions will be applied to all of the participating districts, but in a staggered fashion. This type of design is similar to a single-subject multiple-baseline design. Each district will receive both SAU and ISS, but on varying schedules. For example, one district might start with SAU, then switch to ISS shortly thereafter. Another might begin with SAU and switch to ISS much later in time. The design is strengthened by randomizing sequence assignment. Within every DHS region, there will be three possible sequence assignments for each district: early, mid, or late-year ISS implementation (i.e., point at which the switch from SAU to ISS occurs). The implementation point for each district is determined by a blocked randomization procedure that uses approximately balanced assignment permutations (i.e., blocking is used to ensure large imbalances in the number of districts at each of the three switch points do not occur). Finally, partly due to capacity limitations and partly due to enhanced rigor, clients within an implementation district/sub-district are also randomized to receive (or not receive) ISS services. An adaptive randomization procedure is used to determine whether an eligible client in an active implementation district/sub-district is assigned to ISS. The probability of this assignment depends on the number of available slots at the moment eligibility is determined.

Effectiveness of ISS will be examined by stratifying on region and examining intra-region differences among ISS and SAU districts. Because of the longitudinal aspect of the design, two-thirds of the districts

---

[1] Campbell, D. T. (1986). Relabeling internal and external validity for applied social scientists. *New Directions for Program Evaluation*, *1986*(31), 67-77.

[2] Maxwell, S. E., Delaney, H. D., & Kelley, K. (2017). *Designing experiments and analyzing data: A model comparison perspective*. Routledge.

(those assigned to mid or late-year transition points) will also serve as their own control, enabling examination of pre-ISS and post-ISS outcome change. Effectiveness within and between districts will be based solely on those families deemed eligible for ISS services. As described earlier, eligibility designations will be based on predictive risk models that predict which families are most likely to benefit from early in-home ISS.

Oklahoma considered a standard two-group randomized design for the demonstration, but felt the interrupted time-series (stepped-wedge) offered a number of practical and scientific advantages. Like a simple two-group design, the latter design allows direct comparison of the two intervention conditions. Similar to a simple two-group randomized trial, reasonable causal conclusions can be drawn from this type of design particularly when the sequence is randomized rather than self-selected or naturalistic.[3] Because the design uses each district as its own control, it may offer advantages for dealing with the site-to-site variation in outcomes often seen in field effectiveness studies. There are pragmatic advantages as well. Because every district will receive ISS at some point, the design better fits the quality assurance needs and mandates of the regional agencies and state planners. Finally, it is a more cost-efficient design. Because ISS is staggered, a maximum number of districts delivering ISS at one point in time is highly controlled. This (sequential) roll-out design is more feasible compared to an implementation plan that assigns half the state all at once to the newly devised service system. Given the number of consultant and support personnel involved in delivering ISS, the simple two-group design would not be feasible for the budget allowed unless the number of district sites was reduced. Because Oklahoma is interested in testing ISS in a scaled-up application across as many diverse regional settings as possible, we preferred the design that could accommodate a larger number of sites.

The evaluation of the Demonstration project can be classified into three components: process (or implementation) outcomes, effectiveness outcomes, and cost outcomes. Each of these classifications and the operationalized measurement of outcomes within class are described in turn below. But first, details are provided regarding the randomized sequential roll-out of ISS and the determination and selection of ISS eligible families.

Sampling and Randomization Plans

The ISS intervention has been rolled out sequentially across geographic settings, starting with sub-districts in Region 3 (Oklahoma County) and gradually deploying to new regions and districts (within regions) over time. Start dates for ISS services within regions have been randomized and systematically varied in three to six month time-intervals in a stepped-wedge[45], staggered fashion. As such, the proposed evaluation uses a hybrid effectiveness-implementation paradigm[6] that explores the effectiveness of an intervention (the intent of the outcomes study below) while also collecting data on the quality of implementation (part of the process study below). Although this paradigm may sacrifice

---

[3] Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference/William R. Shadish, Thomas D. Cook, Donald T. Campbell*. Boston: Houghton Mifflin.

[4] Handley, M.A.;Schillinger, D.; and Shiboski, S. 2011 *"Quasi-Experimental Designs in Practice-based Research Settings: Design and Implementation Considerations"* The Journal of the American Board of Family Medicine Sep 2011, 24 (5) 589-596; DOI: 10.3122/jabfm.2011.05.110067

[5] Hussey, M. A., & Hughes, J. P. (2007). Design and analysis of stepped wedge cluster randomized trials. *Contemporary clinical trials*, *28*(2), 182-191.

[6] Curran, G. M., Bauer, M., Mittman, B., Pyne, J. M., & Stetler, C. (2012). Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. *Medical care*, *50*(3), 217.

some aspects of tight experimental control and measurement, there are compensatory advantages in terms of ecological validity and the extent to which findings generalize to and are relevant for actual field practice.

Two randomly chosen sub-district units within Region 3 (Oklahoma County) began the implementation of the demonstration. After several months, another two randomly chosen sub-district units within Region 3 were selected and so on throughout Year 1. The sequence of sub-district roll-out was designated at the beginning of a region's implementation year by the BBMC evaluation team using a block randomization procedure. DHS had both child welfare staff and contracted staff in Region 3 trained and ready to deliver services by July 22, 2015.

After year 1, new implementing regions were selected based on ongoing assessments by the Core Team as described in the original IDIR, which is available upon request. To date, services have begun in all five DHS regions. Start dates for these remaining regions were in January (Region 1), July (Region 5), and August (Region 2) of 2017 and April of 2018 (Region 4).

ISS Selection and Services Flow-Chart

In an attempt to control the type and volume of families coming into ISS, districts now institute new case flow processes as demonstrated in Appendix B. Central to this change was a new predictive risk model, PREMISS, which is used to determine ISS eligibility for all families in an implementing region. Archived administrative data were used during the Demonstration's developmental period to inform the PREMISS model which was built to predict which children, between the ages of 0 to 12, historically experience shorter lengths of placements that resulted in successful reunification. More specifically, demographic and case characteristics as well as safety assessment fields from the Assessment of Child Safety (AOCS) instrument were input in a logistic regression model using a cross-validated elastic net optimization algorithm[7] to predict reunification within six months of a child removal. The original Region 3 model was trained using all removals occurring in the State between the dates of July 1, 2011 and December 31, 2012, and that model output a risk score that was dichotomized using a single cut-off to determine ISS eligibility.

Predicting Eligibility through Predictive Analytics and the PREMISS Model

Using a cut-off that selected enough eligible families to meet demands of the randomized design and anticipated service capacity in Region 3 (i.e., 147 families with 294 children[8] receiving ISS in year 1), the model's performance was also evaluated using a test data set that included all removals from Region 3 between the dates of January 1, 2013 and February 1, 2014. Performance metrics comparing the accuracy of the PREMISS decisions to a random selection model and to a model that solely relies on initial CPS worker assessments of safety[9] are provided in Table 3 below. Note, the PREMISS model is designed to behave like a screening measure that limits the number of false negatives and results in a high sensitivity (TPP) and a high negative predictive value (NPV). The PREMISS model outperforms the random selection model on all metrics, and although, the worker decisions produced higher positive predictive value and specificity (TNP) values, this type of selection mechanism would result in too few clients (a projected 77 per year) being served through ISS.

---

[7] Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software*, *33*(1), 1.

[8] These numbers were based on current removal estimates and an anticipated capacity to serve 20% of that count.

[9] In a small number of instances, CPS workers' recommendations to keep the children within the home were not followed.

TABLE 3. PERFORMANCE COMPARISONS FOR POSSIBLE ISS SELECTION METHODS

	PREMISS Selection	Random Selection	Worker Selection
Served Count	441	441	77
PPV	0.13	0.06	0.41
NPV	0.99	0.94	0.97
TPP	0.86	0.41	0.48
TNP	0.62	0.59	0.95

PPV=Positive Predictive Value; NPV=Negative Predictive Value; TPP=True Positive Proportion=Sensitivity; TNP=True Negative Proportion=Specificity

PREMISS decisions only apply to cases where CPS workers/supervisors have recommended child removal. Workers and supervisors are informed of these requirements during trainings that occur prior to implementation. Because of the randomized, stepped-wedge evaluation design, not all PREMISS recommendations for ISS will result in receipt of the new services. Receipt of ISS will require both a PREMISS risk score that falls in the range of eligibility (above the selected cutoff) and assignment of the case to an ISS implementing district unit. Assignment of cases to the district units involves block randomization using an automated system designed by the evaluation team. When the PREMISS score does lie in the eligibility range but the assigned district is not currently implementing ISS, these referrals are assigned as SAU cases. Finally, PREMISS decisions can be overturned in instances where the CPS staff and CSM facilitators determine there are no appropriate safety monitors available; however, these cases are still considered ISS condition participants for purposes of the Intent-to-Treat evaluation.

Evaluation Status

Evaluation Timeframe and Implementation Status
The following section reviews cumulative evaluation activities and findings from the Oklahoma Title IV-E Waiver Demonstration project between July 22, 2015 and October 2nd, 2019. As a reminder, this reporting period marks the 4.25 year operation point in Region 3, Oklahoma County, the 2.75 year mark of activity in Region 1, and the 2.25 year mark for Regions 2 and 5, and the 1.5 year mark for Region 4.

The following list summarizes evaluation activities during this final six month reporting period:

- Routine inspection of PREMISS predictive fidelity continues for the duration of the reporting period
- Error testing code continues to monitor data outliers, missing data, adverse events, and unusual findings
- Transcription of interviews with Administrators from regions 1-5
- Analysis of interviews with Administrators from regions 1-5
- Development of new qualitative interview guides for follow-up evaluation post-Waiver
- Data exchanged and tested for final report
- Assistance with the Program Semi-annual report submitted by OKDHS

- Meetings with DHS continued to discuss project outcomes and the sustainment of ISS following the end of the Demonstration Project.
- Automated code for quantitative data of the process and outcome study reports continues to evolve for expanded reporting needs
- Evaluation staff maintain management responsibilities for data collection reminder system for FCS and PP workers
- Improved modeling approach for all multivariate models
- New modeling approach for safety threats and protective factors
- Finalizing report to ACF for March 2020 submission
- ISS evaluation post-Waiver is underway in Region 3 and 5

High Level Milestones since the Start of the Waiver Evaluation

Since the start of the evaluation, we have accomplished the following major milestones:
- Evaluated 4.5 years of ISS
- Completed cost analysis
- Briefings with OK DHS leadership and state representatives
- Presentations and participation at grantee meetings
- Presentation at the National CW evaluation summit
- Contributed to the sustainability of the ISS post-Waiver
- Preparing for Family First Prevention Services Act review
- Manuscript preparation

Challenges to Implementation of the Evaluation and Planned or Enacted Resolutions

There have been a few major implementation challenges related to the evaluation plan to this point. The biggest challenge to date was an unexpected delay in the finalized versions of data sharing agreements between the evaluator's institution (OUHSC) and DHS. This problem was compounded when evaluators learned in April, 2015, that two separate data sharing agreements would be required: 1) a renewal of a previous agreement covering other OUHSC-DHS projects and amended to include process, outcome, and cost measures for this Waiver Demonstration; and 2) a new agreement covering the SACWIS fields requested as possible inputs to the evaluator's PREMISS eligibility model. Both agreements were submitted for review in May, 2015. Both institutions requested a number of language revisions in an iterative review process that lasted approximately 6 months. Copies of the final versions of the two agreements were received on November 12, 2015 and December 23, 2015, respectively. The late date of approval limited evaluator's ability to reliably monitor and track program outcomes during the early phase of implementation. The late arrival of the data sharing agreements have had minimal to no impact on the evaluation plan. While data exchanges were started much later than planned, evaluation staff have fully caught up with data entry tasks.

Another obstacle encountered related to the implementation of the PREMISS eligibility model. The model was embedded within the State's SACWIS system, KIDS, on July 21, 2015, and this enabled an automated alert system to inform the Core Implementation Team when new cases met the model's eligibility criteria. While this system of detection and alerts has operated without major technological problems to date, there have been a number of feasibility problems with serving all families identified by PREMISS through the ISS model. The most frequent problem encountered has been the involvement

of the court system in a case prior to a PREMISS eligibility notification. DHS has determined, for the moment, that once the courts are involved in a case, ISS is not a practical option for that family (i.e., it would require major changes to legal proceedings). These cases, along with other impractical assignments (e.g., no safety monitors identified) and a moderate number of ISS refusals, have occasionally resulted in significant gaps between ISS worker caseloads and worker capacity, and this has negatively affected the contracted ISS agency's ability to sustain financial viability.

Tracking of outcomes has been difficult under the originally designed procedures. Previously, evaluation team members traveled to DHS office locations to recover necessary data. Coordination of these data visits and tracking of files that were filed on older cases have caused problems in terms of timing and data recovery. To overcome these issues, DHS has provided select evaluation team members with OKDHS email accounts so that DHS staff can scan and forward required data electronically. This new procedure for data exchange began in January, 2016, and marginally improved the fractions of missing data. The impact of the late arriving data sharing agreements on the maintenance of the PREMISS model have also been limited. When issues arose prior to the agreement, problems were quickly resolved through brief troubleshooting calls and meetings.

Another data collection obstacle was caused by delays in IRB approval for the qualitative portions of the evaluation's process study. Most of this delay was intentional, as evaluators chose to wait for the new innovation system to stabilize and for routines to get established before devising meaningful interview probes. Once program stability was evident, interview guides were constructed (see Appendix), and on November 3, 2015, these guides were submitted for IRB approval. Final IRB approval for these additions was received on December, 1, 2015. By this juncture, qualitative data collection for the 1st reporting period was inhibited due to scheduling difficulties in the month of December.

Delays to the construction and approval of the full qualitative protocol did not pose major impediments to the evaluation timeline. The interviews previously planned for the 2nd quarter of the project were scheduled for the first 45 days of the third quarter. By March 14, 2016, all qualitative data collection, coding, analysis, and reporting targets were back on schedule; however, the 3rd and final round of qualitative data collection from the 1st implementation region (Region 3) was intentionally postponed to free up more time for transcribing and coding data from rounds 1 and 2 and to wait for the local ISS contracting agency to fill a provider vacancy. The latter vacancy was filled in the fall of 2016, and the contracted agency requested interviews be delayed until early 2017. Qualitative staff completed Region 3's year 1 collection in May of 2017. These 3rd round interviews have been coded and analyzed and are incorporated into this evaluation report.

An additional change to the qualitative evaluation plan was that we elected not to conduct interviews with court personnel. Given that the project has not focused heavily on court involvement, the investigator team determined those resources would be better spent getting additional input from administrative leadership on policy and practice recommendations for the final report. We also tried to reallocate resources towards the possibility of interviewing parents served by ISS in order to include family voice. We were not able to secure the permissions needed to do so in time for this final report, but maintain a commitment in the follow-up evaluations to attempting to.

As stated above, there have been case flow problems to date related to the PREMISS model case selection procedures. In the majority of the cases where court involvement preceded a PREMISS output, the PREMISS model was not run until after a child safety meeting (CSM) had occurred and an initial decision to pursue child removal had been made. This finding seemed to indicate that retraining of

front-line CPS staff on the PREMISS procedures (which required inputs prior to a CSM) could increase the number of PREMISS selected cases deemed practical for ISS.[10] DHS agreed and conducted re-trainings of CPS units during the months of October and November, 2015. Small improvements (i.e., reduction in rate of court involvement cases prior to PREMISS runs) were noted early by the evaluation staff, but these improvements were not sustained. This issue created problems not only for case flow, but also for analysis. To combat case flow concerns, evaluators initially adjusted the PREMISS selection cutoff to identify more families as eligible for ISS. While this strategy partly addressed concerns over unfilled ISS worker case openings, it also meant a handful of "riskier" families were now getting assigned to and served by ISS. Neither the evaluators nor DHS were comfortable sustaining this practice once caseload capacity was reached, so the original PREMISS cutoff was quickly re-instituted. A second strategy was then installed which did require an adaptation to the assignment procedures.

Previously, the assignment probabilities for SAU and ISS services for those deemed eligible by PREMISS were a function of the number of unfilled ISS case openings (i.e., adaptive randomization is being used). To address the concerns above, this mechanism was slightly modified to, in essence, overbook ISS with two more cases than capacity. In instances where overbooking actually results in too many families accepting ISS, the ISS contracting agency agreed to assign those cases to their acting ISS supervisor, which happened rarely if at all. The most effective solution to date, however, has been the involvement of the Region 3 District Director. Frustrated by the limited progress of CPS protocol compliance following re-trainings, the Core Waiver team enlisted the support of the District Director to spread the word on the importance of providing SACWIS data to PREMISS in time for an eligibility determination. The involvement of the District Director produced immediate gains in compliance, but these have not been sustained to date. Capacity issues continue to be a cyclical concern.

As stated above, the impractical ISS assignments also impact the planned analysis. The planned intent-to-treat analysis will view all these impractical assignments as though they received the ISS intervention. To account for the obvious bias this will have on the (under) estimated treatment effect, evaluators have decided to conduct a subgroup analysis that better reflects the realities of this trial implementation. Because court involvement is tracked within the State's SACWIS records, evaluators can, for example, run subgroup analyses that exclude both ISS and SAU assigned cases who have experienced court involvement prior to a PREMISS notification. Beginning with this report, these subgroup analyses and complier average causal treatment effect models are examined to better estimate the impact of ISS receipt (with or without CHBS) on outcomes.

Implementing regions continue to struggle to submit the PREMISS data prior to the occurrence of the Child Safety Meetings and court involvement (see summaries in later section), and ISS worker capacity issues continue to occur periodically. Staffing has been a minor challenge over the last couple years for a variety of reasons. In January 2017, due to major funding cuts at DHS, the evaluation team was asked to subsume the FCS and PP data collection reminder responsibilities. A new tracking system was developed for this purpose, and an individual from the evaluation team was assigned this new role. On the whole, this transition has been smooth; however, the evaluation staff member assigned these duties did leave the team in the summer of 2017 to pursue graduate studies full-time. Fortunately, a new position was

---

[10] It is important to mention that better adherence to the PREMISS input procedure may not increase the number of practical ISS cases. There is a strong possibility that late arriving inputs signal instances where the CPS worker delayed initiating PREMISS simply because it was obvious the family had no chance to avoid court intervention.

posted and filled relatively quickly, and the new employee was fully trained before their predecessor left the office. This employee left for her own graduate studies this summer, but again, the evaluators were able to fill this position quickly and train a replacement in time for this round of reporting. The evaluation team did also lose two quantitative analysts and a qualitative analyst over the last 30 months (one due to a spouse's job transition, another to a faculty appointment at Georgia Tech, and the other due to acceptance of a graduate studies program). To accommodate these losses, two existing analysts from the evaluator's methodology core were reassigned to this Waiver Demonstration project, and these new individuals now cover two prior members' reporting and analysis responsibilities. In August of 2018, the team also added a new faculty member to replace the 2nd analyst, and she is now fully participating in the evaluation activities.

Missing data for follow-up information continues to be a problem. In some cases (such as the AOCS), we are able to work with individual offices and track down paper forms that had not been submitted to us in a timely manner and are entering them up until the final results for this report are run. In most cases, the information was simply not collected. The amount of missing data for follow-up forms and assessments are embedded in the individual outcome analyses.

Study Groups

Through this reporting period, 4,851 families were identified by the PREMISS model as eligible for ISS. A total of 614 of these PREMISS eligible families were later discovered to meet exclusion criteria (see categories below), thus, only 4,237 families were randomized to either the Intensive Safety Services (ISS) or Services As Usual (SAU) study conditions. To review, the PREMISS model is a predictive risk model developed by the evaluators during the development phase of this Waiver Demonstration project. The model assigns a risk score on a continuous scale, and a cut-point is used to determine eligibility for ISS (the experimental intervention being tested by the demonstration project). Only cases where a CPS worker recommends child removal are run through the PREMISS model's algorithm. Input from the initial CPS investigation, along with prior CPS history and some general family and case characteristics, are utilized in this algorithm. If inputs are missing once the CPS worker electronically records the removal recommendation, the PREMISS result is also missing. CPS workers were trained on the process and timeline for model inputs during the latter months of the project's development phase and immediately prior to initial ISS roll-out. CPS staff do receive refresher trainings from members of the Core Waiver team and strong encouragement from unit supervisors and the District Directors of all region.

Throughout this report, clients are referred to within groups as described above: ISS and SAU. However, within the ISS group, there is also a sub-group of individuals who were assigned to ISS (meaning, the PREMISS model established eligibility, and the family was randomized to ISS), but they did not actually receive ISS (typically due to logistical issues with the timing of the CSM with respect to the PREMISS run). This sub-group is referred to within the report as "ISS not workable", and the sub-group of ISS that received services is referred to as "ISS received".

Among the 4,859 families identified as eligible for ISS, the following study condition assignments and service outcomes were realized.

- 1,527 referrals were assigned to ISS
  - 529 separate referrals received ISS services. We refer to these cases as the "ISS Received" group throughout this report.

- o 316 ISS referrals and 317 cases closed after all requirements were met
  - 289 of these cases stepped down to CHBS
  - 14 of these cases did not step down to CHBS (note: 1 of the cases from the referral split in two did not step down to CHBS)
  - 152 ISS cases closed without meeting all ISS goals (41 of these cases stepped down to CHBS for a brief period of time before closure).
  - Reasons for closure among these cases are:
    * 41 cases were closed after partial completion of ISS goals * 10 families relocated * 14 ended early at the request of Child Welfare (this included cases where the person responsible for child, PRFC, was no longer available for ISS due to incarceration, long-term hospitalization, inpatient treatment, etc.); * 46 where family was uncooperative, refused services, or did not make progress after intake * 78 cases where the children were removed * 18 cases where a change of custody (adoption/guardianship) occurred * 22 cases still active as of 2019-10-02
- o 996 cases randomized to ISS that did not receive ISS. We refer to these cases as the "ISS Not Workable" group throughout the report. Reasons for not receiving ISS are listed below :
  - Court Intervention – 613
  - Client Refused ISS – 72
  - No Safety Monitors – 80
  - Services Not Needed (safety threats do not necessitate ISS) – 105
  - Services Not Needed (child to adoption/guardianship) – 16
  - Client Withdrawn Due to Safety Concerns – 106
  - Client Not Available for ISS – 32
  - Other – 35 (e.g., involved slow referral actions by a child welfare worker which hampered the start of services)
- 2719 referrals were assigned to services as usual (SAU)
- 615 cases met exclusionary criteria[11] and were NOT randomized to either condition

Study Group Demographics and Characteristics

<mark>Error! Reference source not found.</mark> describes the demographic characteristics of the three study groups (ISS Received, ISS Assigned, SAU), as well as the ITT comparison (ITT Not-Workable and SAU combined)[12]. The table presents results of tests of significant differences between the groups. Data come from all referrals prior October 2nd, 2019. This table also examines significant differences across four sets of comparisons.

Table 4 presents the means and standard deviations of characteristics of the victim, perpetrator, and case. We also test for significant differences between the four study groups and these are indicated next to the variable name with subscripts. Some of the means are significantly different across the comparisons. For instance, mean age of the child victim in Services as Usual is significantly higher than

---

[11] Reasons include: family residence not in an ISS-assigned DHS Region; children already victims in an open case; family previously received ISS; family participating in another DHS innovation study.

[12] The count for children is an undercount, because some child dates of birth are missing for the "ISS not workable" group (those randomized to ISS who did not receive ISS but are included in this count).

the other study groups, though the difference is small (Cohen d = 0.09). A key indicator of risk is the count of prior reports, and for this, we find that there all small but significant differences between groups, with the ISS received group having slightly fewer previous reports than the other groups. The difference in prior reports between ISS assigned and SAU was also significant, but again, the risk difference was small (Cohen d = 0.14). Similar differences were observed for our PREMISS eligibility scores, which also reflect risk, but these, too, were very small (ISS vs SAU Cohen d = 0.10). Finally, the number of protective capacities for ISS Not Workable was significantly lower than both ISS Received and SAU, suggesting some selection effects for families in ISS Not Workable. Differences between the non-randomized groupings (comparisons involving ISS received and ISS not workable) are not surprising. The small statistical differences reported for randomized groups are small, did not apply multiple testing adjustments (statistically too liberal), and mask substantial distributional overlap between conditions. Group differences in overlaid plots of prior referrals, for instance, are extremely difficult to visually discern due to near identical distributions. Arguably, these tabled central tendency differences improperly infer imbalance as the result of a few outlier cases (with unusually large prior counts) and the high statistical power (due to large sample sizes) affording detection of exceedingly small (not clinically significant) effect sizes. Thus, effect size comparisons are a worthy consideration for all rows.

TABLE 4 DESCRIPTIVE STATISTICS ON VICTIM, PERPETRATOR AND CASE CHARACTERISTICS BY STUDY GROUP AND SIGNIFICANT DIFFERENCES BETWEEN GROUPS

	ISS Assigned Mean (SD)	ISS Received Mean (SD)	Service as Usual Mean (SD)	ISS Not Workable Mean (SD)
Age of Child[a,c,d]	5.71 (4.51)	5.54 (4.47)	6.14 (4.72)	5.81 (4.53)
Count of Prior Reports[a,b,c,d]	1.41 (1.89)	1.25 (1.69)	1.71 (2.33)	1.50 (1.99)
Count of Perpetrators per Referral[b,c]	2.04 (0.84)	1.89 (0.76)	2.09 (0.88)	2.12 (0.87)
Count of Victims per Referral	2.22 (1.33)	2.29 (1.36)	2.19 (1.27)	2.19 (1.31)
Average Age of Perpetrators[a,b,c]	32.6 (7.38)	31.5 (7.08)	33.3 (7.96)	33.1 (7.47)
Count of Allegations per Referral	4.30 (3.57)	4.26 (3.6)	4.23 (3.35)	4.32 (3.56)
Safety Threats	1.24 (0.48)	1.22 (0.44)	1.24 (0.50)	1.25 (0.51)
Count of Protective Capacities per Referral[b,d]	8.09 (7.85)	9.09 (8.13)	8.27 (7.92)	7.54 (7.64)

PREMISS Risk Scores[a,b,c]	-1.51 (0.74)	-1.43 (0.75)	-1.58 (0.71)	-1.55 (0.74)

Note: superscripts denote differences between group means for that variable are statistically significant at the p < 0.05 level: "a" denotes a significant difference between ISS Assigned and SAU; "b" denotes a significant difference between ISS Received and ISS Not Workable; "c" denotes a significant difference between SAU and ISS Received; "d" denotes a significant difference between SAU and ISS Not Workable. PREMISS risk scores reflect the log-odds prediction that this case will experience reunification within 6 months were it to be served with SAU.

Table 5 presents the frequencies and proportions for characteristics of the victims. We find fewer significant differences between study groups than in the previous table. We do find that black children who are victims of maltreatment are significantly less likely to be in SAU than ISS Assigned and ISS Not Workable. The same finding is true for Hispanic victims, though these differences are not large.

In summary, by making these comparisons, while we have a few significant differences between study groups, the differences are small. In addition, these results point to the fact that analyses may want to control for prior reports and protective capacities.

TABLE 5 FREQUENCIES AND PROPORTIONS OF VICTIM DEMOGRAPHIC CHARACTERISTICS BY STUDY GROUPS AND SIGNIFICANT DIFFERENCES BETWEEN THEM

	ISS Assigned Frequency (Proportion)	ISS Received Frequency (Proportion)	Service as Usual Frequency (Proportion)	ISS Not Workable Frequency (Proportion)
Race of Victim: Asian	13 (0%)	2 (0%)	27 (0%)	11 (1%)
Race of Victim: Black[a,d]	785 (23%)	295 (24%)	1098 (18%)	490 (22%)
Race of Victim: Pacific	14 (0%)	5 (0%)	32 (1%)	9 (0%)
Race of Victim: White	2029 (60%)	748 (61%)	2658 (61%)	1281 (59%)
Race of Victim: Unknown	0 (0%)	0 (0%)	1 (0%)	0 (0%)
Race of Victim: Missing[a,b,c]	555 (16%)	167 (14%)	1158 (19%)	388 (18%)
Hispanic Victim[a,d]	786 (23%)	265 (22%)	1240 (21%)	521 (24%)
Child victim is female	1647 (48%)	577 (47%)	2999 (50%)	1070 (49%)

Note: superscripts denote differences between group means for that variable are statistically significant at the p < 0.05 level. "a" denotes a significant difference between ISS Assigned and SAU. "b" denotes a significant difference between ISS Received and ISS Not Workable. "c" denotes a significant difference between SAU and ISS Received. "d" denotes a significant difference between SAU and ISS Not Workable.

Evaluation Findings to Date

Process Study

The Process Study evaluates aspects of implementation including adherence to model fidelity, staff perceptions and attitudes surrounding implementation, and organizational change. Process Study measures are collected through a combination of methods including administrative data, qualitative interviews, and supplemental quantitative surveys with DHS staff involved in either SAU or ISS case flow and/or service delivery, ISS service contractors, and Core Waiver Demonstration team members. In addition to the minimally required process study measures, the evaluation team collects data on the quality of the DHS worker and client relationship, client satisfaction with services, and overall ratings of district-specific implementation phase and fidelity.

Analysis for the process study includes tracking process and implementation measures as listed in the evaluation plan, additional implementation measures surrounding alliance between client and worker, items surrounding staff perception of implementation, and measures of model fidelity. Methods and outcomes are discussed under each topic and qualitative data are used to inform conclusions where appropriate.

ISS Protocol

As we review adherence and fidelity to the model with these process measures, it is important to lay out the desired protocol and model components.

1) The PREMISS model should be run before the child safety meeting. Eligibility for ISS is determined by ISS.
2) ISS therapist must make face to face contact with the family within 24 hours of receiving the referral.
3) The requirements for FCS and ISS workers to visit children and families/PRFCs include[13]:
    - FCS workers visit families and children at least once every 7 calendar days (Table 12).
    - ISS workers visit PRFCs 3 to 5 times every 7 calendar days.
    - ISS workers visit children at least once every 30 calendar days within the ISS intervention period. These workers visit families to see both children and their parents together. There were no expectations to see children alone.
4) ISS is expected to be a six-week duration.

Case Flow Process with Respect to the PREMISS Model

Understanding the case flow and assignment using the PREMISS model is central to the evaluation. Thus, as part of our process evaluation, we track eligibility and note changes over time in the administrative data and assignment to ISS based on PREMISS. We also track time between the referral and PREMISS notification.

Figure 2 describes the number of PREMISS notifications each month by region. The ups and downs in PREMISS eligible notices are likely due to rating changes by CPS staff or case mixture changes leading to more or less eligible families (i.e., an influx of higher or lower risk referrals than in the past) in the

---

[13] These visits are documented on FCS and ISS Staffing Checklists.

region. Periodic, in-depth investigations of possible PREMISS algorithm malfunctions uncovered no issues with the study's automated eligibility system.

In Region 1, the number of PREMISS notices each month varied more in the first year of implementation (2017), but has been fairly stable for the last 2 years. Since 2017, this number has ranged from 15 to 48 with an average of 33.64 (sd = 7.78) days. The duration of time between referral and PREMISS notice averaged 25.8 days and ranged from 0 to 252 days.

Region 2 began implementation in August 2017. The average number of PREMISS notices in Region 2 has been 24.54 (sd = 5.48) with a range of 15 to 34. The time between CPS referral and PREMISS notifications has averaged 19.83 days and ranged from 0 to 127 days.

In Region 3, the number of PREMISS notifications dropped each month during the beginning of 2017, but started climbing up again. The time between the initial referral and the PREMISS notice averaged 17.3 days (ranging from 0 to 128). Despite the differences in notification frequency over time, this average duration between the referral and the PREMISS notification has remained consistent. As of October 2, 2019, the average number of PREMISS notices in Region 3 is 26.23 (sd = 9.89) with a range of 6 to 51. The time between CPS referral and PREMISS notifications has averaged 17.38 days and ranged from 0 to 150 days.

Region 4 began implementation in April 2018. The average number of PREMISS notices in Region 4 has been 24.22 (sd = 5.82) with a range of 10 to 37. The time between CPS referral and PREMISS notifications has averaged 19.9 days and ranged from 0 to 137 days.

Region 5 began implementation in July of 2017. The average number of PREMISS notices has ranged from 17 to 34 with an average of 25.74 (sd = 5.16). The time between CPS referral and PREMISS notification has averaged 16 days and has ranged from 0 to 86 days.

FIGURE 2 MONTHLY PREMISS NOTIFICATIONS
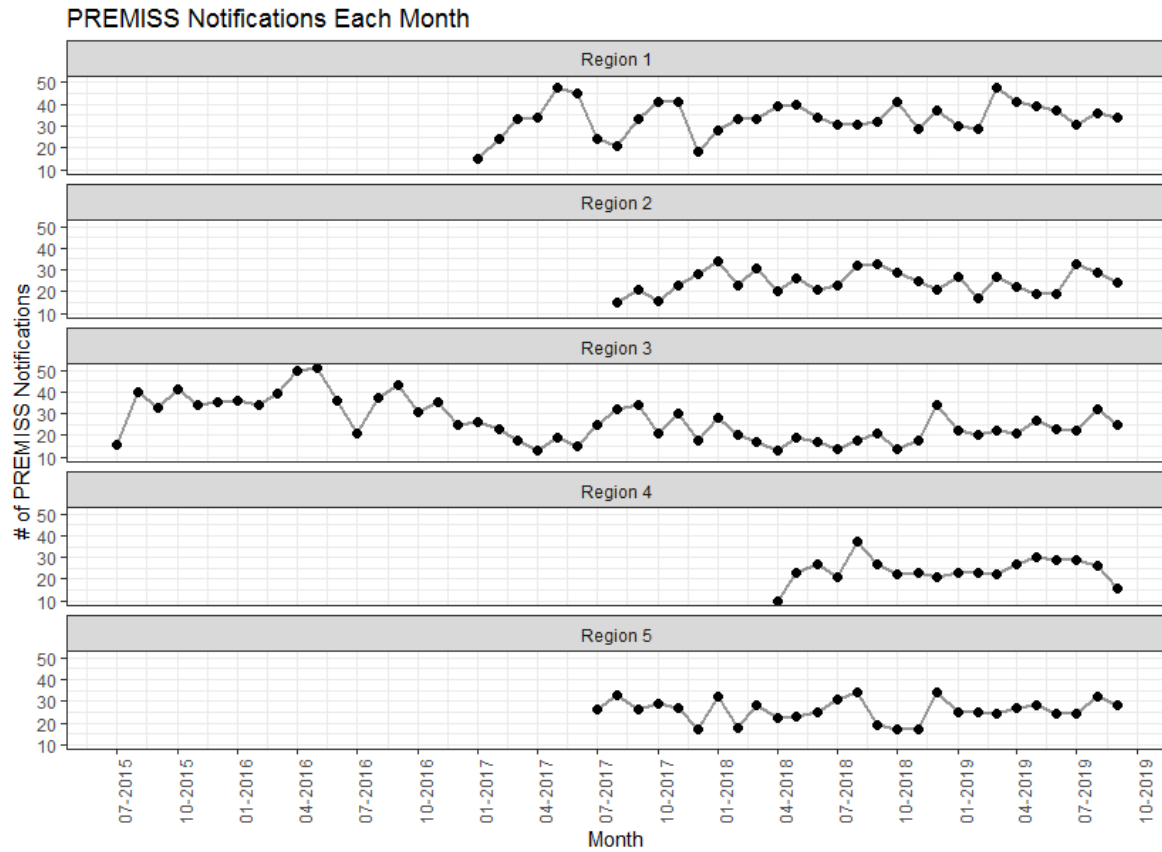
As noted, per protocol, the PREMISS model should be run before a Child Safety Meeting (CSM) takes place. Among those who were determined PREMISS eligible and were known to have participated in a CSM (n = 4,244), PREMISS data were entered, on average, 14.1 (sd = 20), 7.6 (sd = 13.5), 3.8 (sd = 8.2), 9 (sd = 12.3), and 4 (sd = 8.1) days after the CSM in Regions 1, 2, 3, 4, and 5, respectively. Figure 3 shows the fluctuation in the mean difference between the CSM and PREMISS run dates by Region. Averages for all months in all Regions have been in the "wrong" direction with PREMISS data usually arriving after the CSM takes place.

FIGURE 3 MEAN DIFFERENCE IN DAYS BETWEEN CSM AND PREMISS RUN DATES

Mean Difference in Days betwen Child Safety Meeting and PREM-ISS Run Dates
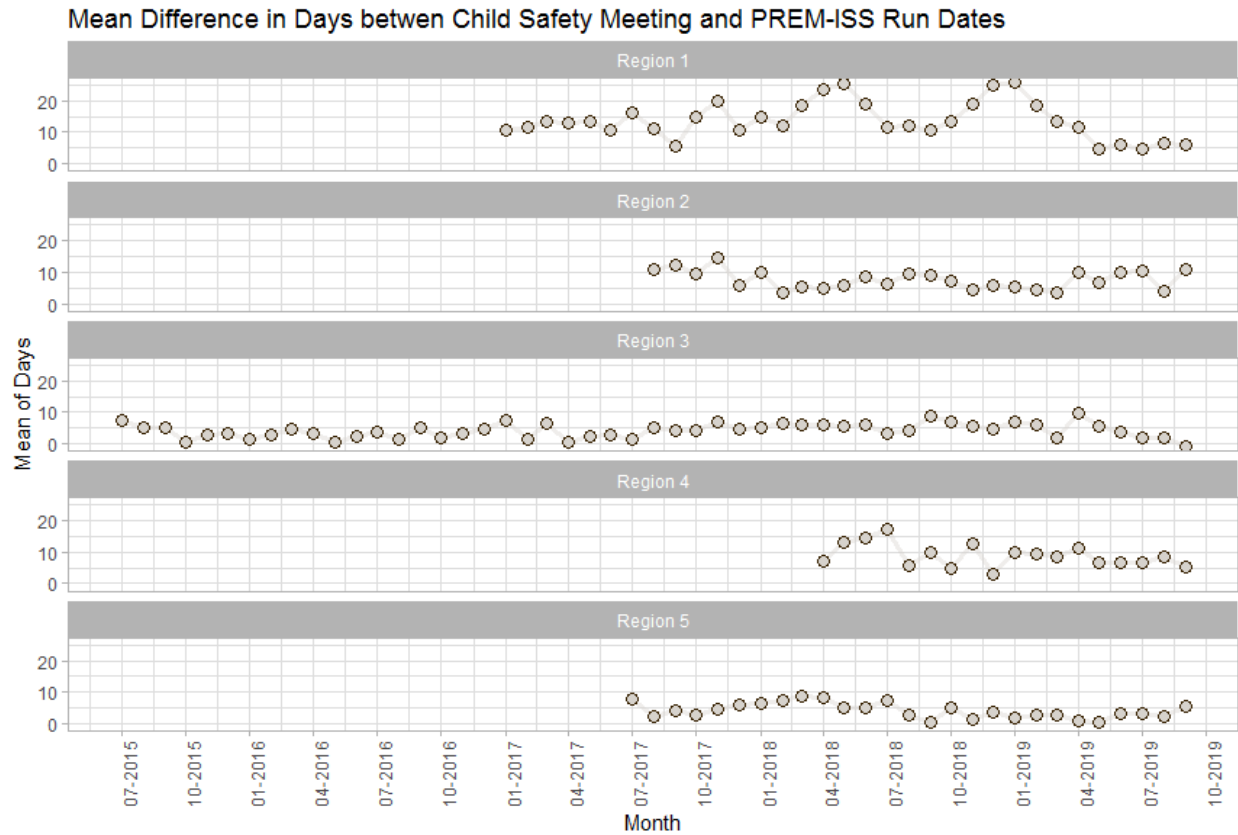
Figure 4 presents the percent of PREMISS eligible referrals that violated protocol – where the CSM preceded the PREMISS model. Evaluators worked directly with the regional leadership to improve these rates. For example, in Region 3, performance was best the month following the issuance of a directive by the Region 3 District Director to populate PREMISS fields more quickly (only 22% of referrals were in the "wrong" direction in May, 2016).

Percent of Eligible Referrals Where CSM Preceeded the PREM-ISS Run
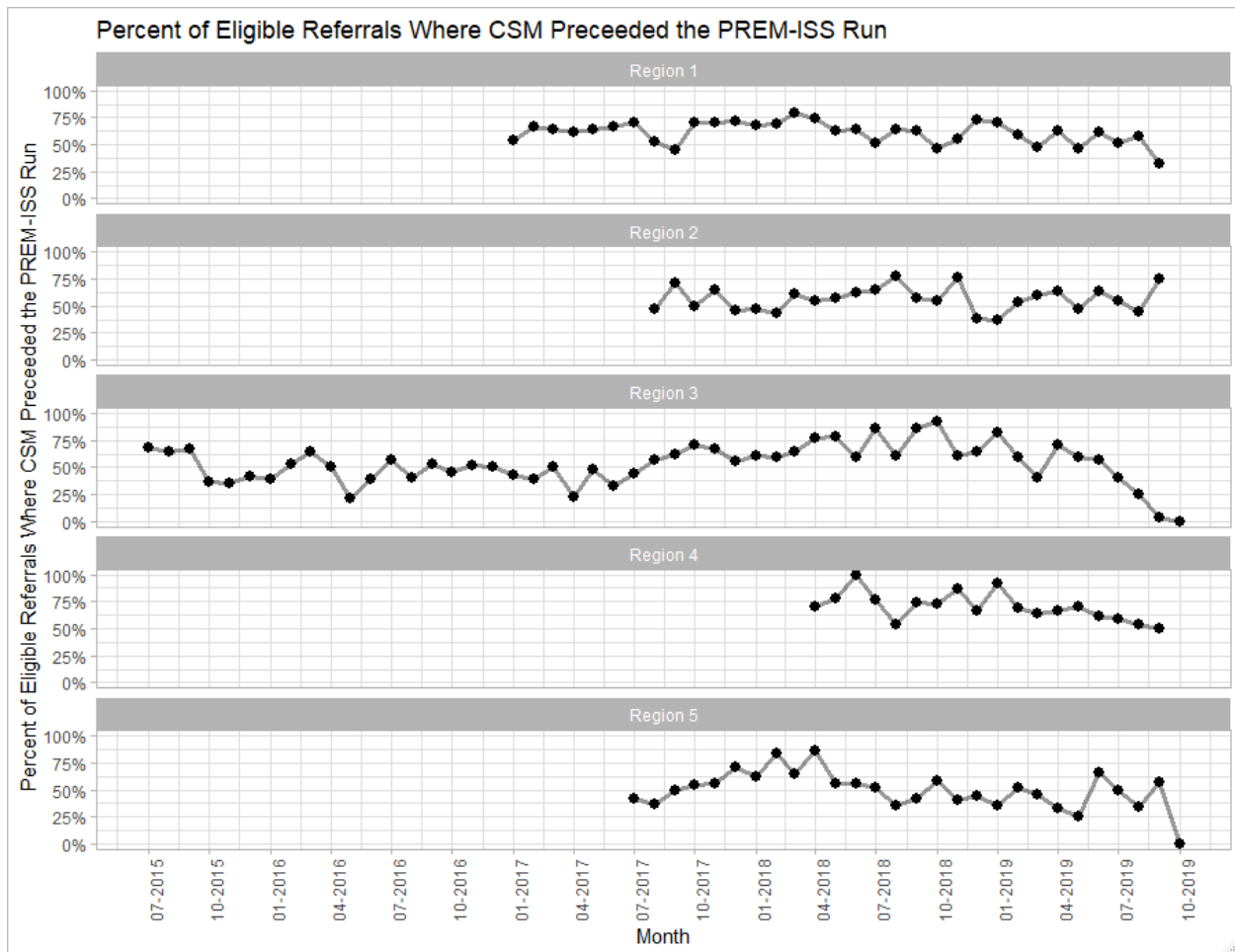
We also examine the time to service for initiation of ISS. The protocol is that ISS make face contact within 24 hours of referral. ISS referral should occur within 24 hours of PREMISS eligibility establishment. Thus, from PREMISS to contact should be about 48 hours. For cases accepting ISS, the average length of time between the PREMISS notice and the ISS service referral was 2.8 (range = 0 - 13), 4 (range = 0 - 60), 3.1 (range = 0 - 51), 3.9 (range = 0 - 10), and 3.5 (range = 0 - 18) days for Regions 1, 2, 3, 4, and 5, respectively. The average length of time between ISS referral and intake was 1.2 (range = 0 - 7), 1.9 (range = 0 - 66), 1.1 (range = 0 - 7), 3.6 (range = 0 - 15), 7 (range = 0 - 364) days for these Regions, respectively. These results show that, on average, there is some room for improvement, but given the short timelines, it seems good faith efforts are being made to establish contact with the family quickly.

ISS was designed to be up to six weeks of intervention. The average length of cases reaching closure was 37.9 (range = 4 - 70), 36.7 (range = 0 - 63), 37 (range = 1 - 65), 38.7 (range = 4 - 56), and 39.6 (range = 7 - 63), respectively. Across all regions, of the 507 cases closed within the reporting period, the reasons for closure are presented in Figure 5. Reasons for closure during the reporting period included requirements completed (n = 314), goals partially completed (n = 41), change in custody (adoption/guardianship; n = 18), closed at the request of child welfare (usually due to incarceration or inpatient treatment of PRFC; n = 14), uncooperative/refused services (n = 46), child(ren) removed (n = 78), family relocated (n = 10). In some instances, multiple reasons may be selected for closure, so the

40

count is greater than the number of referrals included. Of the 314 closures meeting ISS requirements, 303 successfully stepped down to the Comprehensive Home-Based Services (CHBS) program. The average length of time between the PREMISS notice and the start of CHBS for these cases was 34.4 days (range = 0 - 63). [14]

FIGURE 5: REASONS FOR ISS CLOSURE
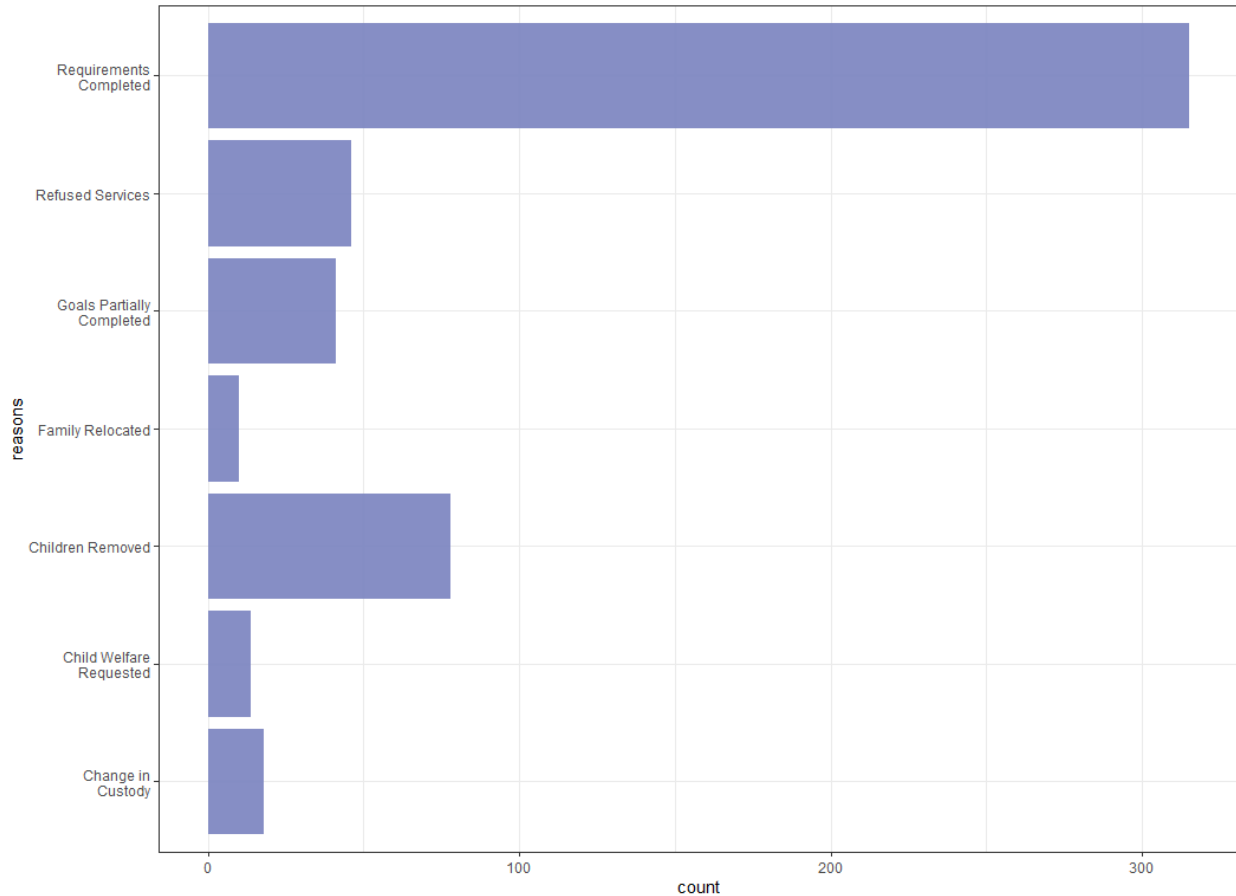
From the beginning of the study to October 2nd, 2019, there have been 996 cases randomized to ISS that were not served. We refer to these cases as the "ISS Not Workable" group throughout the report. Reasons for not receiving ISS are listed below:

- Court Intervention – 613
- Client Refused ISS – 72
- No Safety Monitors – 80
- Services Not Needed (safety threats do not necessitate ISS) – 105
- Services Not Needed (child to adoption/guardianship) – 16
- Client Withdrawn Due to Safety Concerns – 106
- Client Not Available for ISS – 32

---

[14] CHBS typically starts before the ISS case closes.

- Other – 35 (e.g., involved slow referral actions by a child welfare worker which hampered the start of services)
- 2715 referrals were assigned to services as usual (SAU)
- 614 cases met exclusionary criteria and were effectively NOT randomized to either condition.

Model Fidelity

The ISS therapist and FCS specialist were asked to complete a staffing form for each week of the ISS portion of the service. The intent was to help facilitate productive staffing discussions, and also to provide up-to-date information on the progress of the case and utilization of resources and referrals. The ISS version of the form provides information about how many visits and/or calls there were with the parents that week, if there were any new safety threats that were observed, and if any diminished or enhanced protective capacities were noted. They reported on utility of the current safety plan, cooperation with services, and on the status of any referrals made for adjunct treatment services for the children or parents. The FCS version of the form also reported on the number of visits with the family that week, any observed changes to safety threats and protective capacities, updates to the safety plan, cooperation with services, and any updates they may have received from collateral service providers for the family. The ISS therapist provided the FCS specialist with the staffing form weekly, and the FCS specialist was instructed to place the forms in the KIDS record and send a copy to OUHSC. In some cases, the ISS therapist also sent a copy of their form to OUHSC separately.

Table 6 shows a total count of clients receiving ISS at each weekly measurement point, as well as the counts of FCS forms and ISS forms present and missing for each week. Table 7 reports this information by region. These tables also includes the amount of missing data not received by the evaluation team or not completed by the providers because, not unlike other field trials, it is significant. In some cases, missing forms are due to ISS services that ended early (e.g., in the 4th and 5th week) negating the need to complete staffing forms for the final weeks.

TABLE 6- FORMS FOR EACH WEEK OF ISS INVOLVEMENT – ALL REGIONS

Weeks	Total Families in ISS	FCS Forms Available	FCS Forms Missing	ISS Forms Available	ISS Forms Missing	Visits Items Missing on FCS Staffing	Visits to Children Items Missing on ISS Staffing	Visits to PRFCs Items Missing on ISS Staffing
Week-1	551	186	365	383	168	65	46	39
Week-2	532	213	319	396	136	64	50	44
Week-3	509	213	296	380	129	65	47	44
Week-4	480	185	295	347	133	54	40	32
Week-5	442	157	285	255	187	48	21	19
Week-6	363	71	292	89	274	10	11	6

TABLE 7- FORMS FOR EACH WEEK OF ISS INVOLVEMENT BY REGION

Regions N=Total # of Clients	Measurement "Weeks"	Families in ISS in Measurement "Weeks"	FCS Forms Available	FCS Forms Missing	ISS Forms Available	ISS Forms Missing	Visits Items Missing on FCS Staffing	Visits to Children Items Missing on ISS Staffing	Visits to PRFCs Items Missing on ISS Staffing
Region 1 (n = 88)	week_1	89	28	61	15	83	6	6	2
	week_2	85	30	55	13	79	6	10	6
	week_3	83	25	58	12	72	11	7	2
	week_4	76	21	55	12	63	13	6	2
	week_5	71	17	54	8	50	21	3	2
	week_6	59	8	51	2	16	43	1	2
Region 2 (n = 90)	week_1	88	18	70	6	77	11	12	4
	week_2	84	18	66	8	75	9	10	5
	week_3	79	22	57	7	68	11	10	4
	week_4	77	17	60	7	68	9	6	2
	week_5	72	14	58	7	50	22	5	4
	week_6	56	4	52	0	19	37	3	0
Region 3 (n = 234)	week_1	248	100	148	24	187	61	23	21
	week_2	239	123	116	28	202	37	26	19
	week_3	231	126	105	32	197	34	24	23
	week_4	216	109	107	22	178	38	21	15
	week_5	195	93	102	21	132	63	12	8
	week_6	160	38	122	4	46	114	7	1
Region 4 (n = 31)	week_1	31	5	26	1	4	27	0	2
	week_2	30	4	26	1	4	26	0	0
	week_3	29	4	25	1	5	24	0	1
	week_4	29	4	25	0	4	25	1	2
	week_5	28	3	25	1	2	26	0	1
	week_6	23	2	21	1	0	23	0	0
Region 5 (n = 82)	week_1	95	35	60	19	32	63	5	10
	week_2	94	38	56	14	36	58	4	14
	week_3	87	36	51	13	38	49	6	14
	week_4	82	34	48	13	34	48	6	11
	week_5	76	30	46	11	21	55	1	4
	week_6	65	19	46	3	8	57	0	3

Table 8 reports the referrals and acceptance rates by regions. A total of N = 1056 referrals (some multiple referrals per case) have been made to date, of which N = 782 (74.1%) have been accepted by the Person Responsible for Child (PRFC). Among these referrals, 85 clients in region 1, 79 clients in

region 2, 221 clients in region 3, 7 clients in region 4 and 62 clients in region 5 contributed 77, 77, 209, 5 and 46 unique referral cases respectively.

TABLE 8- SERVICES REFERRALS AND ACCEPTANCES BY REGIONS

Total 1056 referred, 782 accepted	Region 1 (total clients=85) 77 referred		Region 2 (total clients=79) 77 referred		Region 3 (total clients=221) 209 referred		Region 4 (total clients=7) 5 referred		Region 5 (total clients=62) 46 referred		Total	
	R	A	R	A	R	A	R	A	R	A	R	A
Any services for children	12% (41)	63% (26)	12% (42)	79% (33)	13% (117)	70% (82)	0% (0)	0% (0)	4% (9)	33% (3)	20% (209)	18% (144)
Mental Health	15% (55)	65% (36)	19% (63)	86% (54)	16% (151)	79% (120)	8% (2)	50% (1)	8% (18)	39% (7)	15% (163)	16% (125)
Substance Abuse	13% (47)	66% (31)	13% (43)	88% (38)	15% (136)	83% (113)	4% (1)	0% (0)	7% (16)	38% (6)	14% (152)	14% (107)
Domestic violence	5% (16)	50% (8)	9% (31)	74% (23)	11% (107)	84% (90)	8% (2)	50% (1)	3% (7)	43% (3)	27% (289)	28% (218)
FRS referral	10% (36)	56% (20)	10% (35)	80% (28)	8% (74)	76% (56)	0% (0)	0% (0)	3% (7)	43% (3)	23% (243)	24% (188)
Total	55% (195)	62% (121)	63% (214)	82% (176)	63% (585)	79% (461)	20% (5)	40% (2)	25% (57)	39% (22)		

Note: R is referral; A is acceptance.

Referral Follow-up

Beginning January 19, 2018, DHS workers started recording follow-up status on referrals made for different services in all five regions. This information is presented in Table 9 and Table 10. Since then, around 228 clients have this follow-up status. A total of 410 referrals were made and 141 (34%) of these referred services were started.

TABLE 9-CROSS -SECTIONAL VIEW OF REFERRAL FOLLOW-UP STATUS FOR NEW CLIENTS ALL REGIONS

Referrals	Referral Count	Intake completed	Intake scheduled	Missed appointment	Services ended	Services started	waiting for response from referral agency	waitlist
Any referrals for Children	78	12	29	1	10	14	57	1
Domestic violence referrals	68	30	35	7	2	31	28	6
Mental health Referrals	125	33	75	6	1	50	42	4
Substance abuse referrals	103	32	67	7	3	37	36	6

Region	Referrals	Referral Count	Intake completed	Intake scheduled	Missed appointment	Services ended	Services started	waiting for response from referral agency	waitlist
Region 1	Any referrals for Children	17	3	5	-	-	4	14	-
(N=47)	Domestic violence referrals	6	2	4	-	-	2	3	-
	Mental health Referrals	22	9	13	3	1	8	5	-
	Substance abuse referrals	21	8	15	4	-	6	8	-
	Total	66	22	37	7	1	20	30	0
Region 2	Any referrals for Children	27	3	11	1	5	6	18	-
(N=64)	Domestic violence referrals	22	4	10	1	1	11	10	-
	Mental health Referrals	45	7	23	1	-	25	20	-
	Substance abuse referrals	32	7	18	1	2	17	11	1
	Total	126	21	62	4	8	59	59	1
Region 3	Any referrals for Children	29	4	9	-	5	4	24	1
(N=69)	Domestic violence referrals	36	21	20	6	1	18	14	6
	Mental health Referrals	48	13	33	2	-	16	17	4
	Substance abuse referrals	40	13	30	2	1	12	16	5
	Total	153	51	92	10	7	50	71	16
Region 4	Any referrals for Children	-	-	-	-	-	-	-	-
(N=5)	Domestic violence referrals	1	1	-	-	-	-	-	-
	Mental health Referrals	1	-	-	-	-	1	-	-
	Substance abuse referrals	1	-	1	-	-	-	-	-
	Total	3	1	1	0	0	1	0	0
	Any referrals for Children	5	2	4	-	-	-	1	-
Region 5	Domestic violence referrals	3	2	1	-	-	-	1	-
(N=41)	Mental health Referrals	9	4	6	-	-	-	-	-
	Substance abuse referrals	9	4	3	-	-	2	1	-
	Total	26	12	14	0	0	2	3	0

Referrals based on the Child Well-Being Scales (CWBS)

In addition to the ISS staffing forms, the referrals to "drug and alcohol treatment", "mental health treatment" and "domestic violence treatment" are also recorded by DHS workers on the Child Well-Being Scales (CWBS), collected at Baseline, Stepdown to CHBS (~6 weeks after group assignment) and at 6 months. Table 11 shows the counts and percentages for the clients who are in need of, and also those receiving services at each assessment point. A total 492 unique clients have measurements across all five regions. Due to attrition and censoring, not all clients have all three assessment points. For example, 477 clients completed the baseline measurement. Among that group, 265 clients reached the CHBS Stepdown, and only 248 reached the six-month measurement point.

TABLE 11- SERVICES REFERRALS AND CURRENT TREATMENT STATUS FROM THE "CHILD WELL-BEING SCALE - CWBS" BY REGIONS

Measurements	Services	Region 1 Total clients = 87		Region 2 Total clients = 73		Region 3 Total clients = 232		Region 4 Total clients = 26		Region 5 Total clients = 74	
		referral	Currently in treatment	referral	Currently in treatment	referral	Currently in treatment	referral	Currently in treatment	referral	Currently in treatment
Baseline		CWBS completed = 85		CWBS completed = 63		CWBS completed= 232		CWBS completed = 24		CWBS completed = 73	
	Domestic Violence	14% (12)	8% (7)	25% (16)	14% (9)	31% (71)	22% (52)	29% (7)	12% (3)	25% (18)	11% (8)
	Mental Health	48% (41)	12% (10)	59% (37)	13% (8)	32% (75)	13% (31)	38% (9)	29% (7)	55% (40)	27% (20)
	Alcohol Abuse	11% (9)	2% (2)	10% (6)	3% (2)	13% (30)	3% (8)	4% (1)	-	7% (5)	1% (1)
	Drug-Use	39% (33)	6% (5)	29% (18)	13% (8)	31% (71)	9% (22)	42% (10)	21% (5)	34% (25)	18% (13)
CHBS Stepdown		CWBS completed = 46		CWBS completed = 49		CWBS completed= 139		CWBS completed = 8		CWBS completed = 23	
	Domestic Violence	26% (12)	26% (12)	33% (16)	31% (15)	45% (63)	41% (57)	38% (3)	38% (3)	22% (5)	4% (1)
	Mental Health	76% (35)	57% (26)	65% (32)	41% (20)	72% (100)	53% (73)	62% (5)	38% (3)	78% (18)	65% (15)
	Alcohol Abuse	11% (5)	11% (5)	6% (3)	4% (2)	15% (21)	14% (20)	-	-	-	-
	Drug-Use	43% (20)	37% (17)	29% (14)	24% (12)	42% (59)	38% (53)	50% (4)	38% (3)	52% (12)	39% (9)
Six Months		CWBS completed = 41		CWBS completed = 35		CWBS completed= 133		CWBS completed = 6		CWBS completed = 33	
	Domestic Violence	32% (13)	22% (9)	29% (10)	20% (7)	36% (48)	32% (43)	17% (1)	17% (1)	18% (6)	6% (2)
	Mental Health	66% (27)	49% (20)	31% (11)	17% (6)	56% (74)	34% (45)	17% (1)	-	42% (14)	30% (10)
	Alcohol Abuse	17% (7)	17% (7)	6% (2)	3% (1)	12% (16)	9% (12)	-	-	3% (1)	-
	Drug-Use	39% (16)	32% (13)	23% (8)	11% (4)	35% (47)	22% (29)	33% (2)	-	33% (11)	18% (6)

Note: Percentage in each cell were calculated from the total clients at each given measurement point, with the numerator as each cell count and denominator was corresponding to "CWBS completed".

Model Fidelity: Visits

As noted previously, ISS has a visit protocol for frequency and type of visits. These visits are documented on FCS and ISS Staffing Checklists. The current protocol for FCS and ISS providers is:

1) FCS workers visit families and children at least once every 7 calendar days. (This policy is the case regardless of whether the family is in ISS or not.)
2) ISS workers visit PRFCs 3 to 5 times every 7 calendar days.
3) ISS workers visit children at least once every 30 calendar days within the ISS intervention period. These workers visit families to see both children and their parents together. There were no expectations to see children alone.

The counts of complete and missing entries are recorded for FCS visits. Table 12 presents the number of families served, the number of completed forms, and the number of missing information on visits for providers who had completed forms. The missing visits ranges from 14% to 35% for individual weeks and 30% across all weeks of service. Table 13 presents this information by region.

TABLE 12- COUNT OF FAMILIES, FORM COMPLETION, AND MISSING VISIT INFORMATION PER PROTOCOL FOR FCS VISITS TO FAMILIES AND CHILDREN-ALL REGIONS.

Measurements	Families in ISS	Therapists who completed forms	Visit Missing count	percentage missing
Week-1	551	186	65	35%
Week-2	532	213	65	31%
Week-3	509	213	65	31%
Week-4	480	185	54	29%
Week-5	442	157	48	31%
Week-6	363	71	10	14%
Total	2877	1025	307	30%

TABLE 13 - COUNT OF FAMILIES, FORM COMPLETION, AND MISSING VISIT INFORMATION PER PROTOCOL FOR FCS VISITS TO FAMILIES BY WEEKS AND REGIONS

Regions	Measurements	Families in ISS	Therapists who completed forms	Visit Missing Count	Percentage missing
Region 1	Week-1	89	28	15	54%
	Week-2	85	30	13	43%
	Week-3	83	25	12	48%
	Week-4	76	21	12	57%
	Week-5	71	17	8	47%
	Week-6	59	8	2	25%
	Total	463	129	62	48%
Region 2	Week-1	88	18	7	39%
	Week-2	84	18	8	44%
	Week-3	79	22	7	32%
	Week-4	77	17	7	41%
	Week-5	72	14	7	50%
	Week-6	-	-	-	-
	Total	400	89	36	40%
Region 3	Week-1	248	100	24	24%
	Week-2	239	123	28	23%
	Week-3	231	126	32	25%
	Week-4	216	109	22	20%
	Week-5	195	93	21	23%
	Week-6	160	38	4	11%
	Total	1289	589	131	22%
Region 4	Week-1	31	5	1	20%
	Week-2	30	4	1	25%
	Week-3	29	4	1	25%
	Week-4	-	-	-	-
	Week-5	28	3	1	33%
	Week-6	23	2	1	50%
	Total	141	18	5	28%
Region 5	Week-1	95	35	18	51%
	Week-2	94	38	15	39%
	Week-3	87	36	13	36%
	Week-4	82	34	13	38%
	Week-5	76	30	11	37%
	Week-6	65	19	3	16%
	Total	499	192	73	38%

Table 14 and Table 15 report on visits according to protocol for ISS therapist visits to PRFCs overall and by region, respectively. We report both visit information that is missing (i.e., the therapist did not complete the visit item) and visit violations (i.e., the therapists recorded fewer than 3 visits with a family). There were 190 visit violations. More specifically, 49 violations were reported in Region 1, 31 violations were recorded in Region 2, 76 violations in Region 3, 8 violations in Region 4 and 26 violations for visits in Region 5. It should be noted that although having less than 3 visits in a week technically is a violation of the original protocol, there are practice variations that allow for these deviations. For example, when there was a holiday during a given week of treatment, or a provider was on vacation, supervisors may have approved less visits. In other instances, the case goals may have been met and the case was getting ready to step down to CHBS, and the supervisor and ISS therapist may have agreed to

have less visits. Therefore, although there may be some true violations of not having at least three visits a week, some of the others may have been deliberate and occurred with supervisory approval.

TABLE 14: COUNT OF FAMILIES, FORM COMPLETION, AND MISSING VISIT INFORMATION PER PROTOCOL FOR ISS VISITS TO PRFCS-ALL REGIONS.

Measurements	Families in ISS	Therapists who completed forms	Missing		Violation	
			count	*Missing percentage*	*count*	*Violation percentage*
Week-1	551	383	41	11%	27	7%
Week-2	532	396	45	11%	43	11%
Week-3	509	380	44	12%	40	11%
Week-4	480	347	31	9%	41	12%
Week-5	442	255	20	8%	25	10%
Week-6	363	89	6	7%	14	16%
Total	2877	1850	187	10%	190	10%

TABLE 15: COUNT OF FAMILIES, FORM COMPLETION, AND MISSING VISIT INFORMATION PER PROTOCOL FOR ISS VISITS TO PRFCs BY WEEKS AND REGIONS

Regions	Measurements	Families in ISS	Therapists who completed forms	Missing count	Missing percentage	Violations count	Violations percentage
Region 1	Week-1	89	83	2	2%	10	12%
	Week-2	85	79	6	8%	10	13%
	Week-3	83	72	2	3%	8	11%
	Week-4	76	63	2	3%	11	17%
	Week-5	71	50	2	4%	9	18%
	Week-6	59	16	2	12%	1	6%
	Total	463	363	16	4%	49	13%
Region 2	Week-1	88	77	4	5%	4	5%
	Week-2	84	75	5	7%	9	12%
	Week-3	79	68	4	6%	6	9%
	Week-4	77	68	2	3%	7	10%
	Week-5	72	50	4	8%	2	4%
	Week-6	56	19	-		3	16%
	Total	456	357	19	5%	31	9%
Region 3	Week-1	248	187	22	12%	10	5%
	Week-2	239	202	20	10%	13	6%
	Week-3	231	197	23	12%	16	8%
	Week-4	216	178	14	8%	20	11%
	Week-5	195	132	9	7%	11	8%
	Week-6	160	46	1	2%	6	13%
	Total	1289	942	89	9%	76	8%
Region 4	Week-1	31	4	2	50%	1	25%
	Week-2	30	4	NA		2	50%
	Week-3	29	5	1	20%	3	60%
	Week-4	29	4	2	50%	1	25%
	Week-5	28	2	1	50%	1	50%
	Week-6	-	-	-	-	-	-
	Total	147	19	6	32%	8	42%
Region 5	Week-1	95	32	11	34%	2	6%
	Week-2	94	36	14	39%	9	25%
	Week-3	87	38	14	37%	7	18%
	Week-4	82	34	11	32%	2	6%
	Week-5	76	21	4	19%	2	10%
	Week-6	65	8	3	38%	4	50%
	Total	499	169	57	34%	26	15%

Table 16 reports on both visit-item missingness and approximate violations for ISS workers visiting children. The table reports the number of times a provider did not visit a child before closing an ISS case that lasted 4 weeks or more. Similar to the situation of PRFCs visits, the violations happened when the ISS workers did not visit children at least once during the period where documentation was available. ISS workers did not complete the child visits on schedule 37 times across all regions, reflecting an 8% visit to children violation. Regional variation in visit violation to children ranged 4% to 15% within regions. Visit violation rates to children were generally low, but could still be improved.

| Measurements | Families in ISS | Missing | | Violation | |
		count	percentage	count	Violation percentage
Region 1	86	5	6%	4	5%
Region 2	84	7	8%	3	4%
Region 3	223	13	6%	19	9%
Region 4	21	16	76%	1	5%
Region 5	68	25	37%	10	15%
Total	482	66	14%	37	8%

Table 17 presents the mean and median number of visits by FCS and ISS workers by region per week.

TABLE 17- DESCRIPTIVE INFORMATION OF THE NUMBER OF TIMES DHS WORKERS VISITED OR CALLED CLIENTS

| | FCS workers visited families and children (times a week) | | | ISS specialists visited PRFCS (times a week) | | | ISS specialist called families (times a week) | | |
Measurement	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD
Region1	1.28	1.03	0.51	3.23	3.25	0.38	1.53	1.17	0.65
Region2	1.48	1.50	0.44	3.36	3.25	0.45	1.33	1.00	0.47
Region3	1.41	1.25	0.54	3.68	3.75	0.58	1.74	1.00	1.15
Region4	1.18	1.04	0.32	2.46	2.25	0.52	1.50	1.50	0.00
Region5	1.24	1.00	0.42	3.07	3.00	0.73	1.65	1.50	0.63
Total	1.37	1.20	0.51	3.47	3.40	0.58	1.63	1.00	0.91

Note: SD = Standard Deviation.

In general, these summary statistics reflect good, but not perfect adherence to the protocols.

Healthy Relationships

To address issues of interpersonal conflict in the home, the *Healthy Relationships* curriculum can be incorporated into the ISS or the CHBS interventions at the discretion of the service provider and the provider's supervision team. Data from all previous fiscal years are available to the evaluators for the purpose of current report. Clients and families from data sources were matched against counts of clients determined to be in ISS. Of the 558 ISS cases, a total of 72 families received either partial or the complete Healthy Relationships training module (Table 18). (Note: Healthy Relationships is administered at some point during CHBS delivery).

TABLE 18- HEALTHY RELATIONSHIPS FOR ISS FAMILIES

Regions	Families in ISS	Healthy Relationship Module	Count	Percentage[15]
Region 1 ISS	89	Complete Training	1	0.28%
		Partial Training	6	1.69%
Region 2 ISS	90	Complete Training	1	0.28%
		Partial Training	1	0.28%
Region 3 ISS	251	Complete Training	27	7.61%
		Partial Training	16	4.51%
Region 4 ISS	31	Complete Training	6	1.69%
		Partial Training	1	0.28%
Region 5 ISS	97	Complete Training	9	2.54%
		Partial Training	4	1.13%
Total	558		72	20.28%

A total of 3,828 clients were referred for CHBS after the referral that triggered their study inclusion. Among them, 2,801 were from the "SAU" group and eight of them received "complete or partial training" from the Healthy Relationship module.

Qualitative Interviews

One-on-one interviews with child protection services (CPS) investigation and assessment teams, ISS workers, and FCS workers took place in all regional districts after ISS was initiated. Interviews for the qualitative assessment targeted key process and implementation outcomes at two contextual levels. The conceptual model of implementation that these levels are based on, suggest there are inner and outer contexts to sustainability, funding, and policy regarding factors to implementation. [16] Specifically, these levels include: 1) Outer context - funding issues (sources, perceptions of funding security, plans for sustained funding), sociopolitical issues (impact of other federal initiatives, Pinnacle Plan, pending and enacted legislation), client demands and satisfaction (stigmatization, cultural relevance and sensitivity), inter-organizational collaborations and networking (number, type, and strength of partnerships, service and judicial system collaborations, availability and use of technical assistance), and high-level leadership; and 2) Inner context – absorptive capacity (workforce education, workloads, effective supervision), staff and administrative structuring, culture (support for identifying and implementing best practices, QA/QI support and expectations, readiness to change, policies, procedures, goals), climate (support for creativity, innovation-value fit, learning orientation, teamwork, speed of action, tolerance for mistakes, openness), local leadership (internal program champions, open communication), adopter characteristics (values and goals, social and professional networks, perceived need for and risk of change). Interviewees were asked to describe the perceived impact each of these contextual factors have had on ISS and SAU implementation.

---

[15] Percentages are calculated from the number of families in ISS within each region.

[16] Aarons, A., Hurlburt, M., & Horwitz, S. (2011). Advancing a Conceptual Model of Evidence-Based Practice Implementation in Public Service Sectors. Administration and Policy in Mental Health. 38, 4-23. doi: 10.1007/s10488-010-0327-7.

Methods

Procedures

A purposive sampling method was used with CPS workers (N=38), FCS workers (N=41), ISS workers (N=16), ISS supervisors (N=8), FCS supervisors (N=29) and administrators (N=6). CPS and FCS workers and FCS supervisors were those who had direct contact with the waiver program and ISS workers were providing direct intensive safety services.

CPS, FCS, and ISS workers, FCS and ISS supervisors, and administrators were recruited for semi-structured interviews with a total of 138 interviews conducted. The research team contacted program supervisors, explained the purpose of the study, and asked supervisors to advertise the current study. Next, a qualitative team member recruited workers, supervisors, and administrators for interviews. CPS, FCS, ISS, and administrative interviews were conducted during normal workday hours and at times convenient for participants. The qualitative study evaluation methods were approved by OUHSC and OSU University Institutional Review Boards (IRBs).

Individual interviews lasted from approximately 1 to 1.5 hours each. Individual interviews were conducted by trained qualitative interviewers. All interviews were transcribed and cross-checked by qualitative team members to ensure accuracy. During transcription, all identifying information that could be linked to informants was removed.

Individual Interview Guide Development

Semi-structured individual interview guides were developed by the lead qualitative researcher in conjunction with all key study personnel and reviewed by the qualitative research team. The purpose of the semi-structured guides was to assess implementation of the waiver program from the perspective of CPS, FCS, and ISS workers, FCS and ISS supervisors, and administrators.

Data Analysis

Qualitative data analysis of the transcriptions was conducted using NVivo 10 and 11 software. A template approach was used to identify broad themes influenced by the semi-structured interview guide. [17] Specifically, a codebook was developed collaboratively by at least three trained qualitative researchers. Thematic analysis was then used to define specific themes and were coded as sub-codes within the broader categories. Training in developing the coding scheme and codebook took approximately two days for each interview guide. Upon completion of the codebook two qualitative researchers coded transcriptions using the created templates. Once coded, transcriptions were compared for inter-coder reliability. To assess for reliability, simple percent agreement was used between the two coders. Simple agreement is a commonly used method for assessing reliability in qualitative studies[18][19]. Average inter-coder reliability was over 85%, which is considered a good standard to meet.

---

17 Patton, M. Q. (2002). Qualitative analysis and interpretation. In Qualitative Research and Evaluation Methods (3rd ed., pp. 431-540). Thousand Oaks, CA: Sage Publications.

[18] Lombard, M., Snyder-Duch, J., & Bracken, C. C. (2002). Content analysis in mass communication: Assessment and reporting intercoder reliability. Human Communication Research, 8, 587-604. doi:10.1111/j.1468-2958.2002.tb00826.x

[19] Stemler, S. (2004). A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. Practical Assessment, Research, & Evaluation, 9(4)

Percentages were calculated for every specific and broad theme within the transcription. First, the percentages were calculated for each participant who made at least one comment about each particular theme. This allowed researchers to identify what proportion of participants spoke about a particular subject. Second, the percentage of total comments for each theme was calculated. This gave an indication of the emphasis that each theme received during the interviews. Themes are described in more detail below.

Results

Number and type of staff involved in implementation, including training received, level of experience, education and characteristics

Training Received During Implementation of ISS Services

FCS workers described the training received to implement ISS services with most FCS workers reporting that they received an explanation of ISS services (86% of FCS workers). It was reported that FCS workers receive trainings specifically related to child welfare, child protective services, family centered services, permanency planning, and trauma. Most FCS workers reported that training is accessible (86% of FCS workers). Some FCS workers reported that training prepared them well to work with families enrolled in ISS (38% of FCS workers) while others reported insufficient training (38% of FCS works). Specific areas workers felt prepared in were explaining ISS to families, understanding interpersonal violence, and clarification of the ISS protocol. Some FCS workers also shared that on the job training was beneficial (33% of FCS workers) and that training had improved over time. FCS workers also shared that to improve training there needs to be ongoing refresher trainings focused on protocol changes and training of both FCS and ISS workers. More training on the protocol and paperwork, more hands on training, and more training focused on interpersonal violence was suggested by some FCS workers.

Contextual factors, such as the socio-political, economic and political forces that may have a bearing on the replicability of the intervention or influence the implementation or effectiveness of the demonstration. This discussion will note any possible confounding effects of changes in these systems, or changes resulting from other demonstrations or reforms that were implemented during the title IV-E demonstration

Administrative or Political Issues that Impact the Program

Administrative and political issues that impact the waiver program were discussed among some CPS and FCS workers. Some CPS and FCS workers did not comment on administrative issues or said they were not aware of any issues (81% of CPS workers; 15% of FCS workers). Additionally, some FCS workers reported there were no administrative issues (52% of FCS workers). Among those who FCS workers who said they were, issues included administrative decisions being made without feedback from frontline workers (13% of FCS workers), issues with protocol timelines (9% of FCS workers), and worker turnover (9% of FCS workers).

In terms of political issues impacting the program, CPS and FCS workers mentioned issues with funding (9% of CPS workers; 39% of FCS workers) and the Pinnacle Plan (10% of FCS workers). Specifically, FCS workers reported that the pinnacle plan was a political issue that created motivation of prevention efforts and limited caseload size (17% of FCS workers). Many CPS and FCS workers were unsure of political issues that impact the waiver program (91% of CPS workers; 7% of FCS workers).

Many supervisors reported no current administrative or political issues impacting the program (33% of FCS supervisors) or were unsure (11% of FCS supervisors). Those that did report issues explained concerns about the state budget deficit (26% of FCS supervisors) that has led to decreased DHS staff and reduced service funding. FCS supervisors also explained that there are issues with vendors bidding on state contracts which results in delays in getting ISS services in place and restricts services to one agency. It was also reported that administrative turnover has become an issue.

ISS workers reported that political issues include funding for both DHS and ISS services. It is important to note that some ISS workers had no administrative or political issues to report (14% of ISS workers) or were unsure of issues (21% of ISS workers).

ISS supervisors also reported on several administrative issues. These include deadlines, a hierarchy within DHS that makes it difficult for ISS workers to know who to contact with issues, DHS limiting the agencies that families can work with, some agencies not being able to text clients, and DHS prematurely closing cases without ISS.

In terms of political issues, funding was reported as being a concern. Other issues reported include political issues between CPS and FCS with ISS personnel remaining neutral and policies that hinder families from getting services such as perpetrators not qualifying for service funding.

One of the most salient administrative and political issues discussed from the perspective of the administrators was the potential for children to be harmed while enrolled in ISS. Administrators also reported that funding is a concern due to ongoing difficulties in securing prevention services and issues with individuals understanding federal dollars and what they cover. Funding issues were also explained to be related to illegitimate requests to use Title IV-E funds.

Organizational aspects of the demonstration, such as staff structure, funding committed, administrative structures, and project implementation, including ongoing monitoring, oversight, and problem resolution at various organization levels

CPS Worker Intra-Agency Collaboration Perspective

Some CPS workers reported minimal involvement with FCS workers (26% of CPS workers). Other CPS workers described communication with FCS (55% of CPS workers) that included emails, phone calls, and sharing information. CPS workers also discussed overall collaboration with FCS as successful with all CPS workers discussing positive aspects of collaboration. Specific positives include sharing information, having good communication, having shared goals, being able to collaborate within child welfare, workers supporting each other, and working well with FCS. Negative aspects of collaboration with FCS were shared by 21% of CPS workers and included disagreements about the best option for child safety.

CPS Worker Inter-Agency Collaboration Perspective

Themes regarding between agency interactions with ISS workers was varied with primarily positive responses. CPS workers explained that there is good collaboration with ISS workers and good communication. CPS workers were mixed regarding level of ISS involvement with some indicating high involvement (22% of CPS workers) and some indicating minimal involvement (44% of CPS workers). CPS also indicated the need for more general communication between DHS and ISS and the need for ISS to follow up with CPS after ISS services are complete so they know how the family did in services.

FCS Worker Intra-Agency Collaboration Perspective

Themes emerged regarding within agency collaboration from the perspective of FCS workers. FCS workers reported good overall interagency communication (24% of FCS workers) and good supervisors (16% of FCS workers). It was further reported that there is good communication with the child safety meeting facilitator and among FCS workers. FCS workers also explained there is good communication with the OCS liaison with them connecting families to CHBS. Other positive collaborations include permanency planning consultations and case staffing.

FCS workers were also asked about barriers to within agency collaboration with some workers indicating that there are no barriers to report (50% of FCS workers). Barriers that were reported include issues with delayed involvement of FCS in ISS cases (6% of FCS workers), worker overload (35% of FCS workers), AOCS being incomplete or having errors (6% of FCS workers), and scheduling difficulties (12% of FCS workers).

FCS Worker Inter-Agency Collaboration Perspective

FCS workers reported that overall there is good collaboration with ISS workers (43% of FCS workers). This includes good communication (85% of FCS workers) through email, phone, text, and in person contact. It was also reported that good communication includes relaying critical incidents. Other positive aspects of collaboration include good teamwork, being professional, holding biweekly or weekly staffing, and having good access to supervisors.

Many FCS workers (61%) reported that all was going well in terms of between agency collaboration. Other FCS workers explained that there are issues with lack of information sharing, the need for FCS to be involved sooner, issues with group staffings, and issues with ISS overstepping decision making boundaries. Other themes that emerged among FCS workers regarding improvement of collaboration between agencies consist of increasing collaboration on decision making, emailing contact summaries, and specifying resource funding availability. It is also important to note that many FCS workers reported no need for improvement (45% of FCS workers).

FCS Supervisor Intra-Agency Collaboration Perspective

Themes of within agency collaborations that are going well include problem solving with CPS. FCS workers also explained that collaboration with the state office is going well (40% of FCS supervisors) with this office being easy to work with and good at problem solving. Supervisors also explained that FCS workers are working well with on-call FCS staff and are sharing information well. Other within agency roles that were explained to be going well were FCS supervisors sharing ideas and staffing cases. Child safety meeting and continued quality improvement units were reported to have open communication throughout the agency. It was also reported that the OCS liaison is accessible and supportive.

Areas of collaboration that are not going well include the transfer of information between CPS and FCS and the assigned FCS worker not attending the child safety meeting. It was further explained that there are high caseloads and poor communication with administrators.

FCS Supervisor Inter-Agency Collaboration Perspective

FCS supervisors reported that collaboration with the ISS agency included effective communication (82% of FCS supervisors) and effective staffing (43% of FCS supervisors). FCS supervisors also reported that through the ISS program there is another person to consult with, that attendance in meetings is good, and ISS works well as a team.

ISS workers reported on collaborations with DHS with workers providing information on what is going well and not well. ISS workers reported positive collaboration with CPS workers including good communication. Alternately, some ISS workers reported difficulties making contact with CPS workers. It was also described that there are times that not all information on the family is provided to ISS workers making it difficult to make initial contact.

In terms of collaboration with FCS workers, many ISS workers explained that collaboration can vary depending on the worker (43% of ISS workers). It was further explained that there is good communication with FCS workers (50% of ISS workers) and that staffing is going well (38% of ISS workers). Additionally, it was reported that there is frequent contact (25% of ISS workers), information is shared (19% of ISS workers), and that the collaboration with FCS is beneficial to the case (19% of ISS workers). On the other hand, ISS workers reported some negative aspects of collaboration with FCS including inconsistent communication (36% of ISS workers), no sense of urgency (14% of ISS workers), lack of understanding the treatment relationship between ISS worker and family (14% of ISS workers), and different expectations of the role of the ISS worker (14% of ISS workers).

Lastly, ISS workers discussed collaboration with the OCS liaison with workers reporting that the liaison is available, holds workers accountable, and provides answers to their questions. When probed of any negative collaboration aspects with the OCS liaison it was mainly reported there were none with a few ISS workers explaining there was limited contact.

Aspects of service delivery including procedures for determining eligibility, referring subjects for services, the array of services available, the number of children/families served and the type and duration of services received

CPS and FCS Opinions of ISS

CPS Opinion of ISS Services and Providers

When discussing the ISS program, CPS workers reported that it was a new resource that focuses on prevention. It is also important to note that some CPS workers reported they were not familiar with the ISS program (13% of CPS workers) and had few ISS cases.

Almost all the CPS workers (97%) familiar with the ISS program reported positive aspects of the program and workers (97% of CPS workers). In terms of positive attributes of ISS services, CPS workers reported that the program has increased intensity of services through conducting more frequent home visits and providing transportation for families. It was further reported that the ISS program is beneficial because it places another person in the home to ensure child safety and the program is helpful to the family. CPS workers also shared that the program has made services move more quickly through expediting referrals and ISS services beginning quickly. CPS workers also explained that ISS services are able to help with prevention.

CPS workers also commented on the specific positive qualities of ISS workers. Positive characteristics of ISS workers include having good communication through quick response and speaking to families in a way they can understand. It was further reported that the ISS worker is helpful and active with the family. CPS workers also explained that ISS workers provide support for families and assist in engaging and enrolling families in the services they need.

Some CPS workers reported no negative aspects of ISS (32%) while CPS workers did report that many workers are unfamiliar with the ISS program, there is a short time frame of services, and that ISS therapists are not always available due to the intensity of their caseloads and the limited number of ISS therapists.

FCS Opinion of ISS Services and Providers

FCS workers reported positive aspects of ISS including that the program provides extra support for the FCS worker. It was also reported that the high intensity of ISS is good, with ISS being in the home more frequently than FCS. CPS workers explained that ISS services link families to resources and is effective at monitoring and stabilizing the family while also encouraging behavior change. It was also reported that the program is effective at engaging families quickly. The most frequent negative aspect of the ISS program as described by CPS workers is the paperwork. Specifically, CPS workers feel the paperwork is redundant, pointless, and that terminology used in paperwork is different between ISS and FCS. Other negative characteristics of the ISS program is FCS workers reporting it is more work, overwhelming, and not appropriate for all families.

FCS workers also disclosed positive aspects of ISS workers, which included that they are "fantastic", have good communication skills, and are dedicated to families. It was also reported that ISS workers were good at building rapport and have a counseling background. FCS workers also reported that ISS is able to learn more about the family given their frequency of home visits and goals of the program.

Identify barriers of implementation, steps taken to address these barriers, and any lessons learned during implementation

Decision Making Changes

Both CPS and FCS workers reported on decision making changes that have taken place throughout implementation of the ISS program. Themes related to decision making included thinking differently about the removal of children (58% of CPS workers). Specifically, it was reported that ISS is another service offered to families, they are able to gather more family information through the ISS framework, and there are more individuals to monitor family safety.

From a FCS perspective, decision making changes include shifts in decision making (56% of FCS workers) including FCS and ISS working together and consulting with CPS to determine what is appropriate for families. This shift to FCS and ISS working together includes regularly discussing family status and home observations.

Changes to Create Program Success: Organizational and Practice

Organizational Changes

Many CPS workers reported no organizational changes or not being aware of organizational changes as a result of ISS. Of those that did indicate organizational changes to create success it was reported that ongoing buy-in from administration as well as providing clarity about the ISS program.

Similar to CPS workers, many FCS workers reported no organizational changes or not being aware of organizational changes. It was also reported by FCS workers that there has been a reduction of paperwork as well as lower caseloads and more updated Assessment of Child Safety (AOCS).

ISS workers reported organization changes in personnel that include hiring additional supervisors, beginning the stepdown to CHBS more quickly, and revising paperwork for the ISS program. Some ISS workers also reported no organizational changes to have taken place.

In terms of FCS supervisors, it was reported that organizational changes such as starting the stepdown process to CHBS sooner, decreasing paperwork, adding additional staff, and having open communication were all were positive organizational structural changes. It is also important to note that many FCS supervisors reported no changes or not being aware of changes.

Practice Changes

With regard to practice changes, many CPS workers explained they are not aware of any practice changes. CPS workers aware of practice changes reported that they now "click the button" for intervention within the PREMISS model, they conduct the child safety meeting more quickly, look more broadly for safety plan monitors to assist families, no longer discuss cases with child safety meeting facilitator prior to the child safety meeting, invite FCS workers to the child safety meeting, and provide contact information of the family to ISS worker. These are all critical changes to the success of ISS that the evaluation team tried to address early on and that these interviews suggest have been at least somewhat effective.

Similar to CPS workers, many FCS workers reported no practice changes to ensure program success. Some FCS workers did report practice changes related to paperwork, but rather changes made to clarify paperwork and assist in getting paperwork completed in a timely manner. It was also reported that there is an accelerated timeline for the child safety meeting, ISS introduces CHBS worker sooner, and that there are weekly meetings with ISS families.

ISS workers reported practice changes that include varying visit times based on family needs, expediting contact with the family, starting the referral process sooner, and emphasizing self-care. Some ISS workers also reported there were no changes in practice to ensure success.

FCS supervisor practice changes included improving communication, making changes to timelines, and improving the process of transferring cases.

Implementation: Strengths and Concerns
Strengths of Implementation

In terms of implementation strengths, CPS workers explained that the training was useful and the program allowed for additional services to be offered to families. It was also reported that the ISS program had a positive impact on families. Specifically, it was reported that the ISS model allows for children to not be removed from their home.

Strengths from an FCS worker perspective include having the support of ISS workers. This support includes ISS workers connecting families to resources, dedicating time to the family, and providing a professional experience. It was also reported by FCS workers that a strength of the ISS program is the design, quick implementation of the ISS program, listening to feedback from DHS workers, and families getting immediate services.

FCS supervisors reported strengths in the design of the ISS project and the following of its protocols. It was also reported by FCS supervisors that ISS workers are in the home quickly and engaging the family. Supervisors also explained that a strength was the initial training for the ISS program.

ISS workers reported positive aspects of implementation including quick engagement of families, administrators asking for feedback from ISS workers/supervisors, and that the program keeps children in their home.

Strengths of implementation was also discussed from an administrative perspective and included the ISS program having master's level clinicians and providers that can offer transportation. It was also reported that DHS has a high a level of involvement with the ISS team, the administrative team has the flexibility to make changes as implementation progresses, and administration is dedicated to the waiver project.

Concerns of Implementation

CPS workers voiced concern over the design of ISS services through explaining that the stepdown to CHBS happens too quickly for some families and the length of the program is too short. CPS workers also reported that the investigation time frame is too short and that the program is not available to all families. It was also reported that there are issues with the PREMISS model including the computer selecting families and that the PREMISS button should be labeled more clearly to indicate the importance to ISS. In terms of execution, CPS workers indicated that there is not enough training and that when training happens, it is overwhelming, and there is a time lag between training and implementation of the program. No concerns or unsure of concerns was reported by CPS workers regarding the management of ISS services.

From an FCS worker perspective, several concerns of implementation were discussed. Specifically, FCS workers (similar to CPS workers) reported design concerns of selecting the appropriate families for ISS services. FCS workers also reported concerns with documentation of services as usual (SAU) cases, lack of training, and an overall frustration with the research aspect of the ISS program. Additionally, FCS workers reported concern that services should be lengthened. Other concerns related to execution of ISS services included delayed FCS involvement, insufficient training with workers not being prepared to do the work, ISS worker turnover, and the sudden implementation of the ISS program. There were limited concerns with the management of ISS services with the lack of feedback being the only concern reported.

ISS worker concerns with implementation were similar to CPS and FCS workers in that there are concerns that six weeks is not enough time for families to make adequate behavior change. ISS workers also reported concerns with redundant and excessive paperwork. In terms of problems with ISS execution, it was reported that the ISS workers are not always present at the child safety meeting. In terms of management, an ISS worker reported there are inconsistencies in FCS units and their ability to implement ISS services, there is poor communication of feedback during case reviews, and there are limited referrals at times.

The theme of concerns over length of services was also echoed by FCS supervisors with supervisors reporting six weeks may not be enough time for families to make adequate behavior change. Supervisors also expressed concern over excessive paperwork, ISS not being appropriate for all families randomized, and issues with workers not understanding the PREMISS selection process. Concerns related to project execution included the gap between training and implementation of the service, CPS not transferring cases to FCS in a timely manner, more training needed, and the infrequency of ISS cases. In terms of concerns of the management of the ISS program, FCS supervisors explained that staff turnover, there not being a dedicated FCS team to ISS services and concerns that multiple ISS cases can be overwhelming.

Administrators explained that concerns of implementation included the design not being financially viable in rural locations and issues with rolling out implementation in several regions simultaneously. In terms of administrator concerns on the execution of implementation, it was disclosed that there was inadequate initial training, difficulties getting referrals, expansion of criteria model allowing increased family risk to participate in the program, and fear that a critical incident could jeopardize the future of the program.

What Needs to Happen for Program Success

For program success, CPS workers reported the need for clear program guidelines that explain inclusion and exclusion criteria. It was also explained that DHS needs the ability to directly refer families to the ISS program. Other recommendations for success include more ISS workers to serve families and some CPS workers indicating that nothing is need for success. Key areas of focus for ongoing success included leadership, support and, supervision. In terms of leadership it was reported that people are needed to market or "push" the ISS program. It was also reported that FCS supervisors are needed and more time is needed to meet with workers. Lastly, in terms of support it was reported that more training is needed with CPS workers needing more education on ISS services.

FCS workers reported that for future success there needs to be increased collaboration between FCS and the family, as well as FCS and ISS. Further, FCS workers explained that paperwork needs to be reduced and lower caseloads are needed. FCS workers also reported that ISS services need to be longer.

ISS workers reported the need for good leadership, supervision, and support to ensure program success. ISS workers also discussed the need for good communication from the beginning of the case and continued collaboration throughout. It was also reported that DHS needs to set clear expectations for the family. Other recommendations for success include hiring more ISS therapists, more training focused on clarifying the rolls of ISS, CHBS, and FCS, and maintaining good leadership. In terms of support, ISS workers explained that continued, ongoing support is needed.

FCS supervisors reported that for future success the process of selecting ISS family's needs to change (28% of FCS supervisors). Specifically, it was reported that workers need to have the ability to refer families directly to ISS services and have more availability of services for families in need. It was also reported that the documentation requirements for ISS need to change (14% of FCS supervisors), there needs to be improved communication (10% of FCS supervisors), and that they need to "click the button" for the PREMISS model to run (10% of FCS supervisors). Some FCS supervisors also explained that no changes are needed for success (14% of FCS supervisors) and that the program is already successful (21% of FCS supervisors).

Administrators reported that for ongoing success there needs to be ongoing funding availability for prevention services and accountability for workers at all levels.

Overall Recommendations on Policy and Practice

In summary, qualitative interviews focused on the implementation of the ISS program and suggest many aspects of the program that are successful. Data also illuminates some ways that practice and child safety has been improved. These interviews also suggest concrete areas for improvement and suggestions for program redesign, including lengthening the service term for some families.

Specific to recommendations on policy and practice, workers suggested several key areas for improvement. These recommendations include gathering feedback from frontline workers regarding

implementation changes. A broader communication issue was reported in interviews with ISS supervisors when describing that the hierarchy within DHS makes it difficult for ISS workers to know who to contact with issues. Workers also explained that protocol timelines can be challenging. Broader concerns were related to funding and budget deficits as well as service delays caused by the bidding process with state contracts. Staff turnover was also cited as an issue at the CPS and FCS worker level as well as the administrative level. One of the biggest issues discussed was the potential for children to be harmed while enrolled in the ISS program with ongoing monitoring needed.

In terms of overall strengths of implementation of the ISS program, workers explained that the program helps families and allows for children to not be removed from the home. Training was also reported as a strength. Concerns focused on stepdown to CHBS happening too quickly and the program being too short for some families to see change. Although training was indicated as a strength, it was also discussed as a concern. Therefore, the importance of ongoing training on the ISS program, the importance of the PREMISS model, and high frequency of training is evident. Ongoing training will help workers feel confident in how to complete tasks for the program in a timely manner so referrals are not delayed and help them feel comfortable that the correct families are being referred to the program. It is also important to note that some administrators voiced concern that the design may not be as financially viable in rural areas.

Supplemental Survey Following Qualitative Interviews

In addition to qualitative interviews, staff who participated were invited to take a brief online survey addressing a series of outcomes related to the implementation process. Invitations were sent to 111 participants, and a total of 31 individuals (including ISS workers, DHS workers, and supervisors) responded. The evaluation team attempted to increase response rate by sending follow-up invitations and reaching out via email. Because of this low response rate, no results are presented by worker type. The measures included in this survey are listed in *Measures Appendix* and include items from the Children's Service Survey, the Implementation Innovation Questionnaire, the Evidence-Based Practice Attitudes Scale (EBPAS)[20], and items relating to staff turnover intentions. The measures were chosen to address each of the following process implementation outcome constructs:

- staff attitudes toward the intervention
- measures of staff-perceived job autonomy
- staff turnover intentions
- dispositional innovativeness (predisposition to change)
- system network collaboration
- team climate
- perceived leadership quality

Each of these measures are described in their respective sections. Note: Results from the workers' survey are not changed from the January 2019 semi-annual report, because no additional responses have been received.

---

[20] Aarons, G. A., Glisson, C., Hoagwood, K., Kelleher, K., Landsverk, J., & Cafri, G. (2010). Psychometric properties and U.S. national norms of the evidence-based practice attitude scale (EBPAS). *Psychological Assessment, 22*(2), 356-365. doi:http://dx.doi.org/10.1037/a0019188

Implementation Innovation Questionnaire

The Innovation Implementation Questionnaire (IIQ) was developed to address 24 dimensions that influence individuals in the adoption of an innovation, as well as the degree to which each dimension affects how consistently the program is implemented with the way it was intended or originally described. The first 19 dimensions of the IIQ were applicable to all individuals who participated in the qualitative interviews, including ISS Therapists, ISS Supervisors, FCS workers, FCS supervisors, and DHS administrators/program managers.

The response scale for each item 1 to 5, with 5 being the most favorable view for each given aspect, although higher scores on certain items may not be "favorable" (e.g., how much has job changed). For each of the IIQ dimensions, a sub-question was asked regarding the level of influence each respective item has on the degree to which the program has been implemented as presented in the trainings. The response scale for the influence items is also a 1 to 5 response scale from 1-very little/no influence to 5 - very high influence.

Although IIQ items are somewhat heterogeneous, results were aggregated into four broad categories for reporting purposes. The categories chosen by the evaluation team include Logistics (i.e., difficulty of adopting program, changes in workload and job duties, and programmatic supports), Interpersonal Interactions (i.e., perceptions surrounding team's views of ISS, level of interaction within and outside of team/organization), Buy In (i.e., overall belief that program supports the organization's mission, belief that the intervention will be effective, personal priority, etc.), and Leadership Support (i.e., how supported workers feel by leadership, how empowered to make adaptations where necessary, the level of priority leadership places on the intervention). Because the presence of incentives was not relevant given the constraints of the program, we have opted not to report on these items to prevent artificially lowering items in a subscale.

Because the total number of participants within each group was small, results are presented in aggregate for all job titles on the table below. The scores presented in the table are means for all items within the subscale.

As seen in Table 19, scores tended to be highest on the items regarding interpersonal interactions and personal buy-in. The lowest mean score was on the Leadership Support subscale.

TABLE 19 - IIQ SUB-SCALE ITEM MEANS (N = 31)

	Mean (sd)
Logistics (4 items)	3.5 (sd = .84)
Interpersonal Interactions (4 items)	4.0 (sd = .66)
Buy-In (5 items)	3.9 (sd = .63)
Leadership Support (4 items)	3.3 (sd = .81)

TABLE 20 - IIQ ITEM-LEVEL MEANS

	mean	SD	min	max
1. Degree to which staff are included in organizational feedback loop	3.30	1.32	1.0	1.0
2. Similarity of innovation to mission	4.22	0.58	3.0	3.0
3. Degree to which change has impacted work	3.63	1.47	2.0	2.0
4. Organizational changes prior to implementation	3.96	1.32	2.0	2.0
5. Number respected peers showing support for implementation	4.12	1.17	1.0	1.0
6. Number respected peers in opposition to program	4.56	0.87	1.0	1.0
7. Level of interaction with team	3.43	1.53	1.0	1.0
8. Perception of evidence that program fidelity will support goals of program	4.05	1.07	2.0	2.0
9. Level of organization priority	3.88	0.67	3.0	3.0
10. Level of personal priority	3.92	0.81	3.0	3.0
11. Level of support for program by leadership	3.88	0.83	3.0	3.0
12. Level of change in intra-organizational interactions	4.00	1.22	1.0	1.0
13. Ease of learning delivery of program	3.83	0.82	2.0	2.0
14. Level of alignment with existing regulations	4.61	0.66	3.0	3.0
15. Incentives to implement	1.42	1.07	1.0	1.0
16. How many tools/supports are available to support implementation?	2.63	1.24	1.0	1.0
17. How empowered do you feel to adapt ISS to make it work better?	2.08	1.02	1.0	1.0
18. Does program address key issue(s)?	2.54	1.35	1.0	1.0
19. Incentives contingent on performance	1.29	0.91	1.0	1.0

Process Construct: Team Climate

As seen in the table above, responses regarding team interaction and support (items 5, 6 (reversed coded), 7, 11, and 12) show relatively positive responses (mean score range for those items 3.4-4.6) indicating a culture of coordination and interconnectedness amongst coworkers and between staff.

Process Construct: Staff Autonomy

The responses on the item regarding *how individuals perceive their ability to make necessary adaptations to suit the needs of the client* indicates mixed responses. About half of respondents indicated a positive response on this item, but half responded with neutral or negative/strongly negative responses.

Staff Attitudes Toward the Intervention, Dispositional Innovativeness

To address staff attitudes toward the intervention, only two of the EBPAS scales were used for the purposes of this survey: the Openness scale - how willing the individual is in general to try new interventions and evidence-based practices; and the Divergence scale - how much staff perceive the new practice is less useful or important than their own experience or current practice. Scores are presented based on the agency for which the individual worked at the time of the survey. Respondents were asked to rate their level of agreement with each of the statements corresponding to these scales. The response options were a 5-point Likert-type scale follows: 1 = not at all, 2 = to a slight extent, 3 = to a moderate extent, 4 = to a great extent, and 5 = to a very great extent.

Response rate to this survey has been consistently low. Therefore, for this measure, only descriptive information is provided. See Table 21Table 20.

TABLE **21** EBPAS **S**UBSCALE **M**EANS BY **W**ORKER **T**YPE

Agency	Openness Scale (Mean; sd; n)	Divergence Scale (Mean; sd; n)
All	3.94; sd =0.57; n=27	2.09; sd =0.67; n=28
ISS	5.00; sd =0; n=3	1.17; sd =0.14; n=3
FCS	3.83; sd =0.45; n=15	1.95; sd =0.55; n=16
CPS	3.72; sd =0.51; n=8	2.69; sd =0.5; n=8

These analyses are in response to the Process study constructs of staff attitudes toward the intervention and dispositional innovativeness. Overall, staff tended to have high mean scores for Openness across the board. In general, these scores indicate a favorable attitude toward the intervention and high levels of dispositional innovativeness/flexibility. There were also moderately low scores on the Divergence scale, indicating the staff do not perceive great personal conflict between their current practice and the adoption of the ISS innovation. Unsurprisingly, ISS respondents tended to report the highest agreement with Openness and the lowest with Divergence, indicating the greatest receptivity to the program. FCS respondents indicated the lowest agreement with Openness and the highest with Divergence indicating the least receptivity.

Turnover Intentions

The following bar plots in Figure 6 show the turnover intentions of survey respondents. It appears a large percentage of those surveyed, particularly ISS workers, indicate they are actively seeking another job and would be willing to quit for a better job. However, nobody indicated they were seriously considering quitting at the time of the survey (by selecting "Strongly Agree" to any of these items). These concerns have been brought to light with DHS staff; however, DHS staff directly involved with the Title IV-E Waiver implementation believe they do not have the capacity to implement changes for these concerns outside of letting higher-level administrators know about the issues. The evaluation team has relayed these concerns to OKDHS leadership.

FIGURE 6 BAR PLOT TURNOVER INTENTIONS BY WORKER TYPE
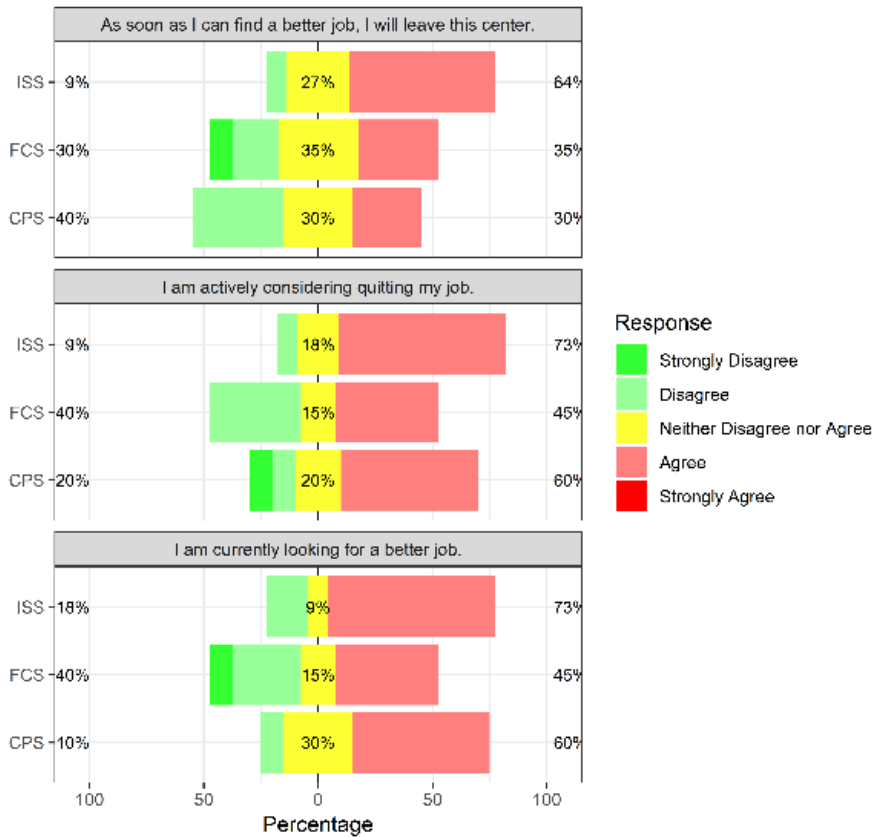
Working Alliance Inventory (WAI)

The WAI[21] is a scale administered to measure levels of coordination, positive experience, and overall alliance between a provider and a client. The procedure for obtaining WAI forms from clients at Stepdown is to provide them with a WAI form and a pre-addressed, stamped envelope to provide some assurance of anonymity in responses regarding their workers. At the Stepdown event, providers (for ISS clients - the ISS provider, for SAU clients - the FCS worker) fill out a provider version of the WAI form. At 6 months, the CHBS worker fills out the same form. There have been some inconsistencies in receiving these forms throughout the project. The percentages vary greatly with a consistently higher compliance rate for provider forms, and a higher rate for the ISS received group. SAU and ISS Not workable tended to have the lowest percentage of forms received.

The number of forms received for parents and providers is shown in Table 22.

---

[21] Horvath, A. O., & Greenberg, L. S. (1989). Development and validation of the Working Alliance Inventory. *Journal of counseling psychology*, *36*(2), 223.

TABLE 22 - WORKING ALLIANCE MEASUREMENT SAMPLE SIZES

Group Actual	Measurement point	Parent WAI Scores Received	Provider WAI Scores Received
ISS received	Stepdown	323	300
ISS received	Six Months	124	112
SAU	Stepdown	286	229
SAU	Six Months	89	65
ISS not workable	Stepdown	77	62
ISS not workable	Six Months	7	7

This 12-item measure is divided into three subscales: 1. Goals - How well the parent/provider believes the treatment will meet goals; 2. Tasks - how well parent/provider agree on steps to take to meet the goals of treatment; 3. Bond - How well the parent/provider perceive their relationship. Each item presents a Likert-type responses scale of 1 - least favorable to 7 - most favorable.

Two items needed to be reverse coded so that the response scale was consistent for scoring purposes. For these two items, it was observed that some clients may have selected the negative response inadvertently when trying to give across-the-board positive responses. These individuals were identified by determining if other items in that subscale were both rated either 6 or 7. If this was the case, these items were not reverse coded to account for this. This reduced the number of individuals with high scores on two other items in the subscale but low scores on these two items from about 9% to about 5%.

To account for missingness, if there were less than three items total missing from the form, the mean for other items within that subscale was used to impute values for missing items. Subscale scores were calculated by summing the responses from the scale for each item.

To illustrate the differences between groups, the boxplots below, separated by region, provide a visual representation of the spread of scores. The correlation coefficient for provider-client pairs is r = 0.47, indicating modest levels of agreement. Data at the six-month measurement period are scarce, but the findings that the ISS group reports higher levels of working alliance remain true at the six months measurement, even when the ISS client's worker is no longer an ISS therapist.

FIGURE 7 - PARENT TOTAL SCORES ALL REGIONS COMBINED
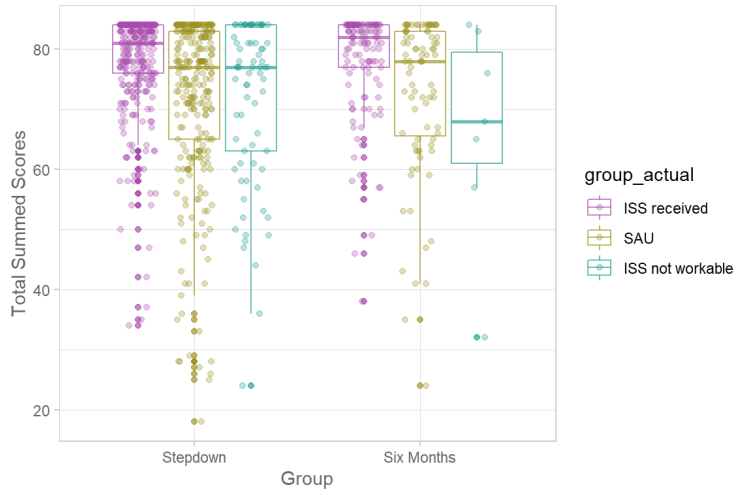
FIGURE 8 - PROVIDER SCORES ALL REGIONS COMBINED

FIGURE 9 - PARENT SCORE DISTRIBUTIONS BY REGION
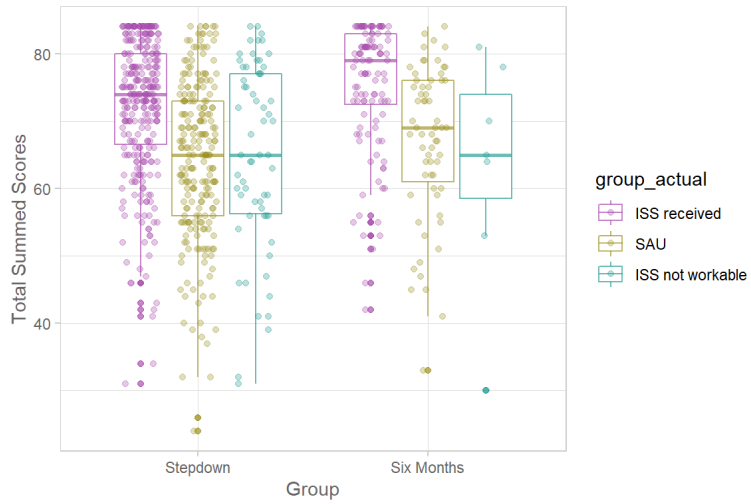
FIGURE 9 - PARENT SCORE DISTRIBUTIONS BY REGION

FIGURE 10 - PROVIDER SCORE DISTRIBUTIONS BY REGION
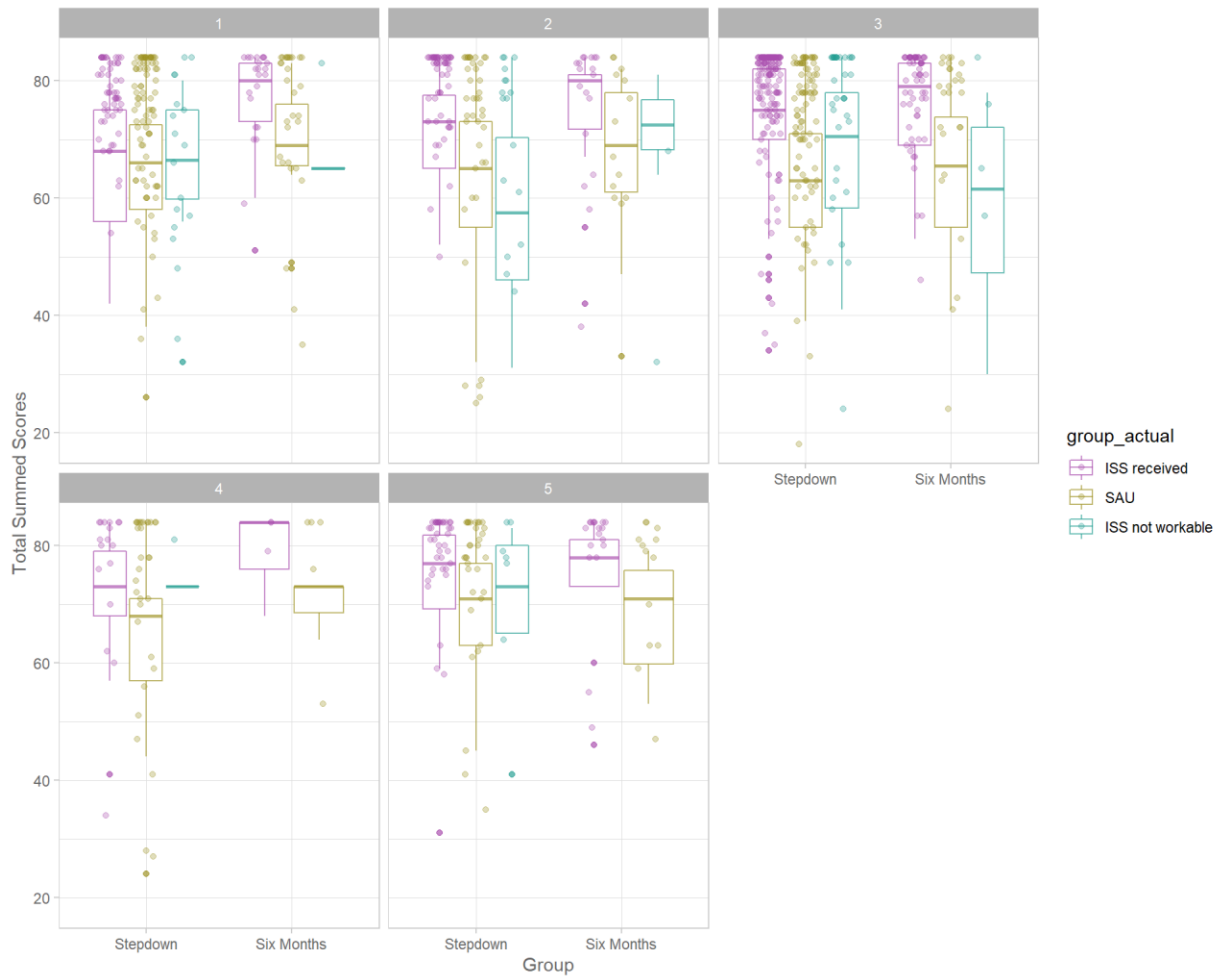
Analysis – WAI Total Score

As visualized in the plots above, ISS consistently had higher scores. GLM regression models are shown below in Table 23, but because the error distribution was non-normal, and the data were far left-skewed, we also show results for a more conservative regression model using maximum likelihood estimation to estimate the differences between ISS and other groups. As seen below, parents belonging to the SAU group scored about 3.28 (95% CI [-4.54; -2.02]) points lower on total WAI score than ISS received. Similarly, ISS not workable had a score of 3.57 (95% CI [-5.51; -1.63]) points lower than ISS received. These results were significant at p < .01.

TABLE 23 - WAI PARENT SCORE MODELS SUMMARY

	GLM Regression	Robust Regression
Intercept	77.41*	77.81*
	[75.26; 79.56]	[76.34; 79.28]
SAU (reference group ISS received)	-5.47*	-3.13*
	[-7.34; -3.59]	[-4.42; -1.85]
ISS not workable (reference group ISS received)	-5.59*	-3.57*
	[-8.42; -2.76]	[-5.50; -1.63]
Region 2	-0.05	1.47
	[-2.87; 2.77]	[-0.46; 3.40]
Region 3	0.73	1.32
	[-1.47; 2.92]	[-0.18; 2.82]
Region 4	0.02	1.80
	[-5.18; 5.22]	[-1.76; 5.36]
Region 5	1.92	2.63*
	[-1.29; 5.14]	[0.43; 4.83]
Six Months (reference Baseline)	-0.22	-0.24
	[-2.29; 1.86]	[-1.66; 1.18]
AIC	5234.38	5267.00
BIC	5275.06	5307.00
Log Likelihood	-2608.19	-2624.00
Deviance	86265.50	90480.00
Num. obs.	679	679

Model Predicted Means

Table 24 and Figure 11 below shows predicted score means for the different observed groups (using means from Robust Regression). As shown below, the trend in higher scores for ISS Received clients remains at the six-month event, even though the targeted worker on the WAI is no longer the ISS therapist.

FIGURE 11 - MODEL PREDICTED WAI MEAN SCORES BY GROUP & DATA COLLECTION TIME
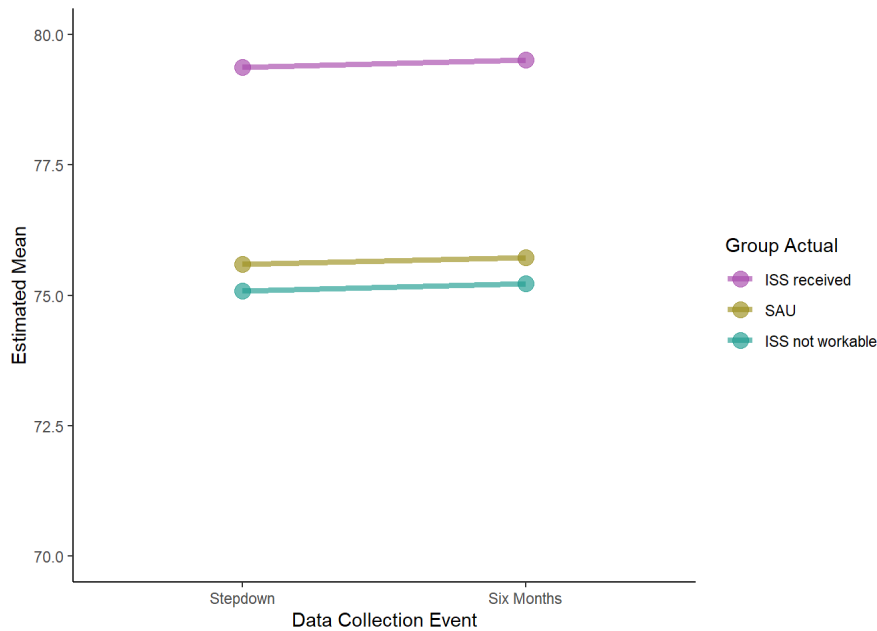
TABLE 24 - WAI MODEL PREDICTED MEAN PARENT SCORE

Group Actual	Data Collection Event	Estimated Mean & 95% Confidence Interval
ISS received	Stepdown	79.37, 95% CI [78.45, 80.3]
SAU	Stepdown	75.6, 95% CI [74.68, 76.52]
ISS not workable	Stepdown	75.09, 95% CI [73.29, 76.9]
ISS received	Six Months	79.51, 95% CI [78.25, 80.77]
SAU	Six Months	75.73, 95% CI [74.43, 77.04]
ISS not workable	Six Months	75.23, 95% CI [73.08, 77.37]

Customer Satisfaction Surveys

Customer satisfaction surveys consisted of 12 Likert-type items on a 4-item response scale with 1 being the lowest (least satisfied) and 4 the highest ratings (most satisfied). These scores were summed for each participant. The sample sizes for the surveys are presented in Table 25. Client perceptions for each of these items are presented on the bar plot below in Figure 12. Overall, client reports of satisfaction on all measures tended to be positive with the ISS received group reporting almost universally positive responses to all items. Percentages of responses rated as negative are shown next to the assignment group on the left, and percentages of responses rated as positive are shown on the right.

TABLE 25 - SAMPLE SIZES FOR SATISFACTION SURVEYS

	Sample Size
ISS Not Workable	64
SAU	240
ISS Received	289

FIGURE 12 - CLIENT SATISFACTION ITEM RESPONSES
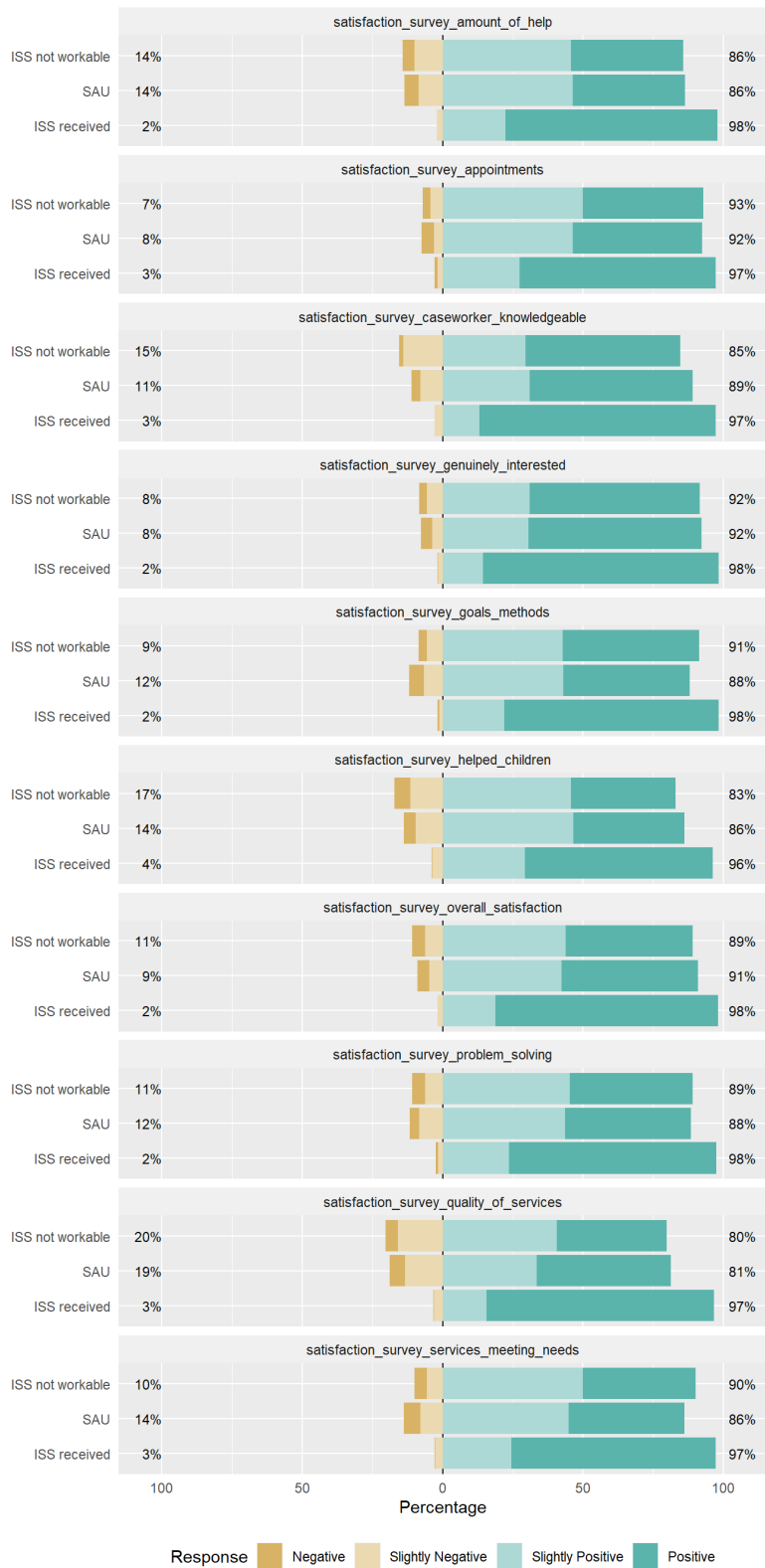
Client Summed Satisfaction Score

To analyze between group differences, a GLM model was used. Additionally, a robust regression using maximum likelihood estimation was also conducted to show a more conservative estimate. As seen below, the mean overall score is 2.86 (95% CI [-3.75; -1.97]) lower for SAU and 2.68 (95% CI [-3.93; -1.43]) lower for ISS Not Workable. These results indicate greater satisfaction for ISS received clients overall compared to other groups.

TABLE 26 CLIENT SATISFACTION MODEL RESULTS

	GLM Model	M-estimate Model
Intercept	34.68*	34.50*
	[33.60; 35.76]	[33.51; 35.48]
SAU	-3.58*	-2.90*
	[-4.59; -2.56]	[-3.83; -1.98]
ISS Not Workable	-3.03*	-2.69*
	[-4.38; -1.67]	[-3.93; -1.45]
Region 2	0.30	0.95
	[-1.08; 1.67]	[-0.30; 2.21]
Region 3	1.91*	2.39*
	[0.74; 3.08]	[1.32; 3.46]
Region 4	4.39*	4.26*
	[2.03; 6.75]	[2.11; 6.42]
Region 5	3.56*	4.09*
	[1.89; 5.22]	[2.57; 5.61]
AIC	2362.74	2370
BIC	2394.71	2402
Log Likelihood	-1173.37	1177
Deviance	8072.62	8218
Num. obs.	402	402

* 0 outside the confidence interval

Estimated Means for Summed Client Satisfaction Items

Means predicted for individual groups by region are presented in Figure 13 and Table 27, ISS is consistently higher across all regions. The differences in scores are small, but significant.

FIGURE 13 - ESTIMATED MEANS CLIENT SATISFACTION BY REGION & GROUP
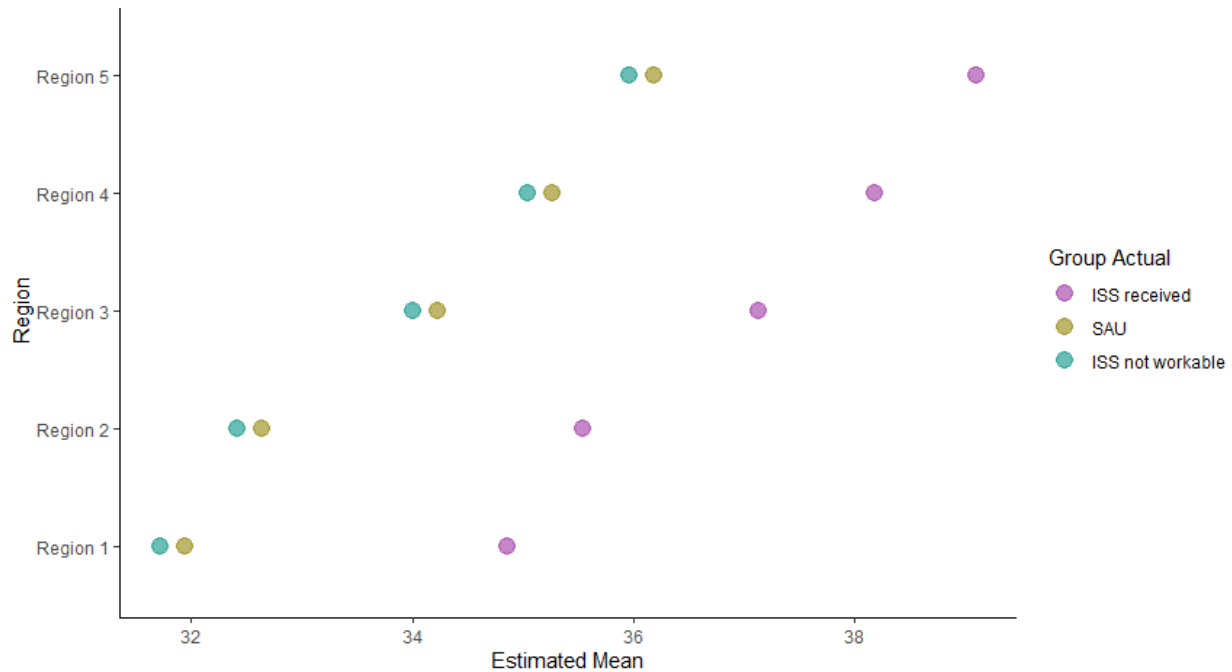
TABLE 27 ESTIMATED MEANS FOR CLIENT SATISFACTION BY REGION & CONDITION

Region	Condition	Estimated Mean [95% CI]
Region 1	ISS received	34.88, 95% CI [33.98, 35.78]
	SAU	31.8, 95% CI [30.97, 32.64]
	ISS not workable	31.77, 95% CI [30.44, 33.11]
Region 2	ISS received	35.5, 95% CI [34.62, 36.38]
	SAU	32.42, 95% CI [31.46, 33.38]
	ISS not workable	32.39, 95% CI [31.1, 33.68]
Region 3	ISS received	37.1, 95% CI [36.45, 37.76]
	SAU	34.03, 95% CI [33.21, 34.85]
	ISS not workable	34, 95% CI [32.83, 35.16]
Region 4	ISS received	38.27, 95% CI [36.86, 39.68]
	SAU	35.19, 95% CI [33.88, 36.51]
	ISS not workable	35.16, 95% CI [33.42, 36.91]
Region 5	ISS received	39.14, 95% CI [38.08, 40.21]
	SAU	36.07, 95% CI [34.97, 37.17]
	ISS not workable	36.04, 95% CI [34.55, 37.53]

In summary, at least as it pertains to client satisfaction and worker alliance, ISS consistently results in higher levels of satisfaction and engagement, which the field knows is key to achieving positive outcomes for the service delivery system.

Outcomes Study

The following section evaluates the effectiveness of ISS utilizing three different approaches: Intent-to-Treat (ITT), As Treated (AT), and Treatment on the Treated (TOT). Intent-to-Treat refers to an analysis that groups the observations by the treatment group assigned (ISS Assigned or SAU). This is a conservative approach for estimating treatment effects and recognizes that in any application of an intervention some people will not start, will drop out, or fail to complete the intervention. However, the impact of those assigned to the treatment but did not receive it at all or in full are included in the estimates. In the Intent-to-Treat approach observations in ISS Received or ISS Not Workable would both be grouped into ISS Assigned, and ISS Assigned treatment effects would be compared to SAU.

An analysis that uses the As Treated approach groups the observations by if they received ISS or not, not adjusting for and ignoring attrition. This type of analysis would compare observations that received ISS with observations that received SAU.

Last, Treatment on the Treated (TOT) analyses accounts for the fact that some people assigned to an intervention may not have received it and, in some cases (not ours), that some people assigned to the control group may have received the intervention. The TOT is a form of a local average treatment effect (LATE), which implies that it estimates the effect of treatment on a subgroup of individuals. In this instance, the subgroup represents (a hypothetical group of) families who would actually receive the treatment provided they were assigned to the ISS condition.[22] Sometimes this subgroup is referred to as "compliers," but in our case, these are families who were (1) assigned to ISS, (2) actually offered ISS by DHS (a case determined to be "workable"), (3) accepted these services, and (4) then complied with treatment, i.e., started ISS. TOT analyses attempt to isolate the effect of treatment among those receiving treatment versus those in the control/comparison condition who would have received the treatment if only assigned to the other condition. Implicitly, this analysis eliminates/excludes the outcomes of so called "never-takers" from both our ISS and SAU groups, i.e., families that would not receive treatment even if assigned to the ISS condition. There are number of ways to estimate the TOT. Herein, we use the instrumental variable (IV) approach popularized by economists.[23]

This section evaluates the effectiveness of ISS on eight key areas related to family outcomes. Where it is relevant to make comparisons (Outcomes Questions 1, 2, 6, 7, and 8), findings are presented using Intent-to-Treat (ITT), As-Treated (AT), and Treatment-effect-on-the-Treated (TOT) analyses. Using ITT analysis, we compare group outcomes based on the randomized treatment group, regardless of whether they received the treatment, to Services as Usual. AT and TOT analyses analyzes data across three groups where the randomized group to receive treatment is identified, but is separated between those who did and did not receive the treatment. The reasons families might have been randomized to ISS without receiving ISS were outlined in the introduction to this report. This group is referred to as "ISS Not Workable". The group that was randomized to and received ISS is referred to as "ISS Received", and those randomized to Services as Usual are referred to as "SAU".

---

[22] Imbens, G. W.; Angrist, J. D. (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica, 62*, 467. doi:10.2307/2951620

[23] Angrist, J. D.; Imbens, G. W.; Rubin, D. B. (1996). Identification of Causal Effects Using Instrumental Variables. *Journal of the American Statistical Association, 91*, 444–455. doi:10.1080/01621459.1996.10476902.

Using ITT analysis, we can provide a conservative approach to avoid issues with biased case selection for ISS (e.g., choosing low-risk cases only). To provide a more in-depth view of the effects of ISS on those who actually received ISS, we are also evaluating from the AT and TOT perspective.

Together this suite of analyses should provide us with a realistic range of observed effects on ISS and additional confidence in our results. Because the top reason for "not working" an ISS case was court involvement prior to randomization, we also occasionally present a subgroup ITT analysis that only includes ISS and SAU families who were randomized prior to any court intervention. This final subgroup analysis complements the TOT models, providing an estimate of the treatment effect among those who had a realistic chance at receiving ISS services.

As discussed previously, PREMISS scores are based on logit estimates for likelihood of reunification within an unusually short timeframe. These may be transformed into probabilities by using the formula $1/(1+\exp(-x))$. The mean scores probability of a quick reunification were .21 for ISS received, .20 for SAU, and .21 for ISS not workable. The risk scores and ranges were comparable among all three groups. ISS received had a mean probability of reunification of .21 with probabilities ranging from .06 to .67. SAU had a mean probability of .20 and a range of .05-.65. ISS Not Workable had a mean probability of reunification of .21 with a range of .06-.77. These PREMISS scores are included in analyses for Outcomes 1, 2, 4, and 6 to increase statistical power. More detailed descriptive statistics by group and event can be found under analysis in Outcome 1.

Outcome Results

Question 1: Relative to the SAU condition, does ISS result in greater reductions in the number of safety threats identified in post-intervention AOCS administrations?

> The key takeaways for this question include:
>
> - The greatest predictor in future safety threats i.e. safety threats at stepdown and the six-month observations, was the number of safety threats at the referral AOCS
> - The ISS cohort displayed the lowest values in safety threats at the stepdown and six-month observation periods.

Assessment of Child Safety (AOCS) forms are collected by FCS and/or Permanency Planning at baseline, stepdown, and at the 6 months for both ISS and SAU families. AOCS forms track the number of safety threats and protective capacities for families in both ISS and SAU at each measurement. The three most common safety threat items across the entire cohort reported at baseline include: 1) Person responsible for care (PRFC(s)) will not or cannot control behavior (n=2825), 2) PRFC(s) in the home lack the knowledge, skills, motivation, or abilities to perform parental duties and responsibilities (n=914), and 3) Living arrangements seriously endanger a child's physical health (n=174). This section of the report provides an accounting for the summation of all 9 safety threat items.

The use of the AOCS in practice is as follows. Safety threats is tied to the presence of unsafe behaviors or conditions. The AOCS provide information regarding family functions and protective capacities. The CW

Specialist uses a safety threshold to understand how likely a safety threat is currently or in the future. This is then used to determine if the child is safe, and if protective capacities can be used to manage or control for the safety threat. The presence of a safety threat is then addressed with safety planning to determine if there is a PRFC that has protective capacities to ensure the safety of the child.

Methods

Analyses of covariance (ANCOVAs) were used to assess group differences in number of observed safety threats at each follow up time point while accounting for baseline safety threat counts. The ANCOVA approach yields an interpretable metric which assess mean group differences at each time point while controlling for nuisance variables. While this approach is easy to implement and interpret, it requires the exclusion of participants without observations at both the intake and succeeding time points. Accordingly, the SAU group had 458 cases at the stepdown time point, and 187 at the six-month time point across all regions, whereas the ISS group had 278 cases at the stepdown time point, and 85 at the six-month time point across all regions.

This report presents the descriptive and inferential statistics for regions 1, 2, 3, 4, and 5, separately and all regions combined. TABLE 28-TABLE 33 present means and standard deviations for the number of safety threats, and total number of individuals in Regions 1, 2, 3, 4, and 5 by event for the Intent-to-Treat cohort (ITT). The ITT cohort includes all observations assigned to the ISS treatment paradigm.

 All question 1 analyses are conducted using both ITT and Treatment on the Treated cohorts (TOT). The TOT Cohort includes all participants in the actual treatment received: ISS received, ISS not workable-which is essentially another SAU condition, and service as usual (SAU).

Error! Reference source not found. depicts mean safety threats by each region for the ITT groups, and Figure 14 depicts the same for the TOT groups.

TABLE 28-TABLE 33 present means and standard deviations for the number of safety threats, and total number of individuals in Regions 1, 2, 3, 4, and 5 by event for the Intent-to-Treat cohort (ITT).

TABLE 28 - SUMMARY OF SAFETY THREATS BY EVENT FOR REGION 1– ITT GROUPS

	ISS Assigned		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(0, 4)	1.20 (0.58, 294)	(0, 5)	1.21 (0.63, 817)
Stepdown	(0, 5)	1.09 (0.97, 45)	(0, 3)	1.08 (0.66, 131)
6 Month	(0, 2)	0.15 (0.55, 13)	(0, 3)	0.71 (0.80, 38)

TABLE 29 - SUMMARY OF SAFETY THREATS BY EVENT FOR REGION 2 – ITT GROUPS

	ISS Assigned		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(0, 3)	1.16 (0.56, 221)	(0, 4)	1.11 (0.55, 415)
Stepdown	(0, 3)	1.14 (0.75, 42)	(0, 4)	1.11 (0.77, 66)
6 Month	(0, 2)	1.00 (0.63, 6)	(0, 6)	0.83 (1.21, 30)

TABLE 30 - SUMMARY OF SAFETY THREATS BY EVENT FOR REGION 3– ITT GROUPS

	ISS Assigned		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(0, 3)	1.19 (0.58, 680)	(0, 4)	1.16 (0.60, 769)
Stepdown	(0, 8)	1.00 (1.01, 131)	(0, 6)	1.03 (0.98, 148)
6 Month	(0, 2)	0.31 (0.55, 51)	(0, 9)	0.64 (1.27, 70)

TABLE 31 - SUMMARY OF SAFETY THREATS BY EVENT FOR REGION 4– ITT GROUPS

	ISS Assigned		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(0, 2)	0.92 (0.44, 85)	(0, 3)	1.13 (0.49, 350)
Stepdown	(0, 1)	0.57 (0.53, 7)	(0, 4)	1.00 (0.80, 32)
6 Month	(0, 0)	0.00 (0.00, 2)	(0, 1)	0.60 (0.52, 10)

TABLE **32** - SUMMARY OF SAFETY THREATS BY EVENT FOR REGION 5– ITT GROUPS

	ISS Assigned		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(0, 4)	0.97 (0.58, 295)	(0, 4)	1.00 (0.54, 486)
Stepdown	(0, 3)	0.75 (0.85, 53)	(0, 4)	0.88 (0.76, 81)
6 Month	(0, 1)	0.15 (0.38, 13)	(0, 3)	0.44 (0.64, 39)

TABLE **33** - SUMMARY OF SAFETY THREATS BY EVENT FOR REGIONS **1, 2, 3, 4,** AND **5**– ITT GROUPS

	ISS Assigned		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(0, 4)	1.13 (0.58, 1575)	(0, 5)	1.14 (0.58, 2837)
Stepdown	(0, 8)	0.98 (0.93, 278)	(0, 6)	1.03 (0.82, 458)
6 Month	(0, 2)	0.31 (0.56, 85)	(0, 9)	0.64 (1.03, 187)

In this reporting period, we also included referral score, i.e., child's PREMISS score, as an additional covariate in the model to adjust for the level of safety in families in all analyses. This score is a crucial component in determining a child's eligibility for ISS. Therefore, adjusting for such predictor in this treatment comparison is helpful in terms of adding statistical power, but also in terms of avoiding bias, particularly for (non-randomized) group effects estimated in the As-Treated models. Higher PREMISS scores indicate a safer home. TABLE **34**-TABLE **39** present means and standard deviations for the PREMISS scores in Regions 1, 2, 3, 4, and 5 by event for the Intent-to-Treat cohort (ITT).

TABLE **34** - SUMMARY OF PREMISS SCORE BY EVENT AND TREATMENT GROUPS FOR REGION **1**

	ISS Assigned		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(-2.57, 0.71)	-1.45 (0.73, 294)	(-3.17, 1.10)	-1.60 (0.70, 818)
Stepdown	(-2.56, 0.41)	-1.50 (0.71, 45)	(-2.56, 1.10)	-1.58 (0.74, 131)
6 Month	(-2.33, 0.41)	-1.09 (0.89, 13)	(-2.55, 1.10)	-1.64 (0.81, 38)

TABLE **35** - SUMMARY OF PREMISS SCORE BY EVENT AND TREATMENT GROUPS FOR REGION **2**

	ISS Assigned		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(-3.15, 0.73)	-1.28 (0.60, 221)	(-2.39, 0.78)	-1.26 (0.61, 415)
Stepdown	(-3.15, 0.73)	-1.47 (0.58, 42)	(-2.39, 0.78)	-1.14 (0.73, 66)
6 Month	(-1.82, -0.36)	-1.42 (0.53, 6)	(-2.39, 0.19)	-1.17 (0.75, 30)

TABLE 36 - SUMMARY OF PREMISS SCORE BY EVENT AND TREATMENT GROUPS FOR REGION 3

	ISS Assigned		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(-2.81, 1.23)	-1.40 (0.72, 687)	(-3.00, 1.08)	-1.47 (0.70, 786)
Stepdown	(-2.64, 1.23)	-1.29 (0.76, 131)	(-2.90, 0.63)	-1.40 (0.74, 148)
6 Month	(-2.64, 0.64)	-1.28 (0.83, 51)	(-2.47, 0.63)	-1.31 (0.74, 70)

TABLE 37 - SUMMARY OF PREMISS SCORE BY EVENT AND TREATMENT GROUPS FOR REGION 4

	ISS Assigned		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(-2.21, 0.04)	-1.47 (0.56, 85)	(-3.10, 0.49)	-1.49 (0.58, 350)
Stepdown	(-2.09, -0.69)	-1.57 (0.62, 7)	(-2.21, 0.11)	-1.51 (0.55, 32)
6 Month	(-1.86, -1.64)	-1.75 (0.16, 2)	(-2.21, -0.80)	-1.63 (0.48, 10)

TABLE 38 - SUMMARY OF PREMISS SCORE BY EVENT AND TREATMENT GROUPS FOR REGION 5

	ISS Assigned		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(-1.95, 0.82)	-0.86 (0.59, 307)	(-2.21, 1.44)	-0.95 (0.58, 495)
Stepdown	(-1.65, 0.82)	-0.86 (0.66, 53)	(-1.91, 0.59)	-1.02 (0.52, 81)
6 Month	(-1.61, 0.52)	-0.79 (0.71, 13)	(-1.91, -0.11)	-1.10 (0.45, 39)

TABLE 39 - SUMMARY OF PREMISS SCORE BY EVENT AND TREATMENT GROUPS FOR REGIONS 1, 2, 3, 4, AND 5

	ISS Assigned		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(-3.15, 1.23)	-1.29 (0.71, 1594)	(-3.17, 1.44)	-1.39 (0.69, 2864)
Stepdown	(-3.15, 1.23)	-1.27 (0.73, 278)	(-2.90, 1.10)	-1.36 (0.72, 458)
6 Month	(-2.64, 0.64)	-1.19 (0.81, 85)	(-2.55, 1.10)	-1.33 (0.72, 187)

TABLE 40-TABLE 45 present means and standard deviations for the number of safety threats, and total number of individuals in Regions 1, 2, 3, 4, and 5 by event for the AS Treated cohort (AT).

TABLE 40 - SUMMARY OF SAFETY THREATS BY EVENT FOR REGION 1 – AS TREATED GROUPS

	ISS Not Workable		ISS Received		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(0, 4)	1.22 (0.64, 205)	(0, 2)	1.15 (0.41, 89)	(0, 5)	1.21 (0.63, 817)
Stepdown	(0, 5)	1.44 (1.20, 18)	(0, 3)	0.85 (0.72, 27)	(0, 3)	1.08 (0.66, 131)

| 6 Month | (0, 0) | 0.00 (0.00, 2) | (0, 2) | 0.18 (0.60, 11) | (0, 3) | 0.71 (0.80, 38) |

TABLE 41 - SUMMARY OF SAFETY THREATS BY EVENT FOR REGION 2 – AS TREATED GROUPS

	ISS Not Workable		ISS Received		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(0, 3)	1.17 (0.61, 132)	(0, 2)	1.15 (0.49, 89)	(0, 4)	1.11 (0.55, 415)
Stepdown	(1, 3)	1.48 (0.59, 23)	(0, 3)	0.74 (0.73, 19)	(0, 4)	1.11 (0.77, 66)
6 Month	(1, 2)	1.25 (0.50, 4)	(0, 1)	0.50 (0.71, 2)	(0, 6)	0.83 (1.21, 30)

TABLE 42 - SUMMARY OF SAFETY THREATS BY EVENT FOR REGION 3 – AS TREATED GROUPS

	ISS Not Workable		ISS Received		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(0, 3)	1.16 (0.62, 432)	(0, 3)	1.23 (0.51, 248)	(0, 4)	1.16 (0.60, 769)
Stepdown	(0, 5)	1.18 (0.91, 49)	(0, 8)	0.89 (1.05, 82)	(0, 6)	1.04 (0.98, 149)
6 Month	(0, 2)	0.62 (0.65, 13)	(0, 2)	0.21 (0.47, 38)	(0, 9)	0.64 (1.27, 70)

TABLE 43 - SUMMARY OF SAFETY THREATS BY EVENT FOR REGION 4 – AS TREATED GROUPS

	ISS Not Workable		ISS Received		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(0, 2)	0.85 (0.49, 54)	(0, 2)	1.03 (0.31, 31)	(0, 3)	1.13 (0.49, 350)
Stepdown	(1, 1)	1.00 (NA, 1)	(0, 1)	0.50 (0.55, 6)	(0, 4)	1.00 (0.80, 32)
6 Month	NA	NA	(0, 0)	0.00 (0.00, 2)	(0, 1)	0.60 (0.52, 10)

TABLE 44 - SUMMARY OF SAFETY THREATS BY EVENT FOR REGION 5 – AS TREATED GROUPS

	ISS Not Workable		ISS Received		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(0, 4)	0.89 (0.63, 201)	(1, 3)	1.16 (0.40, 94)	(0, 4)	1.00 (0.54, 486)
Stepdown	(0, 3)	1.27 (1.10, 11)	(0, 2)	0.62 (0.73, 42)	(0, 4)	0.88 (0.76, 81)
6 Month	NA	NA	(0, 1)	0.15 (0.38, 13)	(0, 3)	0.44 (0.64, 39)

TABLE 45 - SUMMARY OF SAFETY THREATS BY EVENT FOR REGIONS 1, 2, 3, 4, AND 5 – AS TREATED GROUPS

	ISS Not Workable		ISS Received		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(0, 4)	1.10 (0.63, 1024)	(0, 3)	1.18 (0.47, 551)	(0, 5)	1.14 (0.58, 2837)
Stepdown	(0, 5)	1.30 (0.92, 102)	(0, 8)	0.79 (0.89, 176)	(0, 6)	1.03 (0.82, 459)
6 Month	(0, 2)	0.68 (0.67, 19)	(0, 2)	0.20 (0.47, 66)	(0, 9)	0.64 (1.03, 187)

TABLE 46-TABLE 51 present means and standard deviations for the PREMISS scores in Regions 1, 2, 3, 4, and 5 by event for the As Treated cohort (AT).

TABLE 46 - SUMMARY OF PREMISS SCORE BY EVENT AND TREATMENT GROUPS (AS TREATED) FOR REGION 1

	ISS Not Workable		ISS Received		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(-2.57, 0.70)	-1.49 (0.74, 205)	(-2.56, 0.71)	-1.37 (0.71, 89)	(-3.17, 1.10)	-1.60 (0.70, 818)
Stepdown	(-2.46, -0.63)	-1.61 (0.63, 18)	(-2.56, 0.41)	-1.43 (0.76, 27)	(-2.56, 1.10)	-1.58 (0.74, 131)
6 Month	(-1.11, -0.97)	-1.04 (0.10, 2)	(-2.33, 0.41)	-1.10 (0.97, 11)	(-2.55, 1.10)	-1.64 (0.81, 38)

TABLE 47 - SUMMARY OF PREMISS SCORE BY EVENT AND TREATMENT GROUPS (AS TREATED) FOR REGION 2

	ISS Not Workable		ISS Received		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(-2.33, 0.45)	-1.33 (0.57, 132)	(-3.15, 0.73)	-1.21 (0.63, 89)	(-2.39, 0.78)	-1.26 (0.61, 415)
Stepdown	(-1.99, -0.36)	-1.43 (0.45, 23)	(-3.15, 0.73)	-1.52 (0.72, 19)	(-2.39, 0.78)	-1.14 (0.73, 66)
6 Month	(-1.82, -0.36)	-1.35 (0.67, 4)	(-1.55, -1.54)	-1.55 (0.00, 2)	(-2.39, 0.19)	-1.17 (0.75, 30)

TABLE 48 - SUMMARY OF PREMISS SCORE BY EVENT AND TREATMENT GROUPS (AS TREATED) FOR REGION 3

	ISS Not Workable		ISS Received		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(-2.81, 1.23)	-1.44 (0.70, 434)	(-2.77, 0.64)	-1.33 (0.74, 253)	(-3.00, 1.08)	-1.47 (0.70, 786)
Stepdown	(-2.47, 1.23)	-1.40 (0.81, 49)	(-2.64, 0.64)	-1.22 (0.72, 82)	(-2.90, 0.63)	-1.41 (0.75, 149)
6 Month	(-2.44, -0.04)	-1.37 (0.84, 13)	(-2.64, 0.64)	-1.25 (0.83, 38)	(-2.47, 0.63)	-1.31 (0.74, 70)

TABLE 49 - SUMMARY OF PREMISS SCORE BY EVENT AND TREATMENT GROUPS (AS TREATED) FOR REGION 4

	ISS Not Workable		ISS Received		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(-2.21, 0.04)	-1.44 (0.60, 54)	(-2.20, -0.52)	-1.52 (0.48, 31)	(-3.10, 0.49)	-1.49 (0.58, 350)
Stepdown	(-2.05, -2.05)	-2.05 (NA, 1)	(-2.09, -0.69)	-1.49 (0.63, 6)	(-2.21, 0.11)	-1.51 (0.55, 32)
6 Month	NA	NA	(-1.86, -1.64)	-1.75 (0.16, 2)	(-2.21, -0.80)	-1.63 (0.48, 10)

TABLE 50 - SUMMARY OF PREMISS SCORE BY EVENT AND TREATMENT GROUPS (AS TREATED) FOR REGION 5

	ISS Not Workable		ISS Received		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(-1.70, 0.80)	-0.88 (0.57, 209)	(-1.95, 0.82)	-0.82 (0.64, 98)	(-2.21, 1.44)	-0.95 (0.58, 495)
Stepdown	(-1.65, -0.27)	-1.22 (0.47, 11)	(-1.61, 0.82)	-0.76 (0.68, 42)	(-1.91, 0.59)	-1.02 (0.52, 81)
6 Month	NA	NA	(-1.61, 0.52)	-0.79 (0.71, 13)	(-1.91, -0.11)	-1.10 (0.45, 39)

TABLE 51 - SUMMARY OF PREMISS SCORE BY EVENT AND TREATMENT GROUPS (AS TREATED) FOR REGION 1, 2, 3, 4, AND 5

	ISS Not Workable		ISS Received		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(-2.81, 1.23)	-1.32 (0.70, 1034)	(-3.15, 0.82)	-1.24 (0.72, 560)	(-3.17, 1.44)	-1.39 (0.69, 2864)
Stepdown	(-2.47, 1.23)	-1.43 (0.68, 102)	(-3.15, 0.82)	-1.18 (0.75, 176)	(-2.90, 1.10)	-1.36 (0.72, 459)
6 Month	(-2.44, -0.04)	-1.33 (0.75, 19)	(-2.64, 0.64)	-1.16 (0.82, 66)	(-2.55, 1.10)	-1.33 (0.72, 187)

Analysis Methods

ANCOVA models were run predicting follow up time point's counts of safety threat. This yielded both a stepdown and six-month model that controlled for counts of referral safety threats and the PREMISS score, while measuring ITT and TOT group differences.

ITT Findings

No significant differences were observed when comparing ISS assigned to SAU at the stepdown measurement. However at the six-month measurement significant differences were observed for region 1 (p = 0.009; TABLE 52) suggesting lower observed safety threats in the ISS cohort (TABLE 53). TABLE 54-TABLE 61 show results for the other regions.

For all regions combined, significant differences were observed only at the six-month observation period (p=0.004;Table 62); this effect was driven by a .3 difference in observed safety threats suggesting lower number of observed safety threats in the ISS assigned cohort (Table 63).

Results using instrumental variable TOT analysis are presented in TABLE 64 and TABLE 65. The first analysis shows the effect of ISS across all observations at stepdown, the second shows results of ISS after excluding participants who had court intervention after ISS was assigned. These results indicate an intended direction in the PREMISS score, higher scores related to greater safety threats across all cohorts, and of the group effects suggesting more observed safety threats in the SAU cohort however, the group effect was not significant.

TABLE 52 - SUMMARY RESULTS - ANCOVA ITT ANALYSIS ON THE NUMBER OF SAFETY THREATS FOR REGION 1

Time point	Variable	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Stepdown	Intake Dosage	1	8.813	8.813	17.056	0.000
Stepdown	Group	1	0.031	0.031	0.060	0.807
Stepdown	PREMISS Score	1	0.002	0.002	0.004	0.947
Stepdown	Residuals	172	88.875	0.517	NA	NA
Six month	Intake Dosage	1	0.582	0.582	1.067	0.307
Six month	Group	1	3.188	3.188	5.842	0.020
Six month	PREMISS Score	1	1.094	1.094	2.005	0.163
Six month	Residuals	47	25.646	0.546	NA	NA

TABLE 53 - MODEL PREDICTED MEANS AND STANDARD ERRORS FOR SAFETY THREATS FOR REGION 1

	ISS	SAU
Stepdown	1.11 (0.11)	1.08 (0.06)
6 Months	0.05 (0.21)	0.75 (0.12)

TABLE 54 - SUMMARY RESULTS - ANCOVA ITT ANALYSIS ON THE NUMBER OF SAFETY THREATS FOR REGION 2

Time point	Variable	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Stepdown	Intake Dosage	1	10.952	10.952	22.622	0.000
Stepdown	Group	1	0.091	0.091	0.188	0.665
Stepdown	PREMISS Score	1	0.041	0.041	0.085	0.771
Stepdown	Residuals	104	50.350	0.484	NA	NA
Six month	Intake Dosage	1	3.337	3.337	2.620	0.115
Six month	Group	1	0.004	0.004	0.003	0.955
Six month	PREMISS Score	1	0.199	0.199	0.156	0.695
Six month	Residuals	32	40.765	1.274	NA	NA

TABLE 55 - MODEL PREDICTED MEANS AND STANDARD ERRORS FOR SAFETY THREATS FOR REGION 2

	ISS	SAU
Stepdown	1.09 (0.11)	1.14 (0.09)
6 Months	0.91 (0.47)	0.85 (0.21)

TABLE 56 - SUMMARY RESULTS - ANCOVA ITT ANALYSIS ON THE NUMBER OF SAFETY THREATS FOR REGION 3

Time point	Variable	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Stepdown	Intake Dosage	1	9.384	9.384	9.835	0.002
Stepdown	Group	1	0.129	0.129	0.135	0.714
Stepdown	PREMISS Score	1	1.017	1.017	1.066	0.303
Stepdown	Residuals	275	262.381	0.954	NA	NA
Six month	Intake Dosage	1	0.137	0.137	0.126	0.723
Six month	Group	1	3.132	3.132	2.894	0.092
Six month	PREMISS Score	1	0.334	0.334	0.308	0.580
Six month	Residuals	117	126.645	1.082	NA	NA

TABLE 57 - MODEL PREDICTED MEANS AND STANDARD ERRORS FOR SAFETY THREATS FOR REGION 3

	ISS	SAU
Stepdown	1 (0.09)	1.03 (0.08)
6 Months	0.32 (0.15)	0.64 (0.12)

TABLE 58 - SUMMARY RESULTS - ANCOVA ITT ANALYSIS ON THE NUMBER OF SAFETY THREATS FOR REGION 4

Time point	Variable	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Stepdown	Intake Dosage	1	6.410	6.410	14.227	0.001
Stepdown	Group	1	0.434	0.434	0.962	0.333
Stepdown	PREMISS Score	1	0.155	0.155	0.344	0.561
Stepdown	Residuals	35	15.770	0.451	NA	NA
Six month	Group	1	0.600	0.600	2.461	0.151
Six month	PREMISS Score	1	0.206	0.206	0.843	0.382
Six month	Residuals	9	2.194	0.244	NA	NA

TABLE 59 - MODEL PREDICTED MEANS AND STANDARD ERRORS FOR SAFETY THREATS FOR REGION 4

	ISS	SAU
Stepdown	0.7 (0.26)	0.97 (0.12)
6 Months	0.03 (0.35)	0.59 (0.16)

TABLE 60 - SUMMARY RESULTS - ANCOVA ITT ANALYSIS ON THE NUMBER OF SAFETY THREATS FOR REGION 5

Time point	Variable	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Stepdown	Intake Dosage	1	0.010	0.010	0.015	0.902
Stepdown	Group	1	0.471	0.471	0.734	0.393
Stepdown	PREMISS Score	1	1.131	1.131	1.762	0.187
Stepdown	Residuals	130	83.440	0.642	NA	NA
Six month	Intake Dosage	1	0.020	0.020	0.055	0.816
Six month	Group	1	0.836	0.836	2.308	0.135
Six month	PREMISS Score	1	0.041	0.041	0.113	0.739
Six month	Residuals	47	17.025	0.362	NA	NA

TABLE 61 - MODEL PREDICTED MEANS AND STANDARD ERRORS FOR SAFETY THREATS FOR REGION 5

	ISS	SAU
Stepdown	0.77 (0.11)	0.87 (0.09)
6 Months	0.14 (0.17)	0.45 (0.1)

TABLE 62 - SUMMARY RESULTS - ANCOVA ITT ANALYSIS ON THE NUMBER OF SAFETY THREATS STATEWIDE

Time point	Variable	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Stepdown	Intake Dosage	1	29.801	29.801	42.363	0.000
Stepdown	Group	1	0.503	0.503	0.714	0.398
Stepdown	PREMISS Score	1	1.697	1.697	2.412	0.121
Stepdown	Residuals	732	514.933	0.703	NA	NA
Six month	Intake Dosage	1	0.316	0.316	0.376	0.540
Six month	Group	1	6.901	6.901	8.221	0.004
Six month	PREMISS Score	1	0.002	0.002	0.002	0.965
Six month	Residuals	267	224.125	0.839	NA	NA

TABLE 63 - MODEL PREDICTED MEANS AND STANDARD ERRORS FOR SAFETY THREATS STATEWIDE

	ISS	SAU
Stepdown	0.98 (0.05)	1.03 (0.04)
6 Months	0.3 (0.1)	0.65 (0.07)

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	0.6277184	0.0361080	17.384476	0.0000000
SAU Received	0.0364942	0.0209840	1.739141	0.0822125
Referral Score	-0.0586165	0.0141506	-4.142343	0.0000363
AOCS Count Safety Threats Referral	0.0594127	0.0234482	2.533784	0.0113836

TABLE 65 - RESULTS USING **SUBGROUP** ANALYSIS EXCLUDING COURT INTERVENTION CASES

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	0.5552781	0.0913732	6.0770330	0.0000000
SAU Received	0.0697189	0.0584746	1.1922942	0.2340977
Referral Score	-0.0751373	0.0354375	-2.1202800	0.0348113
AOCS Count Safety Threats Referral	0.0622003	0.0626186	0.9933205	0.3213624

FIGURE 14- BOXPLOT OF SAFETY THREATS BY MEASUREMENT POINT BY REGION AS-TREATED GROUPS
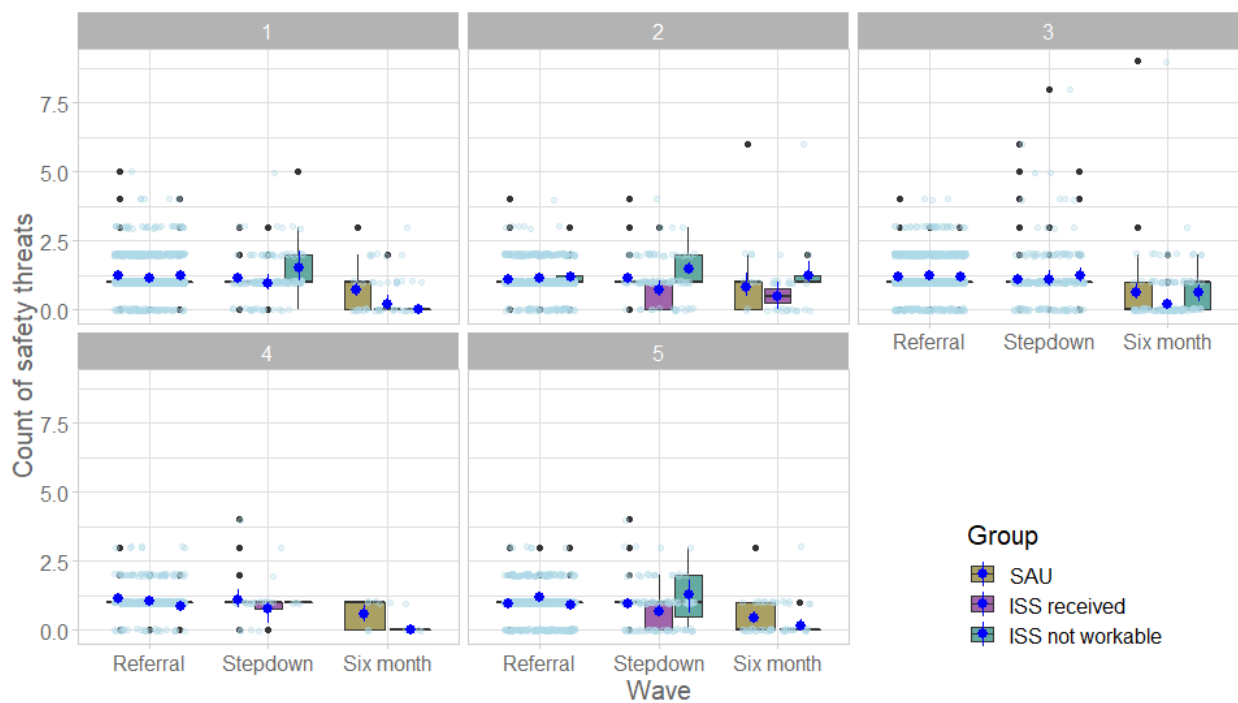
As-Treated Findings

Significant group differences were observed at stepdown for regions: 1 (p = 0.011; TABLE 66), 2 (p = 0.003; TABLE 68), and 5 (p = 0.035; TABLE 70). Through all of these the ISS received group displayed the lowest number of estimated safety threats. No significant group differences were observed within the

regional data at the six-month observation period. However, the direction of the effects remained extremely consistent with the ISS received group displaying the lowest number of estimated safety threats when compared to the SAU group whereas the ISS not workable showed variable results.

For all regions combined there were significant group differences observed at the stepdown observation ($p < 0.001$; TABLE 76). This effect was driven by the differences observed between the ISS received and SAU groups, with the ISS received group displaying an estimated .2 reduction in safety threats and the ISS not received group displaying an estimated increase of .3 safety threats when compared to the SAU group (TABLE 77). This significant group difference was also observed at the six-month observation time point ($p = 0.002$; TABLE 76). The ISS revived group displayed the lowest estimated number of safety threats whereas both SAU and ISS not received displayed relatively similar estimated numbers of safety threats (TABLE 77).

TABLE 66 - SUMMARY RESULTS - ANCOVA AS-TREATED ANALYSIS ON THE NUMBER OF SAFETY THREATS FOR REGION 1

Time point	Variable	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Stepdown	Intake Dosage	1	8.813	8.813	17.874	0.000
Stepdown	Group	2	4.581	2.291	4.646	0.011
Stepdown	PREMISS Score	1	0.013	0.013	0.027	0.869
Stepdown	Residuals	171	84.314	0.493	NA	NA
Six month	Intake Dosage	1	0.582	0.582	1.044	0.312
Six month	Group	2	3.201	1.601	2.872	0.067
Six month	PREMISS Score	1	1.085	1.085	1.947	0.170
Six month	Residuals	46	25.641	0.557	NA	NA

TABLE 67 - MODEL PREDICTED MEANS AND STANDARD ERRORS FOR SAFETY THREATS FOR REGION 1

	ISS not workable	ISS received	SAU
Stepdown	1.5 (0.17)	0.85 (0.14)	1.08 (0.06)
6 Months	0.01 (0.53)	0.06 (0.24)	0.74 (0.12)

TABLE 68 - SUMMARY RESULTS – ANCOVA AT ANALYSIS ON THE NUMBER OF SAFETY THREATS FOR REGION 2

Time point	Variable	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Stepdown	Intake Dosage	1	10.952	10.952	25.009	0.000
Stepdown	Group	2	5.365	2.682	6.125	0.003
Stepdown	PREMISS Score	1	0.010	0.010	0.023	0.879
Stepdown	Residuals	103	45.108	0.438	NA	NA
Six month	Intake Dosage	1	3.337	3.337	2.608	0.116
Six month	Group	2	1.151	0.575	0.450	0.642
Six month	PREMISS Score	1	0.150	0.150	0.118	0.734
Six month	Residuals	31	39.667	1.280	NA	NA

TABLE 69 - MODEL PREDICTED MEANS AND STANDARD ERRORS FOR SAFETY THREATS FOR REGION 2

	ISS not workable	ISS received	SAU
Stepdown	1.41 (0.14)	0.7 (0.15)	1.14 (0.08)
6 Months	1.2 (0.57)	0.29 (0.82)	0.85 (0.21)

TABLE 70 - SUMMARY RESULTS - ANCVOA AT ANALYSIS ON THE NUMBER OF SAFETY THREATS FOR REGION 3

Time point	Variable	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Stepdown	Intake Dosage	1	9.384	9.384	9.875	0.002
Stepdown	Group	2	2.355	1.178	1.239	0.291
Stepdown	PREMISS Score	1	0.795	0.795	0.837	0.361
Stepdown	Residuals	274	260.376	0.950	NA	NA
Six month	Intake Dosage	1	0.137	0.137	0.127	0.723
Six month	Group	2	4.728	2.364	2.192	0.116
Six month	PREMISS Score	1	0.271	0.271	0.251	0.617
Six month	Residuals	116	125.113	1.079	NA	NA

TABLE 71 - MODEL PREDICTED MEANS AND STANDARD ERRORS FOR SAFETY THREATS FOR REGION 3

	ISS not workable	ISS received	SAU
Stepdown	1.16 (0.14)	0.9 (0.11)	1.03 (0.08)
6 Months	0.61 (0.29)	0.22 (0.17)	0.64 (0.12)

TABLE 72 - SUMMARY RESULTS - ANCOVA AT ANALYSIS ON THE NUMBER OF SAFETY THREATS FOR REGION 4

Time point	Variable	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Stepdown	Intake Dosage	1	6.410	6.410	14.069	0.001
Stepdown	Group	2	0.648	0.324	0.711	0.498
Stepdown	PREMISS Score	1	0.220	0.220	0.483	0.492
Stepdown	Residuals	34	15.491	0.456	NA	NA
Six month	Group	1	0.600	0.600	2.461	0.151
Six month	PREMISS Score	1	0.206	0.206	0.843	0.382
Six month	Residuals	9	2.194	0.244	NA	NA
Stepdown	Intake Dosage	1	6.410	6.410	14.069	0.001

TABLE 73 - MODEL PREDICTED MEANS AND STANDARD ERRORS FOR SAFETY THREATS FOR REGION 4

	ISS not workable	ISS received	SAU
Stepdown	NA	0.62 (0.28)	0.97 (0.12)
6 Months	NA	0.03 (0.35)	0.59 (0.16)

TABLE 74 - SUMMARY RESULTS - ANCOVA AT ANALYSIS ON THE NUMBER OF SAFETY THREATS FOR REGION 5

Time point	Variable	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Stepdown	Intake Dosage	1	0.010	0.010	0.016	0.900
Stepdown	Group	2	4.277	2.138	3.433	0.035
Stepdown	PREMISS Score	1	0.418	0.418	0.671	0.414
Stepdown	Residuals	129	80.348	0.623	NA	NA
Six month	Intake Dosage	1	0.020	0.020	0.055	0.816
Six month	Group	1	0.836	0.836	2.308	0.135
Six month	PREMISS Score	1	0.041	0.041	0.113	0.739
Six month	Residuals	47	17.025	0.362	NA	NA

TABLE 75 - MODEL PREDICTED MEANS AND STANDARD ERRORS FOR SAFETY THREATS FOR REGION 5

	ISS not workable	ISS received	SAU
Stepdown	NA	0.64 (0.12)	0.87 (0.09)
6 Months	NA	0.14 (0.17)	0.45 (0.1)

TABLE 76 - SUMMARY RESULTS - ANCOVA AT ANALYSIS ON THE NUMBER OF SAFETY THREATS STATEWIDE

Time point	Variable	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Stepdown	Intake Dosage	1	29.801	29.801	43.630	0.000
Stepdown	Group	2	17.025	8.513	12.463	0.000
Stepdown	PREMISS Score	1	0.811	0.811	1.187	0.276
Stepdown	Residuals	731	499.297	0.683	NA	NA
Six month	Intake Dosage	1	0.316	0.316	0.381	0.538
Six month	Group	2	10.445	5.223	6.298	0.002
Six month	PREMISS Score	1	0.004	0.004	0.005	0.942
Six month	Residuals	266	220.578	0.829	NA	NA

TABLE 77 - MODEL PREDICTED MEANS, STD. ERR. FOR SAFETY THREATS STATEWIDE

	ISS not workable	ISS received	SAU
Stepdown	1.29 (0.08)	0.8 (0.06)	1.03 (0.04)
6 Months	0.68 (0.21)	0.19 (0.11)	0.65 (0.07)

While base rates of observed safety threats were low at the referral timepoint (modal value = 1) a significant reduction was observed for all cohorts. Statewide aggregate effects display reductions in counts in safety threats across all cohorts. However, the greatest six-month reduction was observed in the ITT ISS cohort and the AT ISS received cohort.

Question 2: Relative to the SAU condition, does ISS increase the number of protective capacities identified in post-intervention AOCS administrations?

The key takeaways for this question include:

- The greatest predictor in future counts of protective capacities i.e. protective capacities at stepdown and the six-month observations, was the number of protective capacities at the referral AOCS
- The ISS cohort displayed the greatest values in protective capacities at the stepdown and six-month observation periods.

Assessment of Child Safety (AOCS) forms are collected by FCS and/or Permanency Planning at baseline, stepdown, and at the 6 months for both ISS and SAU families. AOCS forms track the number of protective capacities for families in both ISS and SAU at each measurement. The three most common protective capacity items across the entire cohort reported at baseline include: 1) PRFC shows support/concern for child(ren)'s health, safety, and well-being (n=2521), 2) PRFC demonstrates a positive perception of child(ren) and has appropriate expectations based upon each child's development (n=2170), and 3) PRFC demonstrates an ability to meet child(ren)'s basic and emotional needs (n=2060). This section of the report provides an accounting for the summation of all 11 protective capacities.

Methods

Analyses of covariance (ANCOVAs) were used to assess group differences in safety threats at each follow up time point while accounting for baseline safety threat counts. The ANCOVA approach yields an interpretable metric which assess group differences at each time point while controlling for nuisance variables, while this approach is easy to implement and interpret, it requires the exclusion of participants without observations at both the intake and succeeding time points. Accordingly, the SAU group had 458 cases at the stepdown time point, and 187 at the six-month time point across all regions, whereas the ISS group had 278 cases at the stepdown time point, and 85 at the 6-month time point across all regions.

Descriptive statistics for protective capacity counts and sample sizes are presented individually by region (regions 1, 2, 3, 4, and 5), as well as in aggregate on TABLE 78 through TABLE 83 using ITT groupings (Group Assignment: ISS vs. SAU).

TABLE 78 - SUMMARY OF PROTECTIVE CAPACITIES OVER TIME FOR REGION 1 – ITT GROUPS

	ISS		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(0, 11)	3.11 (3.54, 294)	(0, 11)	3.16 (3.56, 817)
Stepdown	(0, 11)	3.56 (3.82, 45)	(0, 11)	2.81 (3.28, 131)
6 Month	(0, 11)	6.77 (4.92, 13)	(0, 11)	4.29 (4.01, 38)

TABLE 79 - SUMMARY OF PROTECTIVE CAPACITIES OVER TIME FOR REGION 2– ITT GROUPS

	ISS		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(0, 11)	4.38 (3.52, 221)	(0, 11)	4.46 (3.73, 415)
Stepdown	(0, 11)	4.38 (3.73, 42)	(0, 11)	4.14 (3.18, 66)
6 Month	(1, 11)	9.33 (4.08, 6)	(0, 11)	4.67 (4.14, 30)

TABLE 80 - SUMMARY OF PROTECTIVE CAPACITIES OVER TIME FOR REGION 3– ITT GROUPS

	ISS		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(0, 11)	3.53 (3.31, 680)	(0, 11)	3.66 (3.49, 769)
Stepdown	(0, 11)	4.40 (3.92, 131)	(0, 11)	3.90 (3.94, 148)
6 Month	(0, 11)	6.61 (4.13, 51)	(0, 11)	6.10 (4.58, 70)

TABLE 81 - SUMMARY OF PROTECTIVE CAPACITIES OVER TIME FOR REGION 4– ITT GROUPS

	ISS		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(0, 11)	4.07 (3.52, 85)	(0, 11)	3.61 (3.68, 350)
Stepdown	(2, 11)	7.71 (3.50, 7)	(0, 11)	3.75 (3.83, 32)
6 Month	(11, 11)	11.00 (0.00, 2)	(0, 11)	5.60 (5.17, 10)

TABLE 82 - SUMMARY OF PROTECTIVE CAPACITIES OVER TIME FOR REGION 5– ITT GROUPS

	ISS		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(0, 11)	4.71 (3.90, 295)	(0, 11)	4.01 (3.78, 486)
Stepdown	(0, 11)	6.62 (4.53, 53)	(0, 11)	4.31 (4.34, 81)
6 Month	(0, 11)	8.08 (4.68, 13)	(0, 11)	6.10 (5.02, 39)

TABLE 83 - SUMMARY OF PROTECTIVE CAPACITIES OVER TIME FOR REGIONS 1, 2, 3, 4, AND 5 – ITT GROUPS

	ISS		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(0, 11)	3.82 (3.55, 1575)	(0, 11)	3.69 (3.64, 2837)
Stepdown	(0, 11)	4.77 (4.11, 278)	(0, 11)	3.68 (3.76, 458)
6 Month	(0, 11)	7.15 (4.33, 85)	(0, 11)	5.48 (4.55, 187)

Descriptive statistics for protective capacity counts and sample sizes are presented individually by region (regions 1, 2, 3, 4, and 5), as well as in aggregate in **TABLE 84**-**TABLE 89** using AT groupings.

TABLE 84 - SUMMARY OF PROTECTIVE CAPACITIES OVER TIME FOR REGION 1 – AS TREATED GROUPS

	ISS Not Workable		ISS Received		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(0, 11)	2.84 (3.51, 205)	(0, 11)	3.74 (3.52, 89)	(0, 11)	3.16 (3.56, 817)
Stepdown	(0, 11)	1.94 (3.06, 18)	(0, 11)	4.63 (3.94, 27)	(0, 11)	2.81 (3.28, 131)
6 Month	(0, 11)	5.50 (7.78, 2)	(0, 11)	7.00 (4.75, 11)	(0, 11)	4.29 (4.01, 38)

TABLE 85 - SUMMARY OF PROTECTIVE CAPACITIES OVER TIME FOR REGION 2 - AS TREATED GROUPS

	ISS Not Workable		ISS Received		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(0, 11)	4.08 (3.53, 132)	(0, 11)	4.83 (3.47, 89)	(0, 11)	4.46 (3.73, 415)
Stepdown	(0, 7)	2.39 (2.39, 23)	(2, 11)	6.79 (3.69, 19)	(0, 11)	4.14 (3.18, 66)
6 Month	(1, 11)	8.50 (5.00, 4)	(11, 11)	11.00 (0.00, 2)	(0, 11)	4.67 (4.14, 30)

TABLE 86 - SUMMARY OF PROTECTIVE CAPACITIES OVER TIME FOR REGION 3- AS TREATED GROUPS

	ISS Not Workable		ISS Received		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(0, 11)	3.03 (3.22, 432)	(0, 11)	4.41 (3.30, 248)	(0, 11)	3.66 (3.49, 769)
Stepdown	(0, 11)	2.31 (2.63, 49)	(0, 11)	5.65 (4.03, 82)	(0, 11)	3.90 (3.94, 148)
6 Month	(0, 11)	3.38 (3.88, 13)	(0, 11)	7.71 (3.65, 38)	(0, 11)	6.10 (4.58, 70)

TABLE 87 - SUMMARY OF PROTECTIVE CAPACITIES OVER TIME FOR REGION 4- AS TREATED GROUPS

	ISS Not Workable		ISS Received		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(0, 11)	4.17 (3.89, 54)	(0, 11)	3.90 (2.81, 31)	(0, 11)	3.61 (3.68, 350)
Stepdown	(2, 2)	2.00 (NA, 1)	(5, 11)	8.67 (2.66, 6)	(0, 11)	3.75 (3.83, 32)
6 Month	NA	NA	(11, 11)	11.00 (0.00, 2)	(0, 11)	5.60 (5.17, 10)

TABLE 88 - SUMMARY OF PROTECTIVE CAPACITIES OVER TIME FOR REGION 5- AS TREATED GROUPS

	ISS Not Workable		ISS Received		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(0, 11)	4.52 (4.06, 201)	(0, 11)	5.13 (3.53, 94)	(0, 11)	4.01 (3.78, 486)
Stepdown	(0, 11)	3.09 (4.21, 11)	(0, 11)	7.55 (4.19, 42)	(0, 11)	4.31 (4.34, 81)
6 Month	NA	NA	(0, 11)	8.08 (4.68, 13)	(0, 11)	6.10 (5.02, 39)

	ISS Not Workable		ISS Received		SAU	
	Range	Mean (SD, n)	Range	Mean (SD, n)	Range	Mean (SD, n)
Referral	(0, 11)	3.48 (3.59, 1024)	(0, 11)	4.46 (3.40, 551)	(0, 11)	3.69 (3.64, 2837)
Stepdown	(0, 11)	2.34 (2.82, 102)	(0, 11)	6.17 (4.09, 176)	(0, 11)	3.68 (3.76, 458)
6 Month	(0, 11)	4.68 (4.70, 19)	(0, 11)	7.86 (3.98, 66)	(0, 11)	5.48 (4.55, 187)

Analysis Methods

Differences between groups were tested using analyses of covariance (ANCOVAs). Group differences were tested while controlling for the number of observed protective capacities at referral. In this reporting period, the child's PREMISS score is included as an additional covariate. This score is a crucial component in determining family eligibility for ISS based on a number of safety factors known to be predictive of successful reunification. Therefore, adjusting for this risk score is appropriate for estimation of treatment effects for this outcome.

ITT Findings

Figure 17 displays boxplots for the number of protective capacities for these two conditions over time.

Significant group differences were observed when comparing ISS assigned to SAU at the stepdown measurement for regions 4 and 5 indicating a greater number of observed protective capacities in the ISS assigned group ($p < 0.001$ TABLE 96; $p < 0.001$, TABLE 98).

Significant group differences were only observed in region 3 ($P = 0.008$; TABLE 94) with the ISS group displaying more than half of an increase in predicted protective capacities (TABLE 95).

For all regions combined, significant differences were observed at both the stepdown measurement ($p < 0.001$; TABLE 100) and the 6 month measurement period ($p = 0.001$; TABLE 100). The ISS cohort displayed an estimated one additional protective capacity at the stepdown assessment, and one and a half more protective capacities at the 6 month follow-up assessment when compared to the SAU group (TABLE 101).

Results using instrumental variable TOT analysis are presented in TABLE 102 and TABLE 103. These results indicate an intended direction for the PREMISS score effect indicating more negative values in this score yield lower number of protective capacities, and suggest a significant group received effect across the entire cohort. However, when cases who require court intervention are removed, the group assignment effect in this ITT subgroup model is not significant.

TABLE 90 - SUMMARY RESULTS - ANCOVA ITT ANALYSIS ON PROTECTIVE CAPACITIES COUNT FOR REGION 1

Time point	Variable	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Stepdown	Intake Dosage	1	51.172	51.172	4.513	0.035
Stepdown	Group	1	17.990	17.990	1.587	0.210
Stepdown	PREMISS Score	1	36.583	36.583	3.226	0.074
Stepdown	Residuals	172	1950.255	11.339	NA	NA
Six month	Intake Dosage	1	10.189	10.189	0.544	0.465
Six month	Group	1	52.603	52.603	2.807	0.101
Six month	PREMISS Score	1	0.006	0.006	0.000	0.986
Six month	Residuals	47	880.888	18.742	NA	NA

TABLE 91 – ITT MODEL PREDICTED MEANS AND STANDARD ERRORS FOR PROTECTIVE CAPACITIES FOR REGION 1

	ISS	SAU
Stepdown	3.51 (0.5)	2.83 (0.29)
6 Months	6.68 (1.24)	4.32 (0.71)

TABLE 92 - SUMMARY RESULTS - ANCOVA ITT ANALYSIS ON PROTECTIVE CAPACITIES COUNT FOR REGION 2

Time point	Variable	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Stepdown	Intake Dosage	1	24.301	24.301	2.107	0.150
Stepdown	Group	1	2.697	2.697	0.234	0.630
Stepdown	PREMISS Score	1	2.548	2.548	0.221	0.639
Stepdown	Residuals	104	1199.668	11.535	NA	NA
Six month	Intake Dosage	1	0.598	0.598	0.033	0.857
Six month	Group	1	108.889	108.889	6.014	0.020
Six month	PREMISS Score	1	0.031	0.031	0.002	0.967
Six month	Residuals	32	579.370	18.105	NA	NA

TABLE 93 – ITT MODEL PREDICTED MEANS AND STANDARD ERRORS FOR PROTECTIVE CAPACITIES FOR REGION 2

	ISS	SAU
Stepdown	4.39 (0.53)	4.13 (0.42)
6 Months	9.32 (1.75)	4.67 (0.78)

TABLE 94 - SUMMARY RESULTS - ANCOVA ITT ANALYSIS ON PROTECTIVE CAPACITIES FOR REGION 3

Time point	Variable	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Stepdown	Intake Dosage	1	335.849	335.849	23.683	0.000
Stepdown	Group	1	24.672	24.672	1.740	0.188
Stepdown	PREMISS Score	1	33.723	33.723	2.378	0.124
Stepdown	Residuals	275	3899.849	14.181	NA	NA
Six month	Intake Dosage	1	105.790	105.790	5.719	0.018
Six month	Group	1	9.870	9.870	0.534	0.467
Six month	PREMISS Score	1	26.076	26.076	1.410	0.238
Six month	Residuals	117	2164.330	18.499	NA	NA

TABLE 95 – ITT MODEL PREDICTED MEANS AND STANDARD ERRORS FOR PROTECTIVE CAPACITIES FOR REGION 3

	ISS	SAU
Stepdown	4.42 (0.33)	3.88 (0.31)
6 Months	6.63 (0.6)	6.08 (0.51)

TABLE 96 - SUMMARY RESULTS - ANCOVA ITT ANALYSIS ON PROTECTIVE CAPACITIES FOR REGION 4

Time point	Variable	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Stepdown	Intake Dosage	1	37.560	37.560	2.790	0.104
Stepdown	Group	1	103.653	103.653	7.699	0.009
Stepdown	PREMISS Score	1	5.296	5.296	0.393	0.535
Stepdown	Residuals	35	471.183	13.462	NA	NA
Six month	Intake Dosage	1	3.648	3.648	0.247	0.632
Six month	Group	1	45.352	45.352	3.072	0.118
Six month	PREMISS Score	1	121.911	121.911	8.259	0.021
Six month	Residuals	8	118.089	14.761	NA	NA

TABLE 97 – ITT MODEL PREDICTED MEANS AND STANDARD ERRORS FOR PROTECTIVE CAPACITIES FOR REGION 4

	ISS	SAU
Stepdown	7.94 (1.39)	3.7 (0.65)
6 Months	11.18 (2.88)	5.56 (1.23)

TABLE 98 - SUMMARY RESULTS – ANCOVA ITT ANALYSIS ON PROTECTIVE CAPACITIES FOR REGION 5

Time point	Variable	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Stepdown	Intake Dosage	1	91.859	91.859	4.827	0.030
Stepdown	Group	1	171.751	171.751	9.025	0.003
Stepdown	PREMISS Score	1	11.652	11.652	0.612	0.435
Stepdown	Residuals	130	2474.022	19.031	NA	NA
Six month	Intake Dosage	1	17.531	17.531	0.698	0.408
Six month	Group	1	31.044	31.044	1.236	0.272
Six month	PREMISS Score	1	9.969	9.969	0.397	0.532
Six month	Residuals	47	1180.201	25.111	NA	NA

TABLE 99 - ITT MODEL PREDICTED MEANS AND STANDARD ERRORS FOR PROTECTIVE CAPACITIES FOR REGION 5

	ISS	SAU
Stepdown	6.57 (0.6)	4.34 (0.49)
6 Months	7.72 (1.46)	6.09 (0.83)

TABLE 100 - SUMMARY RESULTS - ANCOVA ANALYSIS ON THE NUMBER OF PROTECTIVE CAPACITIES STATEWIDE

Time point	Variable	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Stepdown	Intake Dosage	1	588.527	588.527	41.312	0.000
Stepdown	Group	1	219.641	219.641	15.418	0.000
Stepdown	PREMISS Score	1	99.502	99.502	6.985	0.008
Stepdown	Residuals	732	10428.047	14.246	NA	NA
Six month	Intake Dosage	1	128.047	128.047	6.480	0.011
Six month	Group	1	148.060	148.060	7.492	0.007
Six month	PREMISS Score	1	12.405	12.405	0.628	0.429
Six month	Residuals	267	5276.395	19.762	NA	NA

TABLE 101 - MODEL PREDICTED MEANS AND STANDARD ERRORS FOR PROTECTIVE CAPACITIES STATEWIDE

	ISS	SAU
Stepdown	4.76 (0.23)	3.69 (0.18)
6 Months	7.06 (0.48)	5.49 (0.33)
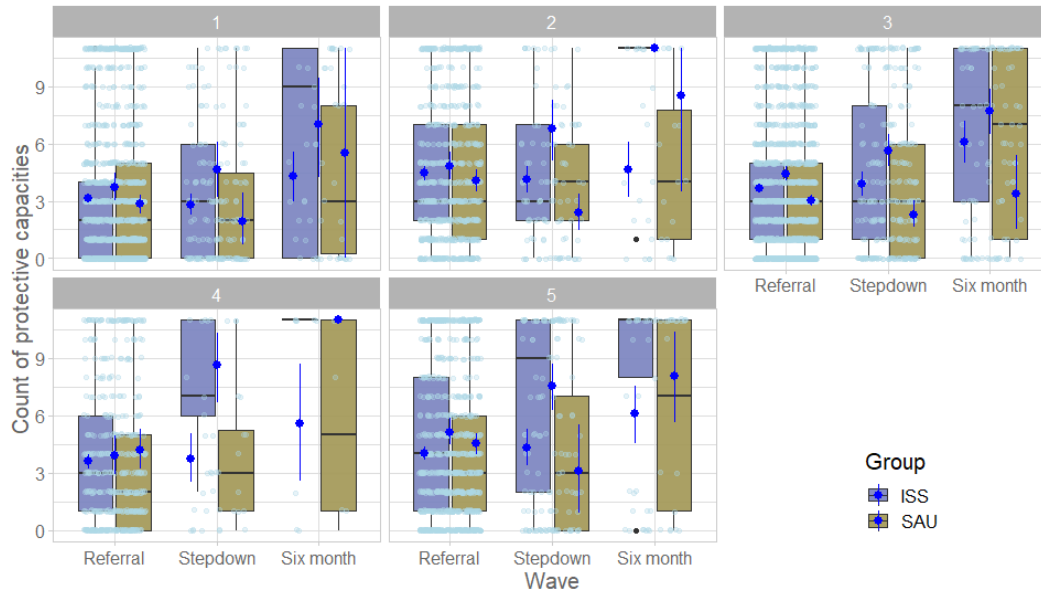
TABLE 102 - RESULTS USING INSTRUMENTAL VARIABLE TOT ANALYSIS ACROSS ENTIRE COHORT

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	0.74	0.033	22.384	0.000
SAU Received	-0.08	0.023	-3.462	0.000
Referral Score	0.014	0.016	0.861	0.389
AOCS Count Safety Threats Referral	0.015	0.003	4.443	0.000

TABLE 103 - RESULTS USING SUBGROUP ANALYSIS EXCLUDING COURT INTERVENTION CASES

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	0.555	0.091	6.077	0.000
SAU Received	0.070	0.058	1.192	0.234
Referral Score	-0.075	0.035	-2.12	0.035
AOCS Count Safety Threats Referral	0.062	0.062	0.993	0.321

As-Treated Findings

Significant group differences were observed at stepdown for all regions: 1 (p = 0.011; TABLE 104), 2 (p < 0.001; TABLE 106), 3 (p<0.001; TABLE 108), 4 (p = 0.011; TABLE 110) and 5 (p < 0.001; TABLE 112). All of these effects displayed identical effects, with the ISS group consistently displaying a greater number of predicted protective capacities at the stepdown observation period than the SAU and ISS not workable group. The greatest effect was shown in region 4 with the ISS received group displaying more than 5 more predicted protective capacities than the SAU group (TABLE 111). For the six-month

observation period significant group differences were observed only region 3 (p = 0.008; TABLE 108) where the ISS received group displayed roughly one and a half more protective capacities than the SAU group, and the ISS not received group displayed roughly 2.3 less protective capacities than the SAU group (TABLE 109)

For all regions combined there were significant group differences observed at the stepdown observation (p < 0.001; TABLE 114). This effect was driven by the differences observed between the ISS received and SAU groups, with the ISS received group displaying an estimated more than 2 greater protective capacities, the ISS not workable group displayed the lowest estimated number of protective capacities with roughly one less than the SAU group (TABLE 115). This significant group difference effect persisted into the six month observation time period (p = 0.001; TABLE 114). Here the ISS received group continued to display the greatest number of protective capacities, with more than 2 greater predicted protective capacities than the SAU group, the ISS not workable group had roughly half of a predicted protective capacity lower than the SAU group (TABLE 115).

TABLE 104 - SUMMARY RESULTS - ANCOVA AT ANALYSIS ON THE NUMBER OF PROTECTIVE CAPACITIES FOR REGION 1

Time point	Variable	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Stepdown	Intake Dosage	1	51.172	51.172	4.655	0.032
Stepdown	Group	2	94.907	47.453	4.317	0.015
Stepdown	PREMISS Score	1	30.337	30.337	2.760	0.098
Stepdown	Residuals	171	1879.585	10.992	NA	NA
Six month	Intake Dosage	1	10.189	10.189	0.534	0.469
Six month	Group	2	55.083	27.542	1.442	0.247
Six month	PREMISS Score	1	0.106	0.106	0.006	0.941
Six month	Residuals	46	878.308	19.094	NA	NA

TABLE 105 – AS TREATED MODEL PREDICTED MEANS AND STD. ERR. FOR PROTECTIVE CAPACITIES FOR REGION 1

	ISS not workable	ISS received	SAU
Stepdown	1.97 (0.78)	4.54 (0.64)	2.82 (0.29)
6 Months	5.6 (3.2)	6.89 (1.37)	4.32 (0.72)

TABLE 106 - SUMMARY RESULTS - GLMM AT ANALYSIS ON PROTECTIVE CAPACITIES FOR REGION 2

Time point	Variable	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Stepdown	Intake Dosage	1	24.301	24.301	2.466	0.119
Stepdown	Group	2	189.886	94.943	9.635	0.000
Stepdown	PREMISS Score	1	0.061	0.061	0.006	0.937
Stepdown	Residuals	103	1014.965	9.854	NA	NA
Six month	Intake Dosage	1	0.598	0.598	0.032	0.858
Six month	Group	2	116.909	58.454	3.171	0.056
Six month	PREMISS Score	1	0.004	0.004	0.000	0.988
Six month	Residuals	31	571.378	18.432	NA	NA

TABLE 107 - AS TREATED MODEL PREDICTED MEANS AND STD. ERR. FOR PROTECTIVE CAPACITIES FOR REGION 2

	ISS not workable	ISS received	SAU
Stepdown	2.48 (0.66)	6.74 (0.73)	4.12 (0.39)
6 Months	8.51 (2.15)	10.98 (3.07)	4.67 (0.79)

TABLE 108 - SUMMARY RESULTS - ANCOVA AT ANALYSIS ON PROTECTIVE CAPACITIES FOR REGION 3

Time point	Variable	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Stepdown	Intake Dosage	1	335.849	335.849	25.085	0.000
Stepdown	Group	2	264.143	132.072	9.865	0.000
Stepdown	PREMISS Score	1	25.663	25.663	1.917	0.167
Stepdown	Residuals	274	3668.438	13.388	NA	NA
Six month	Intake Dosage	1	105.790	105.790	6.122	0.015
Six month	Group	2	173.709	86.854	5.026	0.008
Six month	PREMISS Score	1	22.006	22.006	1.273	0.261
Six month	Residuals	116	2004.561	17.281	NA	NA

TABLE 109 - AS TREATED MODEL PREDICTED MEANS AND STD. ERR. FOR PROTECTIVE CAPACITIES FOR REGION 3

	ISS not workable	ISS received	SAU
Stepdown	2.66 (0.53)	5.46 (0.41)	3.88 (0.3)
6 Months	3.6 (1.16)	7.67 (0.67)	6.08 (0.5)

TABLE 110 - SUMMARY RESULTS - GLMM AT ANALYSIS ON PROTECTIVE CAPACITIES FOR REGION 4

Time point	Variable	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Stepdown	Intake Dosage	1	37.560	37.560	2.915	0.097
Stepdown	Group	2	132.204	66.102	5.129	0.011
Stepdown	PREMISS Score	1	9.764	9.764	0.758	0.390
Stepdown	Residuals	34	438.165	12.887	NA	NA
Six month	Intake Dosage	1	3.648	3.648	0.247	0.632
Six month	Group	1	45.352	45.352	3.072	0.118
Six month	PREMISS Score	1	121.911	121.911	8.259	0.021
Six month	Residuals	8	118.089	14.761	NA	NA

TABLE 111 - AS TREATED MODEL PREDICTED MEANS AND STD. ERR. FOR PROTECTIVE CAPACITIES FOR REGION 4

	ISS not workable	ISS received	SAU
Stepdown	NA	8.8 (1.47)	3.71 (0.64)
6 Months	NA	11.18 (2.88)	5.56 (1.23)

TABLE 112 - SUMMARY RESULTS - GLMM AT ANALYSIS ON PROTECTIVE CAPACITIES FOR REGION 5

Time point	Variable	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Stepdown	Intake Dosage	1	91.859	91.859	5.058	0.026
Stepdown	Group	2	312.392	156.196	8.601	0.000
Stepdown	PREMISS Score	1	2.439	2.439	0.134	0.715
Stepdown	Residuals	129	2342.594	18.160	NA	NA
Six month	Intake Dosage	1	17.531	17.531	0.698	0.408
Six month	Group	1	31.044	31.044	1.236	0.272
Six month	PREMISS Score	1	9.969	9.969	0.397	0.532
Six month	Residuals	47	1180.201	25.111	NA	NA

TABLE 113 – AS TREATED MODEL PREDICTED MEANS AND STD. ERR. PROTECTIVE CAPACITIES FOR REGION 5

	ISS not workable	ISS received	SAU
Stepdown	NA	7.42 (0.67)	4.32 (0.48)
6 Month	NA	7.72 (1.46)	6.09 (0.83)

TABLE 114 - SUMMARY RESULTS – GLMM AT ANALYSIS PROTECTIVE CAPACITIES STATEWIDE

Time point	Variable	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Stepdown	Intake Dosage	1	588.527	588.527	44.425	0.000
Stepdown	Group	2	996.601	498.301	37.614	0.000
Stepdown	PREMISS Score	1	66.466	66.466	5.017	0.025
Stepdown	Residuals	731	9684.124	13.248	NA	NA
Six month	Intake Dosage	1	128.047	128.047	6.606	0.011
Six month	Group	2	269.724	134.862	6.957	0.001
Six month	PREMISS Score	1	10.822	10.822	0.558	0.456
Six month	Residuals	266	5156.315	19.385	NA	NA

TABLE 115 - MODEL PREDICTED MEANS AND STD. ERR. FOR PROTECTIVE CAPACITIES FOR REGIONS STATEWIDE

	ISS not workable	ISS received	SAU
Stepdown	2.59 (0.36)	6.02 (0.28)	3.68 (0.17)
6 Months	4.84 (1.01)	7.71 (0.55)	5.48 (0.32)

FIGURE 16 – AS TREATED BOXPLOT OF PROTECTIVE CAPACITIES BY MEASUREMENT POINT FOR ALL REGIONS

104
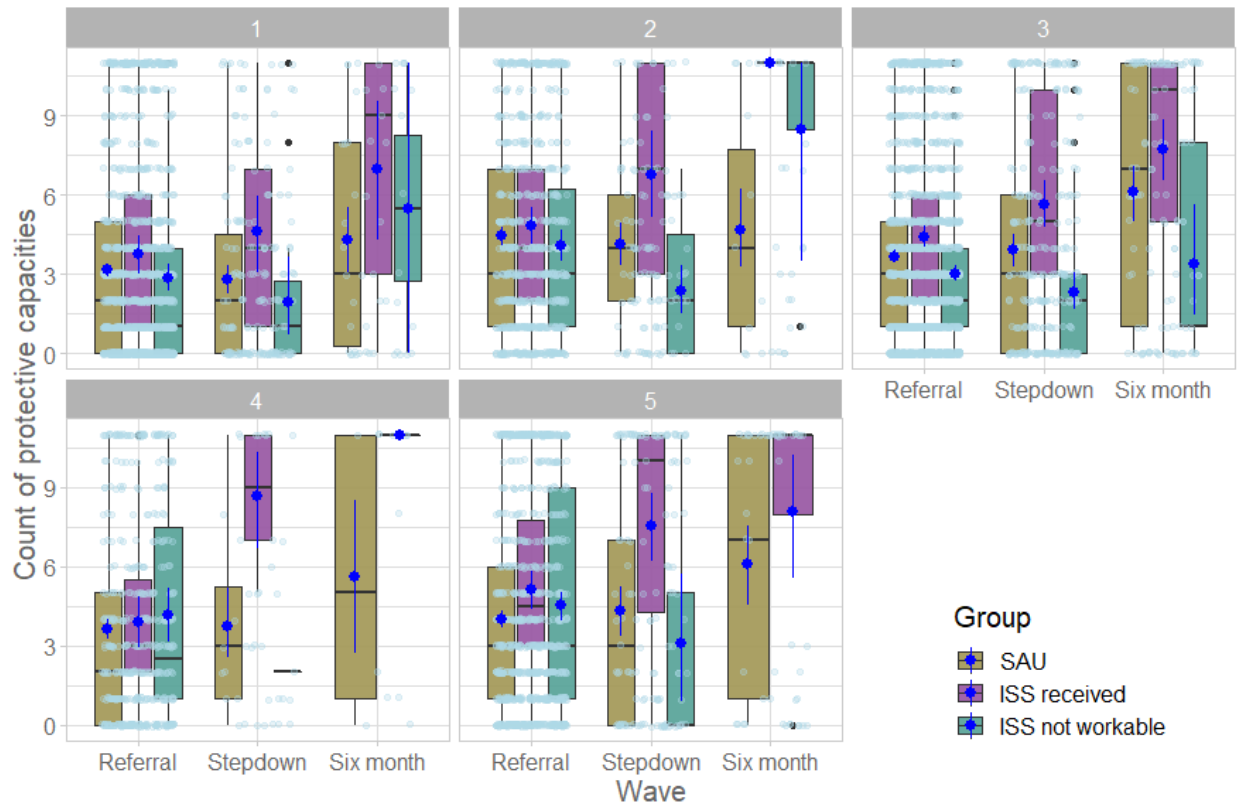
Question 3: Does ISS result in a decline in pre- to post-assessment depressive symptoms, interpersonal conflicts, and substance abuse concerns?

The key takeaways for this question include:

- Caregivers participating in ISS experienced reduced symptoms of depression, interpersonal conflicts, and substance abuse concerns over time
- ISS was particularly effective in reducing the higher rates of clinical depression in female caregivers
- ISS was particularly effective in reducing substance abuse in male caregivers
- Overall (and by Region) the conflict scales are improving (decreasing) for all subscales, suggesting the program is having good effect in decreasing the level of conflict in the client families.
- The percentage of participants with clinically significant drug abuse decreases over the six-month intervention

Methods

The data used to address this question come from the Center for Epidemiological Studies Depression (CESD) Short form, the Child Well-Being Scale (CWBS) and the Conflict Tactics Scale 2 (CTS2) Short Form. The data only include the ISS Received treatment group since this outcome is concerned with the effects of ISS on those who received it. These data were collected at the following intervals: (baseline, 6 weeks, 6 months). Sample sizes varied dependent on which form was considered: 423 CESD, 332 CWBS, 389 CTS2. Only those cases with forms completed at the baseline visit are included in the analysis. A small number (as many as 30 for some measures) had completed forms at six-weeks but not at baseline; these were excluded. The analysis consists primarily of descriptive group statistics with some group comparisons using t-tests for continuous outcomes, Pearson chi-square tests for categorical outcomes. In cases where measures are obtained for both caregivers in one family, both are included in this report. However, this circumstance was relatively rare. Therefore, in most cases the primary caregiver (parent 1) is used in the analyses.

Depressive Symptoms

To gauge depressive symptoms and risk for depression in ISS caregivers, ISS caregivers are administered depression screeners. Prior to January 2016, the Beck Depression Inventory (BDI) was used. This measure was transitioned to the Center for Epidemiological Studies Depression – Short Form (CESD-SF) screener. The BDI data have not changed since presented in previous reports; thus, they are not included here. Only the CESD-SF data are presented here, in addition to information from the Child Well-Being Scale (CWBS).

Center for Epidemiological Studies Depression –Short Form (CESD-SF)

The CESD-SF is collected from ISS caregivers at the baseline, stepdown, and six-months measurement points. The CESD-SF is a 10-item measure. The response scale for the CESD-SF is a Likert-type response scale from 0 to 4 asking about frequency of depressive symptoms with 0 indicating rarely or none of the time and 4 indicating all the time. Possible overall scores range from 0 – 30 (items #5 and #8 are reverse scored), with higher scores indicating more depressive symptoms. A total score of 10 or greater is

considered an indication of clinical depression. Table 116 provides a list of the ten items used and the rating scale.

TABLE 116: DEPRESSIVE SYMPTOM ITEMS INCLUDED IN THE **CESD-SF** AND THEIR SCORING.

During the past week	Rarely or none of the time (less than 1 day)	Some or a little of the time (1-2 days)	Occasionally or a moderate amount of the time (3-4 days)	All of the time (5-7days)
1. I was bothered by things that usually don't bother me…	0	1	2	3
2. I had trouble keeping my mind on what I was doing…	0	1	2	3
3. I felt depressed…	0	1	2	3
4. I felt that everything I did was an effort…	0	1	2	3
5. I felt hopeful about the future…	3	2	1	0
6. I felt fearful…	0	1	2	3
7. My sleep was restless…	0	1	2	3
8. I was happy…	3	2	1	0
9. I felt lonely…	0	1	2	3
10. I could not "get going"…	0	1	2	3

CESD-SF Total Scores for each caregiver

Figure 17 shows the overall distribution of CESD-SF total score at the baseline and when the caregivers are stepped down to CHBS; Figure 18 breaks this information down by region. The figures illustrate median (dark lines across the boxes), the interquartile range (top and bottom edge of the boxes), and the limits for outliers (the ends of the whiskers). Points represent individual scores and are jittered to ensure ease of viewing. The red horizontal line represents the risk score of 10 as the cutoff for clinical depression. The participants who fall above this line are considered depressed.

The average CESD score decreases for both female and male caregivers from baseline to stepdown (~six weeks after baseline) to six months, although the decreases are only significant for female caregivers (p < 0.001 and p = 0.021 for consecutive intervals). The average CESD score for female caregivers is significantly higher than for male caregivers at both baseline (p = 0.0001) and stepdown (p = 0.020).

This suggests that ISS intervention has a positive effect in reducing depression levels, as measured by the CESD, for both female and male caregivers but particularly for the female caregivers were the need is higher.

Figure 18 shows these results broken down by region. The same general trends are seen in all regions.

FIGURE 18: TOTAL CESD SCORES FOR EACH CAREGIVER (BY REGION).
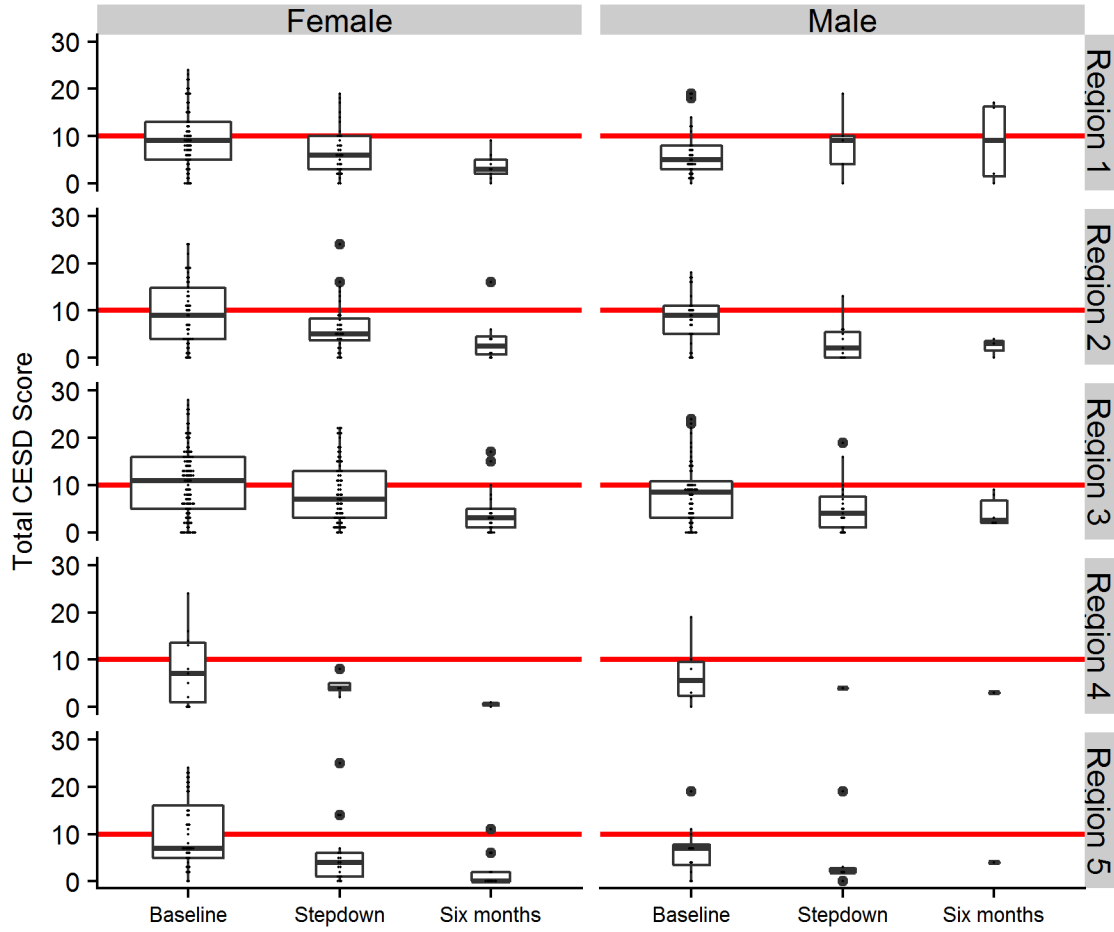
Table 117 presents this information in numerical terms overall and for each region.

	Female caregiver			Male caregiver		
Region	Baseline	Stepdown	Six months	Baseline	Stepdown	Six months
Overall	10.2 ± 7.1 [0 - 28] N=285	7.3 ± 6.0 [0 - 25] N=166	3.4 ± 4.2 [0 - 17] N=53	7.6 ± 5.9 [0 - 24] N=138	5.0 ± 5.5 [0 - 19] N=42	5.0 ± 5.3 [0 - 17] N=15
R1.	9.6 ± 6.4 [0 - 24] N=65	7.1 ± 5.1 [0 - 19] N=35	3.7 ± 2.7 [0 - 9] N=9	6.4 ± 5.2 [0 - 19] N=33	8.4 ± 7.2 [0 - 19] N=5	8.8 ± 9.0 [0 - 17] N=4
R2.	9.4 ± 6.9 [0 - 24] N=50	6.6 ± 5.3 [0 - 24] N=32	4.0 ± 5.3 [0 - 16] N=8	8.4 ± 5.5 [0 - 18] N=27	3.4 ± 4.0 [0 - 13] N=11	2.3 ± 2.1 [0 - 4] N=3
R3.	10.9 ± 7.3 [0 - 28] N=115	8.4 ± 6.6 [0 - 22] N=76	4.2 ± 4.8 [0 - 17] N=21	8.2 ± 6.4 [0 - 24] N=60	5.3 ± 5.4 [0 - 19] N=19	4.3 ± 3.3 [2 - 9] N=6
R4.	8.1 ± 7.9 [0 - 24] N=11	4.5 ± 2.5 [2 - 8] N=4	0.5 ± 0.7 [0 - 1] N=2	7.0 ± 7.0 [0 - 19] N=6	4.0 ± NA [4 - 4] N=1	3.0 ± NA [3 - 3] N=1
R5.	10.6 ± 7.3 [0 - 24] N=44	5.3 ± 6.3 [0 - 25] N=19	2.1 ± 3.5 [0 - 11] N=13	6.5 ± 5.3 [0 - 19] N=12	4.7 ± 7.1 [0 - 19] N=6	4.0 ± NA [4 - 4] N=1

mean ± SD; [min - max]

Change in CESD-SF Total Score

The change in CESD-SF score for female and male caregivers who had CESD-SF total scores for both baseline and stepdown measurements and for both stepdown and six-month measurements are shown for all regions combined in Figure 19. The lines join the pairs of measurements for each caregiver, superimposed on the total CESD-SF scores shown in Figure 17. Green lines indicate that the caregiver had a total CESD-SF score of 10 or above (clinically depressed) at the earlier time-point that then decreased (less depressed). Red lines indicate that the caregiver had a total CESD-SF score of 10 or above at the earlier time-point that then increased (more depressed). Grey lines join caregiver scores that were below 10 at both time-points
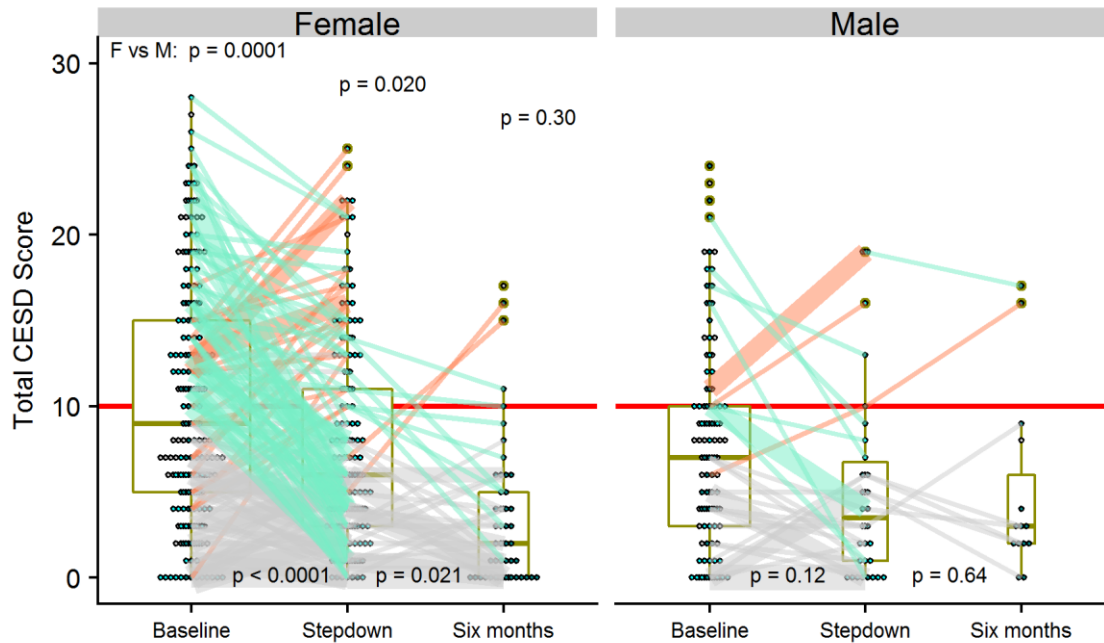
These figures visually display the extent of positive change among caregivers receiving ISS who were clinically depressed to not clinically depressed. While some caregivers increased the number of symptoms (red lines), few transitioned over time to clinical depression.

Below in Figure 20 is a river-plot to show the progression of caregivers, from clinical depression to not clinically depressed as measured by the CESD-SF. To interpret the river-plots, imagine all of the participants stacked in each column categorized by whether they are clinically significant or not clinically significant. Beginning from the left-side (Baseline) column, for each participant a line is drawn from their right-side to their left-side in the next column to the right in their clinical category in that next time-point. Ordering the participants, within their category, to group those with similar progressions from left-to-right then forms the ribbons. Those that do not have a ribbon flowing from their right-side did not appear in the next time-point, possibly having exited the program (for whatever reason).

Figure 20 suggests that the decrease in the number of caregivers who are considered at risk for depression (per their CESD score) is not solely the result of caregivers decreasing their CESD score and progressing to not depressed, but because a large proportion are exiting the program (for whatever reason). However, among those that do not exit early about half are transitioning from depressed to not depressed. Only a small number transition in the reverse and become depressed.

Figure 21 presents the number and percent of female and male caregivers who scored above the risk threshold of a CESD-SF Total score = 10 (red) at each event versus those who did not for all regions combined. The percentage of those above the threshold for clinical depression decreases significantly (p = 0.0001) from nearly 50% (females at baseline) to around 10% after six months of intervention. This decrease was seen in both female and male caregivers, suggesting that the overall impact of the program is positive. Female caregivers showed a significantly (p = 0.0024) higher rate of depression as baseline (49%) compared to male caregivers (33%).

FIGURE 21: INCIDENCE OF RISK FOR DEPRESSION FOR FEMALE AND MALE CAREGIVERS (OVERALL).
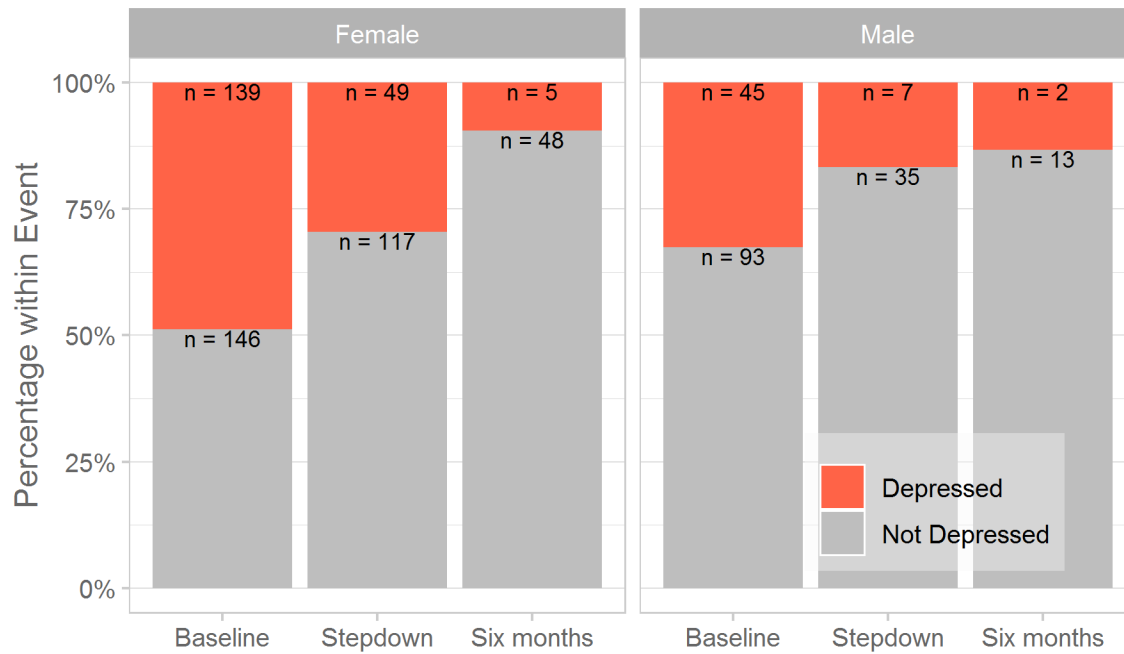
Figure 22 shows a region by region breakdown of these numbers. All gender/time point groups with more than 5 observations showed the same general trend as seen for the overall numbers.

FIGURE 22: INCIDENCE OF RISK FOR DEPRESSION FOR FEMALE AND MALE CAREGIVERS (BY REGION).

CWBS: Observations of Caregiver Distress

To monitor caseworker reports of caregiver depression, distress, and need for services, selected items from the Child Well-Being Scale (CWBS) are discussed in this section. Workers were asked to rate their observation regarding the severity level with which the caregiver is showing distress. This was recorded

at intake/baseline and then again at stepdown to CHBS. At the time of stepdown, CWBS forms were collected both from the ISS workers as they hand the caregivers off to CHBS, and by the CHBS worker during their initial assessment. Note that the CWBS is completed one per family while CESD is completed by both caregivers in the family where possible. The response scale for the item concerning caregiver distress are as follows:

1. Primary caregiver appears to be coping well and handles everyday stress and unexpected events in an appropriate manner.
2. Primary caregiver shows mild level of distress (mild symptoms, some difficulties in home or work, but generally functioning well).
3. Primary caregiver shows moderate level of distress (moderate symptoms, disruptive functioning in more than one area).
4. Primary caregiver shows severe level of distress (distress is debilitating and interferes with functioning at home, work, and the community).

Figure 23 and Figure 24 show the frequencies and percentages of caregivers rated in each category overall and by region. Generally, overall and for each region the cumulative frequency of those reported as showing mild to moderate levels of distress appears to have decreased from initiation to closing of ISS from baseline to stepdown and from stepdown to six months under CHBS.

FIGURE 23: LEVELS OF CAREGIVER DISTRESS MEASURED USING THE CWBS (OVERALL)
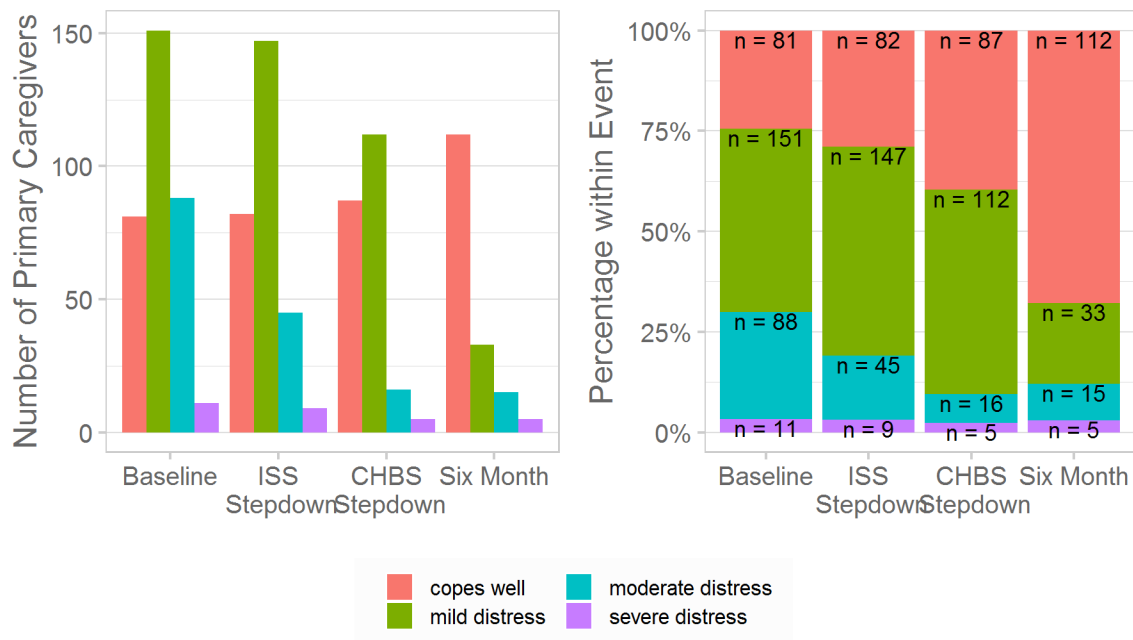
FIGURE 24: LEVELS OF CAREGIVER DISTRESS MEASURED USING THE CWBS (BY REGION).
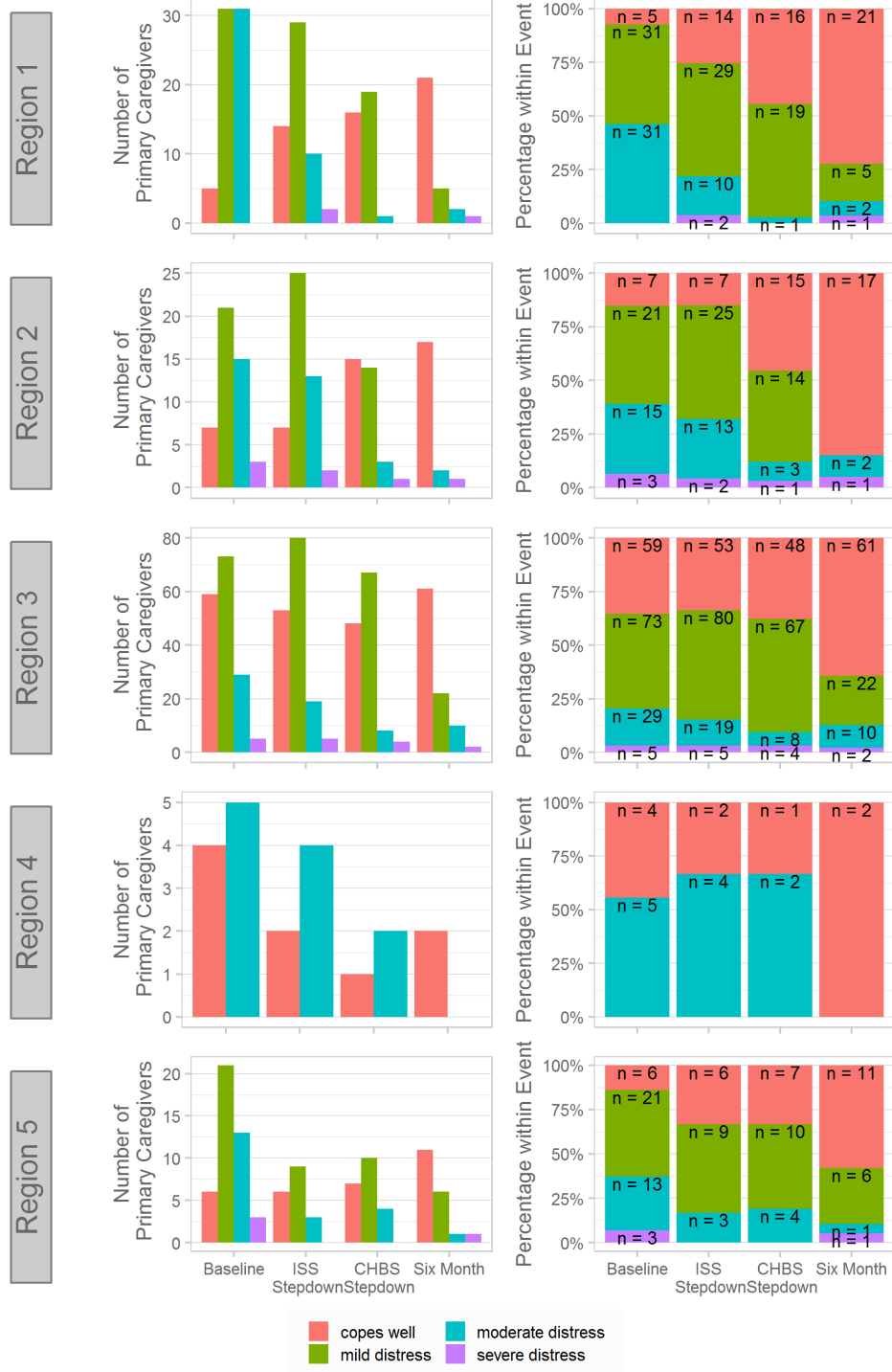
CWBS: depression or anxiety or other mental health problems

ISS workers document whether the primary caregiver has been referred to or is currently in treatment for depression or anxiety or other mental health problems, or whether the caregiver is prescribed medication for depression or anxiety or other mental health problems. Note that these questions are focused on the primary caregiver whether male, female, or missing. (Gender was missing in 1/3 of the records. However, from a survey of the other data forms, the vast majority of primary caregivers are female.) Figure 25 through Figure 34 present the counts and percentages of the responses to these questions over the course of services overall (all regions combined) and by region, as well as the progression of caregivers in referral, treatment and medication are presented. As expected, in general, referral for treatment for depression increases at stepdown and decreases after that.

FIGURE 25: HAS PRIMARY CAREGIVER BEEN REFERRED FOR INDIVIDUAL MENTAL HEALTH TREATMENT? (OVERALL).
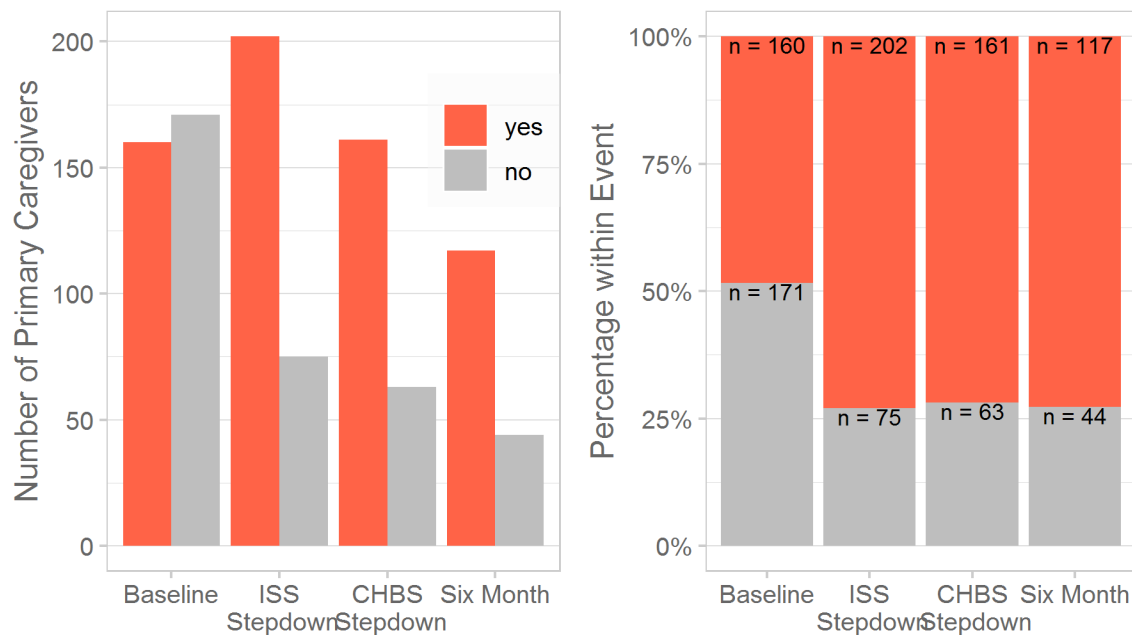
FIGURE 26: PROGRESSION IN REFERRAL FOR PRIMARY CAREGIVERS (OVERALL).
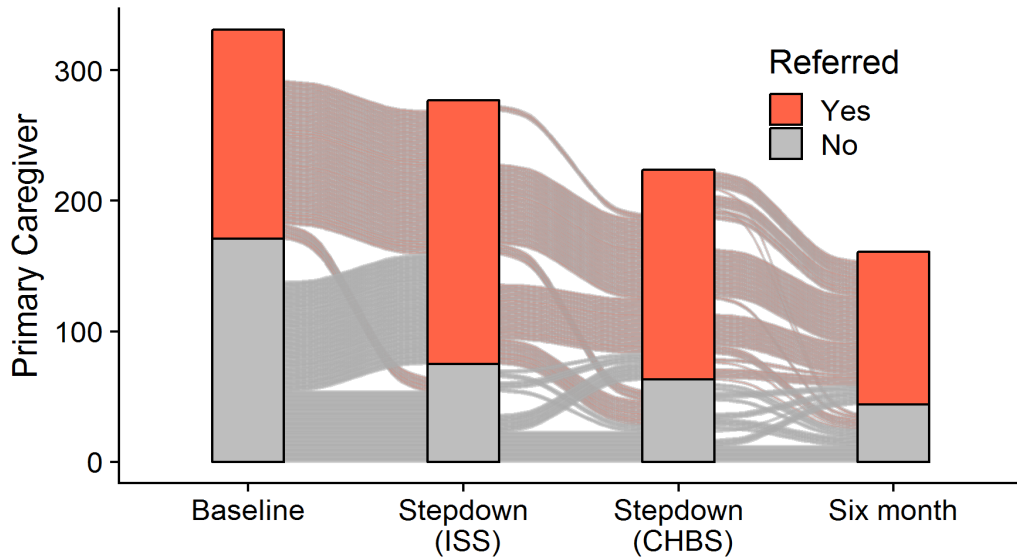
FIGURE 27: PROGRESSION FROM REFERRAL TO TREATMENT FOR PRIMARY CAREGIVERS (OVERALL).
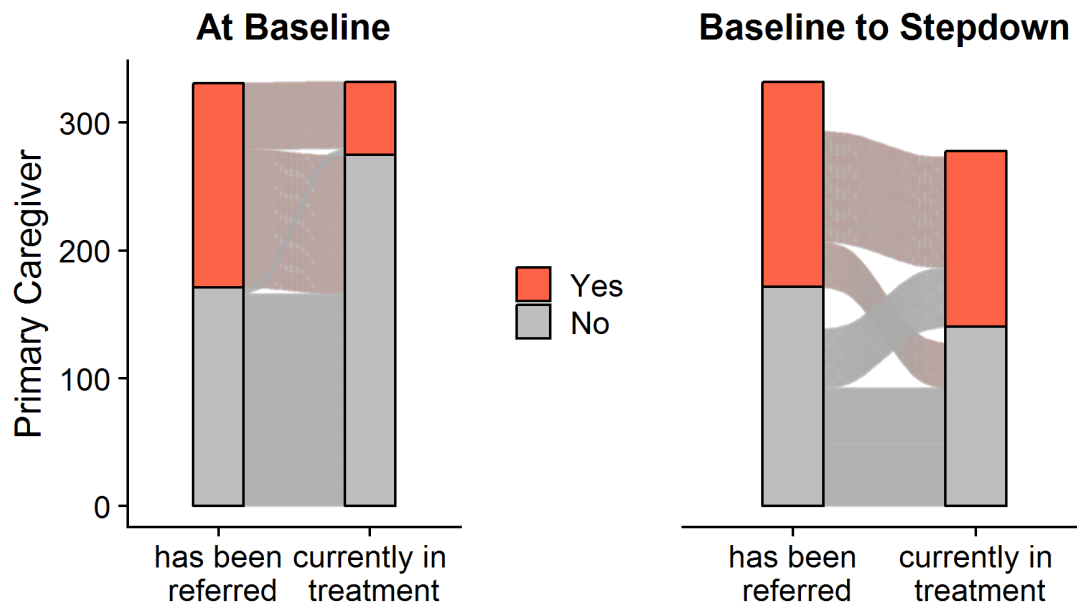
The difficulty in interpreting these data is knowing how either the worker or the primary caregiver interprets "Has primary caregiver been referred for individual treatment?" especially at the Stepdown and six-month visits. The instructions do not specify time periods. Interpretation is somewhat at the discretion of the worker. At Stepdown is the question "Has primary caregiver been referred …" in the time since baseline, in the last month or at any time previously. If it is "at any time previously" regardless of the visit then a "yes" response at Baseline implies a "yes" response at Stepdown. To some

degree this appears to be the case in Figure 27 with the wide ribbon joining most of the Baseline "yes" to the Stepdown "yes".

There is an added complication in that Baseline and Stepdown ISS are completed by the ISS therapist, and Stepdown CHBS and six-month by another provider, the CHBS worker. The CHBS worker should be aware of the referrals made in ISS.

FIGURE 28: HAS PRIMARY CAREGIVER BEEN REFERRED FOR INDIVIDUAL TREATMENT? (BY REGION).
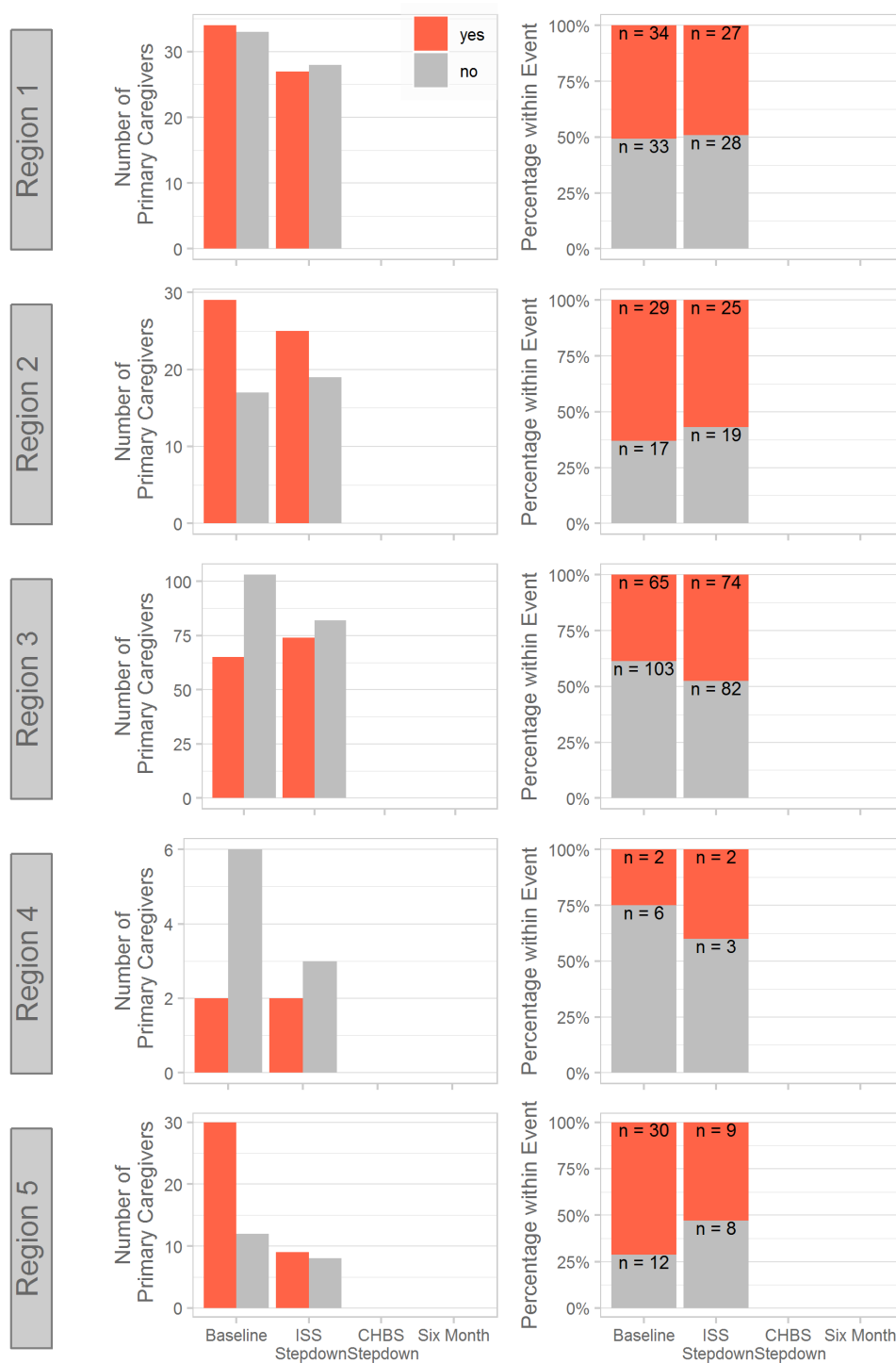
One of ISS's main goals is to get families linked to services, so yes, they should be making referrals. However, often CPS will make referrals during their assessment before the ISS case is even identified and assigned to ISS. There is no way from this measure to know where the referral originated or when.

119

There was an increase in the utilization of services from baseline as evidenced by the question about if they are currently in treatment, which may be the better approach to assessing service needs being identified and met. These results are presented below in Figure 29 to Figure 31.

FIGURE 29: IS PRIMARY CAREGIVER CURRENTLY IN INDIVIDUAL TREATMENT FOR DEPRESSION OR ANXIETY OR OTHER MENTAL HEALTH PROBLEMS? (OVERALL).

FIGURE 30: PROGRESSION IN TREATMENT FOR PRIMARY CAREGIVERS (OVERALL).
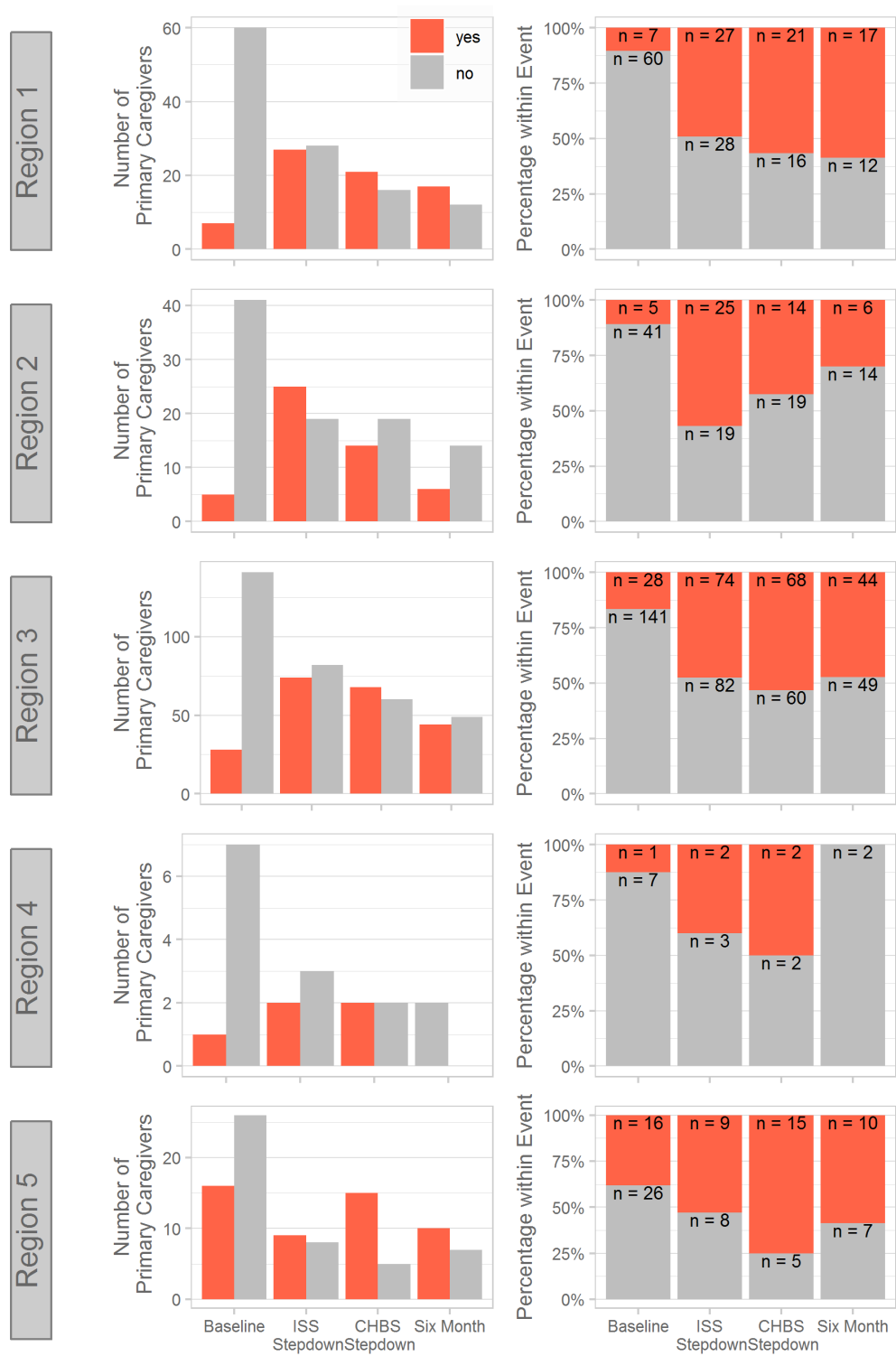
Figure 32 to Figure 34 present information on whether or not the caregiver has been prescribed medication for a mental health-related problem. In general, there is either low knowledge of these prescriptions or low rates of usage of medication, though it does increase overtime.

FIGURE 32: IS PRIMARY CAREGIVER CURRENTLY PRESCRIBED MEDICATION FOR A MENTAL HEALTH RELATED PROBLEM? (OVERALL).
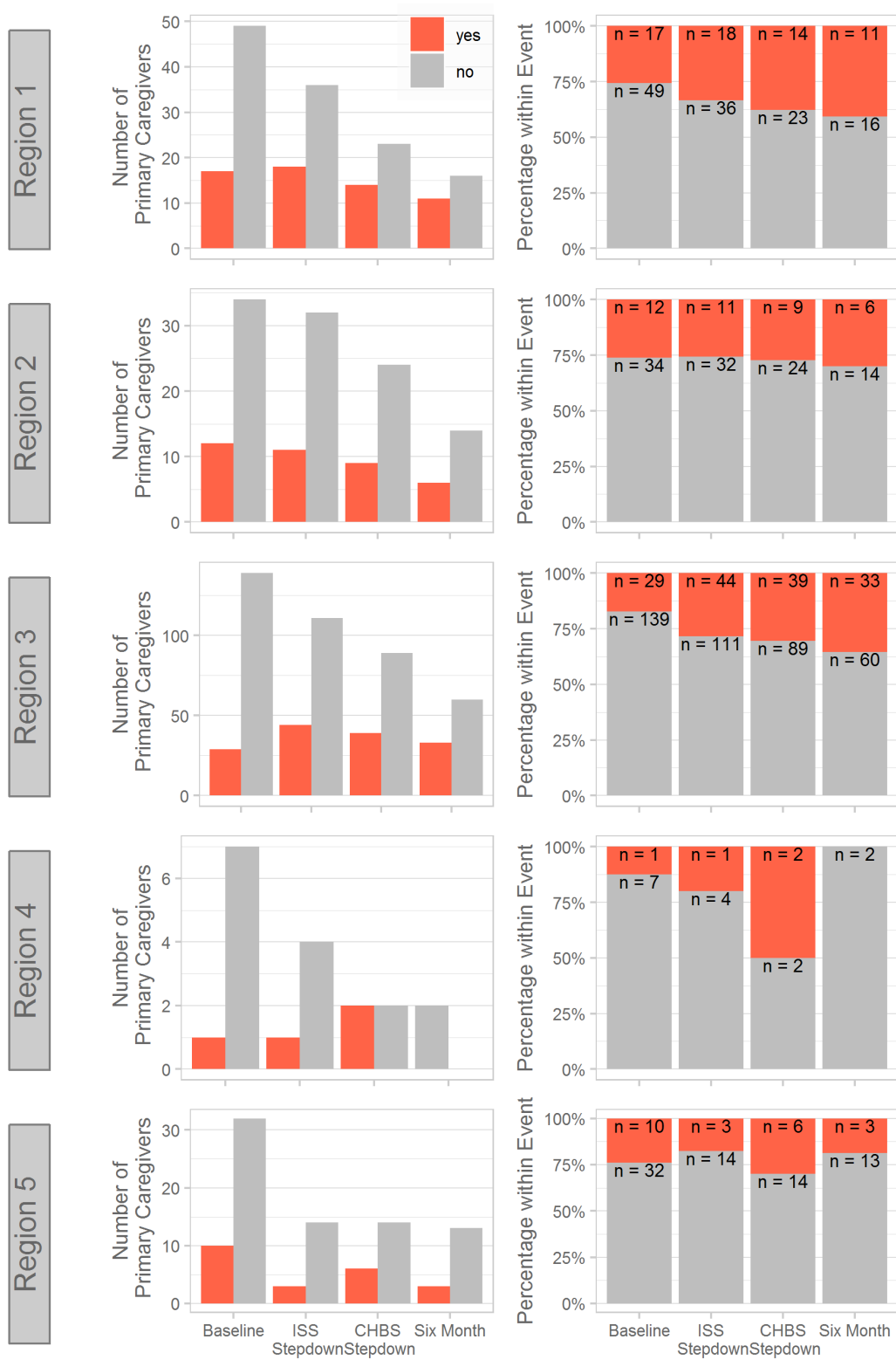
Domestic violence was measured using Conflict Tactics Scale2 (CTS2S) Short Form, as well as some items in the Child Well-Being Scale (CWBS).

CONFLICT TACTICS SCALE2 SHORT FORM (CTS2SF)

CTS2S is an instrument widely used to measure incidence of family violence. Respondents are asked a series of questions about events they may have experienced either as a victim or a perpetrator. Questions fall into five subscales: Injury, Negotiation, Psychological Aggression, Physical Assault, and Sexual Coercion. Specific behaviors are listed, and participants respond based on the frequency of their experience with that behavior. The response scale is:

0   Never has happened, or not in the past six months,

1   Once in the past six months.

2   Twice in the past six months.

3   3-5 times in the past six months.

4   6-10 times in the past six months.

5   11-20 times in the past six months.

6   more than 20 times in the past six months.

Note: "Not in the past six months" is grouped with "Never has happened" to estimate incidence.[24]

Table 118 presents the summary (mean, standard deviation, incidence of the subscales) of the total subscale scores using the summed raw response values within the item type for all regions combined (Overall), while Table 119 – Table 122 break these out into the individual regions (except Region 4 that has little data). The possible range for scores is 0-28 for this subscale (when scored in this manner), with higher scores representing higher relative frequency of this behavior/experience. Also presented are incidence rates for participants within each of the subscales. Incidence in this report is defined by a positive response (scoring 2 or more) to one or more items within each respective subscale, indicating occurrence within the past 6 months.

---

[24] Straus, M. A. (1979). *"Measuring* intrafamily conflict*Intrafamily Conflict and* violence*Violence: The* conflict tactics*Conflict Tactics (CT)* scales.*Scales."* Journal of Marriage and the Family, 75-88.

TABLE 118: CONFLICT TACTICS SCALE 2 SHORT-FORM. AVERAGE SCORE AND INCIDENCE OF CTS2 ITEMS AT BASELINE, STEPDOWN AND SIX MONTH (OVERALL).

	Baseline (n = 389)		Stepdown (n = 344)		6 month (n = 221)	
Subscales	Score	Incidence	Score	Incidence	Score	Incidence
Injury	0.9 ± 1.9	21%	0.8 ± 2.3	16%	0.1 ± 0.6 † ‡	3% † ‡
Negotiation	12.8 ± 7.2	89%	11.3 ± 7.7	81% *	9.3 ± 7.7 ‡	70% ‡
Psychological Aggression	5.8 ± 5.1	71%	4.0 ± 4.9 *	56% *	1.4 ± 2.5 † ‡	30% † ‡
Physical Assault	1.6 ± 3.2	28%	1.3 ± 3.0 *	20%	0.3 ± 1.5 † ‡	5% † ‡
Sexual Coercion	0.2 ± 1.1	6%	0.1 ± 1.0	4%	0.0 ± 0.5	1%

Incidence (Score = 2); * $p < 0.05$ Baseline vs Stepdown; † $p < 0.05$ Stepdown vs 6 month; ‡ $p < 0.05$ Baseline vs 6 month

TABLE 119: CONFLICT TACTICS SCALE 2 SHORT-FORM. AVERAGE SCORE AND INCIDENCE OF CTS2 ITEMS AT BASELINE, STEPDOWN AND SIX MONTH (REGION 1).

	Baseline (n = 68)		Stepdown (n = 63)		6 month (n = 35)	
Subscales	Score	Incidence	Score	Incidence	Score	Incidence
Injury	0.5 ± 1.3	16%	0.0 ± 0.2	0% *	0.0 ± 0.0	0%
Negotiation	13.4 ± 6.3	95%	12.0 ± 8.0	79%	9.9 ± 7.2 ‡	74% ‡
Psychological Aggression	5.2 ± 4.4	77%	3.4 ± 4.3	57%	2.2 ± 2.6 † ‡	47% † ‡
Physical Assault	0.9 ± 2.0	19%	0.1 ± 0.4	4%	0.0 ± 0.0 ‡	0% ‡
Sexual Coercion	0.1 ± 0.5	4%	0.1 ± 0.6	4%		

Incidence (Score = 2); * $p < 0.05$ Baseline vs Stepdown; † $p < 0.05$ Stepdown vs 6 month; ‡ $p < 0.05$ Baseline vs 6 month.

TABLE 120: CONFLICT TACTICS SCALE 2 SHORT-FORM. AVERAGE SCORE AND INCIDENCE OF CTS2 ITEMS AT BASELINE, STEPDOWN AND SIX MONTH (REGION 2).

Subscales	Baseline (n = 52)		Stepdown (n = 51)		6 month (n = 25)	
	Score	Incidence	Score	Incidence	Score	Incidence
Injury	0.7 ± 1.7	14%	0.5 ± 1.1	16%	0.1 ± 0.2	0%
Negotiation	12.5 ± 7.4	82%	12.0 ± 7.9	82%	12.1 ± 8.0	79%
Psychological Aggression	4.8 ± 4.3	62%	3.3 ± 4.5	42%	2.2 ± 3.6 ‡	44% ‡
Physical Assault	1.2 ± 2.7	20%	0.7 ± 1.9	15%	0.2 ± 0.9	5%
Sexual Coercion	0.1 ± 0.5	2%	0.4 ± 2.1	3%		

Incidence (Score = 2); * $p < 0.05$ Baseline vs Stepdown; † $p < 0.05$ Stepdown vs 6 month; ‡ $p < 0.05$ Baseline vs 6 month.

TABLE 121: CONFLICT TACTICS SCALE 2 SHORT-FORM. AVERAGE SCORE AND INCIDENCE OF CTS2 ITEMS AT BASELINE, STEPDOWN AND SIX MONTH (REGION 3).

Subscales	Baseline (n = 206)		Stepdown (n = 180)		6 month (n = 128)	
	Score	Incidence	Score	Incidence	Score	Incidence
Injury	1.1 ± 2.1	24%	1.1 ± 2.9	21%	0.2 ± 0.7 † ‡	5% † ‡
Negotiation	13.0 ± 7.1	89%	11.5 ± 7.8	81%	9.1 ± 7.8 ‡	70% ‡
Psychological Aggression	6.2 ± 5.5	72%	4.6 ± 5.2 *	61%	1.1 ± 2.3 † ‡	23% † ‡
Physical Assault	2.0 ± 3.7	33%	1.8 ± 3.6	27%	0.4 ± 1.7 † ‡	6% † ‡
Sexual Coercion	0.3 ± 1.3	7%	0.1 ± 0.4 *	4%	0.1 ± 0.7	1%

Incidence (Score = 2); * $p < 0.05$ Baseline vs Stepdown; † $p < 0.05$ Stepdown vs 6 month; ‡ $p < 0.05$ Baseline vs 6 month.

TABLE 122: CONFLICT TACTICS SCALE 2 SHORT-FORM. AVERAGE SCORE AND INCIDENCE OF CTS2 ITEMS AT BASELINE, STEPDOWN AND SIX MONTH (REGION 5).

	Baseline (n = 51)		Stepdown (n = 42)		6 month (n = 30)	
Subscales	Score	Incidence	Score	Incidence	Score	Incidence
Injury	0.9 ± 1.8	24%	0.3 ± 0.9	10%	0.0 ± 0.0	0%
Negotiation	11.2 ± 8.1	78%	7.7 ± 6.8	71%	6.8 ± 7.3	53%
Psychological Aggression	5.2 ± 5.2	67%	2.9 ± 4.1	43%	0.8 ± 1.5 ‡	29% ‡
Physical Assault	1.2 ± 2.5	22%	0.7 ± 1.9	14%	0.4 ± 1.5	6%
Sexual Coercion	0.3 ± 1.1	6%	0.2 ± 0.9	5%		

Incidence (Score = 2); * $p < 0.05$ Baseline vs Stepdown; † $p < 0.05$ Stepdown vs 6 month; ‡ $p < 0.05$ Baseline vs 6 month.

Overall (and by Region) the conflict scales are improving (decreasing) for all subscales, suggesting the program is having good effect in decreasing the level of conflict in the client families.

CHILD WELL-BEING SCALE (CWBS)

To monitor caseworker reports of family conflict, selected items from the Child Well-Being Scale (CWBS) are discussed in this section. Workers were asked to rate their observation regarding the severity level of conflict between caregivers. This was recorded at intake/baseline and then again at stepdown to CHBS. At the time of stepdown, CWBS forms were collected both from the ISS workers as they hand the clients off to CHBS, and by the CHBS worker during their initial assessment.

Caregiver - Partner Conflict Issues

The response levels for the item concerning caregiver-partner control are:

1   Caregiver and partner resolve issues well. No concerning areas of conflict.

2   Minor difficulties with resolving conflict.

3   Conflict becomes heated at times or chronic low-level conflict occurring

4   Domestic disputes/Conflict in the home is severe enough to cause minor harm to or fear in family members.

5   Caregiver and/or child have been harmed by or are experiencing significant fear associated with domestic violence in last month.

The results for the analyses on Caregiver-Parent Conflict Issues are presented in Figure 35 and Table 123 to Table 127.

FIGURE 35: LEVELS OF PARENTAL CONFLICT MEASURED USING THE CWBS (OVERALL).

TABLE 123: CAREGIVER/PARTNER CONFLICT ISSUES MEASURED USING THE CWBS (OVERALL).

		ISS		CHBS	
		Baseline (n = 215)	Stepdown (n = 191)	Stepdown (n = 121)	Six months (n = 90)
Average score (mean ± SD)		2.5 ± 1.5	2.3 ± 1.3	2.0 ± 1.2	1.5 ± 0.9
1	Caregiver and partner resolve issues well. No concerning areas of conflict.	74 (34%)	62 (32%)	58 (48%)	62 (69%)
2	Minor difficulties with resolving conflict.	44 (20%)	55 (29%)	27 (22%)	16 (18%)
3	Conflict becomes heated at times or chronic low-level conflict occurring	35 (16%)	36 (19%)	18 (15%)	7 (8%)
4	Domestic disputes/Conflict in the home is severe enough to cause minor harm to or fear in family members.	30 (14%)	23 (12%)	13 (11%)	4 (4%)
5	Caregiver and/or child have been harmed by or are experiencing significant fear associated with domestic violence in last month.	32 (15%)	15 (8%)	5 (4%)	1 (1%)

TABLE 124: CAREGIVER/PARTNER CONFLICT ISSUES AS MEASURED USING THE CWBS (REGION 1).

		ISS		CHBS	
		Baseline (n = 39)	Stepdown (n = 33)	Stepdown (n = 26)	Six months (n = 15)
	Average score (mean ± SD)	2.3 ± 1.3	2.7 ± 1.2	1.7 ± 1.1	1.3 ± 0.5
1	Caregiver and partner resolve issues well. No concerning areas of conflict.	13 (33%)	5 (15%)	15 (58%)	10 (67%)
2	Minor difficulties with resolving conflict.	12 (31%)	11 (33%)	7 (27%)	5 (33%)
3	Conflict becomes heated at times or chronic low level conflict occurring	8 (21%)	10 (30%)	1 (4%)	0 (0%)
4	Domestic disputes/Conflict in the home is severe enough to cause minor harm to or fear in family members.	2 (5%)	3 (9%)	2 (8%)	0 (0%)
5	Caregiver and/or child have been harmed by or are experiencing significant fear associated with domestic violence in last month.	4 (10%)	4 (12%)	1 (4%)	0 (0%)

TABLE 125: CAREGIVER/PARTNER CONFLICT ISSUES AS MEASURED USING THE CWBS (REGION 2).

		ISS		CHBS	
		Baseline (n = 29)	Stepdown (n = 32)	Stepdown (n = 20)	Six months (n = 13)
	Average score (mean ± SD)	2.5 ± 1.1	2.6 ± 1.1	1.8 ± 0.9	1.2 ± 0.6
1	Caregiver and partner resolve issues well. No concerning areas of conflict.	6 (21%)	6 (19%)	10 (50%)	11 (85%)
2	Minor difficulties with resolving conflict.	10 (34%)	8 (25%)	6 (30%)	1 (8%)
3	Conflict becomes heated at times or chronic low level conflict occurring	6 (21%)	12 (38%)	3 (15%)	1 (8%)
4	Domestic disputes/Conflict in the home is severe enough to cause minor harm to or fear in family members.	7 (24%)	5 (16%)	1 (5%)	0 (0%)
5	Caregiver and/or child have been harmed by or are experiencing significant fear associated with domestic violence in last month.	0 (0%)	1 (3%)	0 (0%)	0 (0%)

TABLE 126: CAREGIVER/PARTNER CONFLICT ISSUES AS MEASURED USING THE CWBS (REGION 3).

		ISS		CHBS	
		Baseline (n = 116)	Stepdown (n = 106)	Stepdown (n = 63)	Six months (n = 52)
Average score (mean ± SD)		2.7 ± 1.6	2.2 ± 1.3	2.1 ± 1.3	1.7 ± 1.0
1	Caregiver and partner resolve issues well. No concerning areas of conflict.	42 (36%)	44 (42%)	28 (44%)	33 (63%)
2	Minor difficulties with resolving conflict.	17 (15%)	29 (27%)	12 (19%)	9 (17%)
3	Conflict becomes heated at times or chronic low level conflict occurring	16 (14%)	12 (11%)	12 (19%)	6 (12%)
4	Domestic disputes/Conflict in the home is severe enough to cause minor harm to or fear in family members.	17 (15%)	12 (11%)	8 (13%)	3 (6%)
5	Caregiver and/or child have been harmed by or are experiencing significant fear associated with domestic violence in last month.	24 (21%)	9 (8%)	3 (5%)	1 (2%)

TABLE 127: CAREGIVER/PARTNER CONFLICT ISSUES AS MEASURED USING THE CWBS (REGION 5).

		ISS		CHBS	
		Baseline (n = 25)	Stepdown (n = 16)	Stepdown (n = 11)	Six months (n = 9)
Average score (mean ± SD)		2.3 ± 1.5	2.0 ± 1.2	2.4 ± 1.5	1.4 ± 1.0
1	Caregiver and partner resolve issues well. No concerning areas of conflict.	11 (44%)	7 (44%)	5 (45%)	7 (78%)
2	Minor difficulties with resolving conflict.	4 (16%)	5 (31%)	1 (9%)	1 (11%)
3	Conflict becomes heated at times or chronic low level conflict occurring	4 (16%)	1 (6%)	2 (18%)	0 (0%)
4	Domestic disputes/Conflict in the home is severe enough to cause minor harm to or fear in family members.	3 (12%)	3 (19%)	2 (18%)	1 (11%)
5	Caregiver and/or child have been harmed by or are experiencing significant fear associated with domestic violence in last month.	3 (12%)	0 (0%)	1 (9%)	0 (0%)

Caregiver - Partner Control Issues

The response levels for the item concerning caregiver-partner control are:

1 Balance of power adequate between caregiver and partner.

2 Caregiver or partner exhibits more power/control in the relationship than the other.

3 Caregiver or partner has been controlling that has caused restrictions in the other's choices, behaviors, and options.

4 Level of control partner/caregiver has caused minor problems for caregiver or family members.

5 Caregiver and/or child have been harmed by or are experiencing fear associated with power and control issues.

The results for the analyses on Caregiver-Partner Control Issues are presented in Figure 36 and Table 128 to Table 132 below.

FIGURE 36: LEVELS OF PARENTAL CONTROL MEASURED USING THE CWBS (OVERALL).

133
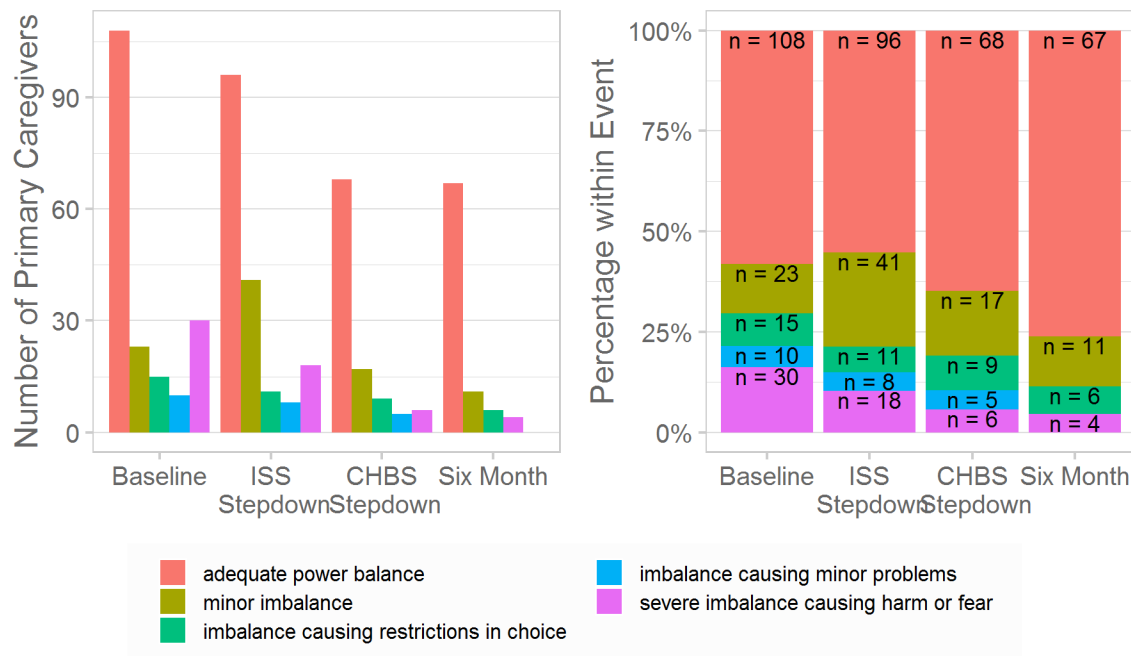
TABLE 128: CAREGIVER/PARTNER CONTROL ISSUES FROM THE CWBS (OVERALL).

	ISS		CHBS	
	Baseline (n = 186)	Step-down (n = 174)	Stepdown (n = 105)	Six months (n = 88)
Average score (mean ± SD)	2.1 ± 1.5	1.9 ± 1.3	1.7 ± 1.2	1.4 ± 1.0
1 Balance of power adequate between caregiver and partner.	108 (58%)	96 (55%)	68 (65%)	67 (76%)
2 Caregiver or partner exhibits more power/control in the relationship than the other.	23 (12%)	41 (24%)	17 (16%)	11 (12%)
3 Caregiver or partner has been controlling that has caused restrictions in the other's choices, behaviors, and options.	15 (8%)	11 (6%)	9 (9%)	6 (7%)
4 Level of control partner/caregiver has caused minor problems for caregiver or family members.	10 (5%)	8 (5%)	5 (5%)	0 (0%)
5 Caregiver and/or child have been harmed by or are experiencing fear associated with power and control issues.	30 (16%)	18 (10%)	6 (6%)	4 (5%)

TABLE 129: CAREGIVER/PARTNER CONTROL ISSUES FROM THE CWBS (REGION 1).

	ISS		CHBS	
	Baseline (n = 36)	Step-down (n = 33)	Stepdown (n = 22)	Six months (n = 15)
Average score (mean ± SD)	1.6 ± 1.2	2.0 ± 1.3	1.4 ± 0.9	1.1 ± 0.4
1 Balance of power adequate between caregiver and partner.	25 (69%)	17 (52%)	17 (77%)	13 (87%)
2 Caregiver or partner exhibits more power/control in the relationship than the other.	6 (17%)	9 (27%)	4 (18%)	2 (13%)
3 Caregiver or partner has been controlling that has caused restrictions in the other's choices, behaviors, and options.	1 (3%)	1 (3%)	0 (0%)	0 (0%)
4 Level of control partner/caregiver has caused minor problems for caregiver or family members.	1 (3%)	3 (9%)	0 (0%)	0 (0%)
5 Caregiver and/or child have been harmed by or are experiencing fear associated with power and control issues.	3 (8%)	3 (9%)	1 (5%)	0 (0%)

TABLE 130: CAREGIVER/PARTNER CONTROL ISSUES FROM THE CWBS (REGION 2).

		ISS		CHBS	
		Baseline (n = 17)	Step-down (n = 26)	Stepdown (n = 19)	Six months (n = 13)
	Average score (mean ± SD)	1.9 ± 1.2	1.8 ± 1.1	1.6 ± 0.9	1.2 ± 0.6
1	Balance of power adequate between caregiver and partner.	10 (59%)	14 (54%)	11 (58%)	11 (85%)
2	Caregiver or partner exhibits more power/control in the relationship than the other.	1 (6%)	7 (27%)	5 (26%)	1 (8%)
3	Caregiver or partner has been controlling that has caused restrictions in the other's choices, behaviors, and options.	3 (18%)	3 (12%)	2 (11%)	1 (8%)
4	Level of control partner/caregiver has caused minor problems for caregiver or family members.	3 (18%)	1 (4%)	1 (5%)	0 (0%)
5	Caregiver and/or child have been harmed by or are experiencing fear associated with power and control issues.	0 (0%)	1 (4%)	0 (0%)	0 (0%)

TABLE 131: CAREGIVER/PARTNER CONTROL ISSUES FROM THE CWBS (REGION 3).

		ISS		CHBS	
		Baseline (n = 105)	Step-down (n = 97)	Stepdown (n = 53)	Six months (n = 49)
	Average score (mean ± SD)	2.3 ± 1.7	1.9 ± 1.3	1.8 ± 1.3	1.6 ± 1.2
1	Balance of power adequate between caregiver and partner.	57 (54%)	58 (60%)	33 (62%)	35 (71%)
2	Caregiver or partner exhibits more power/control in the relationship than the other.	10 (10%)	18 (19%)	8 (15%)	6 (12%)
3	Caregiver or partner has been controlling that has caused restrictions in the other's choices, behaviors, and options.	9 (9%)	7 (7%)	5 (9%)	4 (8%)
4	Level of control partner/caregiver has caused minor problems for caregiver or family members.	5 (5%)	3 (3%)	3 (6%)	0 (0%)
5	Caregiver and/or child have been harmed by or are experiencing fear associated with power and control issues.	24 (23%)	11 (11%)	4 (8%)	4 (8%)

		ISS		CHBS	
		Baseline (n = 23)	Step-down (n = 14)	Stepdown (n = 11)	Six months (n = 10)
Average score (mean ± SD)		1.9 ± 1.3	1.9 ± 1.4	2.0 ± 1.5	1.3 ± 0.7
1	Balance of power adequate between caregiver and partner.	13 (57%)	7 (50%)	7 (64%)	8 (80%)
2	Caregiver or partner exhibits more power/control in the relationship than the other.	5 (22%)	5 (36%)	0 (0%)	1 (10%)
3	Caregiver or partner has been controlling that has caused restrictions in the other's choices, behaviors, and options.	2 (9%)	0 (0%)	2 (18%)	1 (10%)
4	Level of control partner/caregiver has caused minor problems for caregiver or family members.	1 (4%)	0 (0%)	1 (9%)	0 (0%)
5	Caregiver and/or child have been harmed by or are experiencing fear associated with power and control issues.	2 (9%)	2 (14%)	1 (9%)	0 (0%)

Caregiver - Partner Communication Issues

The response levels for the item concerning caregiver-partner communication are:

1. Caregiver has positive communication skills with others. Feelings are expressed in a calm way. Problems are addressed with problem solving skills.

2. Minor difficulties with communication noted.

3. Caregiver has difficulties communicating effectively routinely with others.

4. Primary caregiver has had frequent verbal arguments with others. This has caused problems for the family (e.g., loss of friendships, loss of jobs). Legal intervention has not occurred.

5. Primary caregiver has been in trouble with the law or had other major problems due to poor conflict resolution skills and aggressive behaviors.

The results for the analyses on Caregiver-Partner Communication Issues are presented in Figure 37 and Table 133 to Table 137.

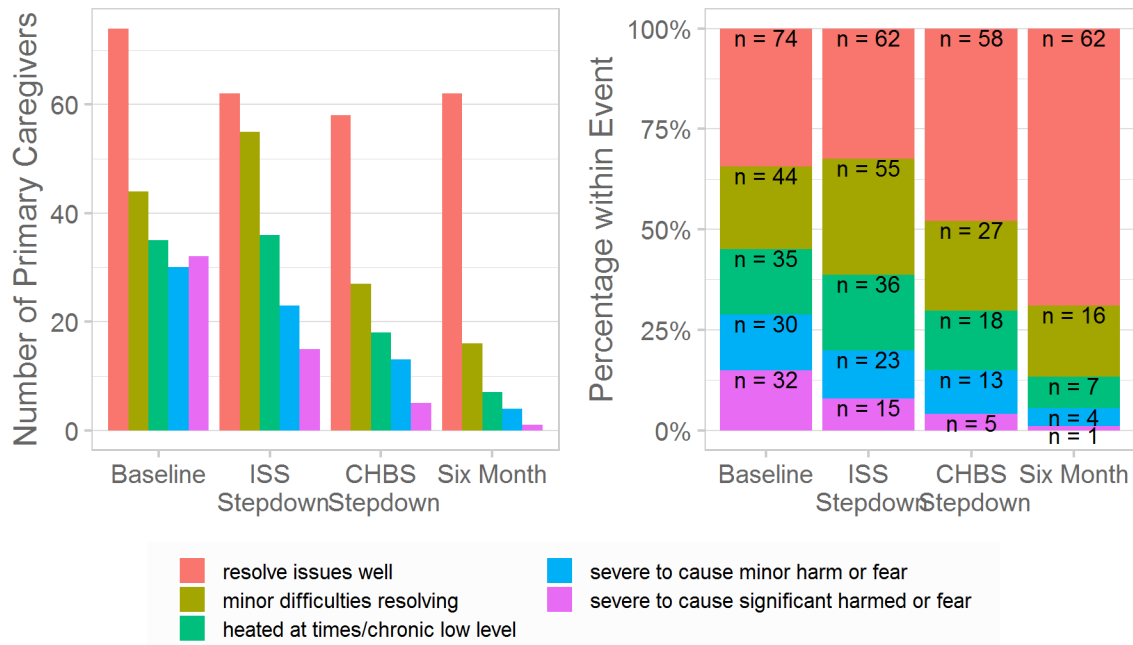**FIGURE 37: LEVELS OF PARENTAL COMMUNICATION AS MEASURED USING THE CWBS (OVERALL).**
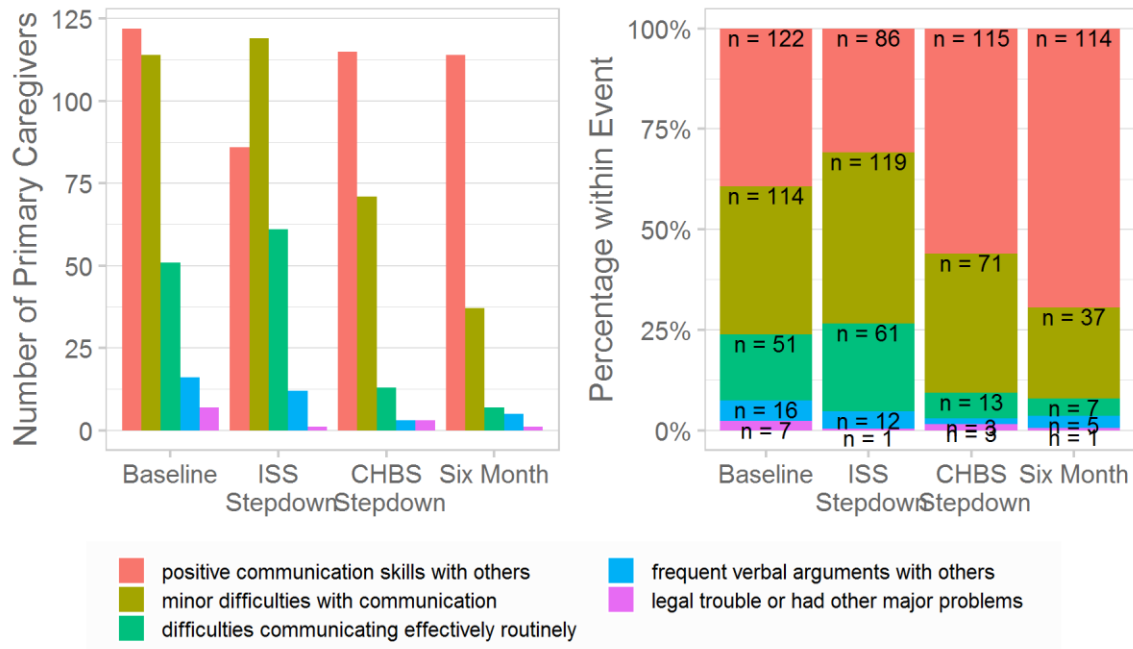
TABLE 133: CAREGIVER/PARTNER COMMUNICATION ISSUES MEASURED USING THE CWBS (OVERALL).

		ISS		CHBS	
		Baseline (n = 310)	Step-down (n = 279)	Stepdown (n = 205)	Six months (n = 164)
	Average score (mean ± SD)	1.9 ± 1.0	2.0 ± 0.9	1.6 ± 0.8	1.4 ± 0.8
1	Caregiver has positive communication skills with others. Feelings are expressed in a calm way. Problems are addressed with problem solving skills.	122 (39%)	86 (31%)	115 (56%)	114 (70%)
2	Minor difficulties with communication noted.	114 (37%)	119 (43%)	71 (35%)	37 (23%)
3	Caregiver has difficulties communicating effectively routinely with others.	51 (16%)	61 (22%)	13 (6%)	7 (4%)
4	Primary caregiver has had frequent verbal arguments with others. This has caused problems for the family (e.g., loss of friendships, loss of jobs). Legal intervention has not occurred.	16 (5%)	12 (4%)	3 (1%)	5 (3%)
5	Primary caregiver has been in trouble with the law or had other major problems due to poor conflict resolution skills and aggressive behaviors.	7 (2%)	1 (0%)	3 (1%)	1 (1%)

TABLE 134: CAREGIVER/PARTNER COMMUNICATION ISSUES MEASURED USING THE CWBS (REGION 1).

		ISS		CHBS	
		Baseline (n = 63)	Step-down (n = 54)	Stepdown (n = 37)	Six months (n = 29)
	Average score (mean ± SD)	2.3 ± 1.1	2.3 ± 1.0	1.4 ± 0.8	1.3 ± 0.8
1	Caregiver has positive communication skills with others. Feelings are expressed in a calm way. Problems are addressed with problem solving skills.	14 (22%)	12 (22%)	28 (76%)	22 (76%)
2	Minor difficulties with communication noted.	25 (40%)	20 (37%)	7 (19%)	6 (21%)
3	Caregiver has difficulties communicating effectively routinely with others.	16 (25%)	17 (31%)	1 (3%)	0 (0%)
4	Primary caregiver has had frequent verbal arguments with others. This has caused problems for the family (e.g., loss of friendships, loss of jobs). Legal intervention has not occurred.	5 (8%)	4 (7%)	0 (0%)	0 (0%)
5	Primary caregiver has been in trouble with the law or had other major problems due to poor conflict resolution skills and aggressive behaviors.	3 (5%)	1 (2%)	1 (3%)	1 (3%)

TABLE 135: CAREGIVER/PARTNER COMMUNICATION ISSUES MEASURED USING THE CWBS (REGION 2).

		ISS		CHBS	
		Baseline (n = 42)	Step-down (n = 44)	Stepdown (n = 28)	Six months (n = 19)
	Average score (mean ± SD)	2.4 ± 1.0	2.4 ± 0.7	1.6 ± 0.8	1.4 ± 1.0
1	Caregiver has positive communication skills with others. Feelings are expressed in a calm way. Problems are addressed with problem solving skills.	6 (14%)	5 (11%)	16 (57%)	15 (79%)
2	Minor difficulties with communication noted.	23 (55%)	18 (41%)	9 (32%)	2 (11%)
3	Caregiver has difficulties communicating effectively routinely with others.	7 (17%)	20 (45%)	2 (7%)	0 (0%)
4	Primary caregiver has had frequent verbal arguments with others. This has caused problems for the family (e.g., loss of friendships, loss of jobs). Legal intervention has not occurred.	4 (10%)	1 (2%)	1 (4%)	2 (11%)
5	Primary caregiver has been in trouble with the law or had other major problems due to poor conflict resolution skills and aggressive behaviors.	2 (5%)	0 (0%)	0 (0%)	0 (0%)

TABLE 136: CAREGIVER/PARTNER COMMUNICATION ISSUES MEASURED USING THE CWBS (REGION 3).

		ISS		CHBS	
		Baseline (n = 160)	Step-down (n = 157)	Stepdown (n = 120)	Six months (n = 96)
	Average score (mean ± SD)	1.7 ± 0.8	1.9 ± 0.8	1.6 ± 0.8	1.5 ± 0.8
1	Caregiver has positive communication skills with others. Feelings are expressed in a calm way. Problems are addressed with problem solving skills.	81 (51%)	56 (36%)	62 (52%)	62 (65%)
2	Minor difficulties with communication noted.	54 (34%)	71 (45%)	46 (38%)	25 (26%)
3	Caregiver has difficulties communicating effectively routinely with others.	20 (12%)	23 (15%)	8 (7%)	6 (6%)
4	Primary caregiver has had frequent verbal arguments with others. This has caused problems for the family (e.g., loss of friendships, loss of jobs). Legal intervention has not occurred.	4 (2%)	7 (4%)	2 (2%)	3 (3%)
5	Primary caregiver has been in trouble with the law or had other major problems due to poor conflict resolution skills and aggressive behaviors.	1 (1%)	0 (0%)	2 (2%)	0 (0%)

TABLE 137: CAREGIVER/PARTNER COMMUNICATION ISSUES MEASURED USING THE CWBS (REGION 5).

		ISS		CHBS	
		Baseline (n = 37)	Step-down (n = 18)	Stepdown (n = 17)	Six months (n = 18)
	Average score (mean ± SD)	2.0 ± 1.1	1.6 ± 0.6	1.7 ± 0.7	1.3 ± 0.6
1	Caregiver has positive communication skills with others. Feelings are expressed in a calm way. Problems are addressed with problem solving skills.	16 (43%)	9 (50%)	7 (41%)	13 (72%)
2	Minor difficulties with communication noted.	9 (24%)	8 (44%)	8 (47%)	4 (22%)
3	Caregiver has difficulties communicating effectively routinely with others.	8 (22%)	1 (6%)	2 (12%)	1 (6%)
4	Primary caregiver has had frequent verbal arguments with others. This has caused problems for the family (e.g., loss of friendships, loss of jobs). Legal intervention has not occurred.	3 (8%)	0 (0%)	0 (0%)	0 (0%)
5	Primary caregiver has been in trouble with the law or had other major problems due to poor conflict resolution skills and aggressive behaviors.	1 (3%)	0 (0%)	0 (0%)	0 (0%)

Substance Abuse

The CAGE-A and CAGE-B screening tests were administered to address substance abuse risk in participants. CAGE is derived as a mnemonic acronym for the items asked within each test: Cut-down, Annoyed, Guilty, and Eye-opener. CAGE-A addresses alcohol abuse; CAGE-B addresses abuse of drugs and other substances[25][26]. According to Johns Hopkins University, a positive response on two or more items is considered clinically significant. An 'alcohol' score was calculated by coding 1 for yes, 0 for no for each of the items asked then adding these up. A similar 'drug' score was calculated for the drug abuse items. For each a score of 2 or more was coded as 'Clinically Significant', 'Not Clinically Significant' otherwise.

Analyses were limited to ISS participants enrolled at baseline. Thus, participants in stepdown (4-6 weeks) and six-month follow-up appeared at baseline. Data are initially presented graphically to give an overall view of the results and then tabularly for those wanting additional details. Figure 38 and Figure 41 are bar-charts that show the overall (all regions combined) data as counts and percentages for alcohol and drug abuse, respectively.

Figure 38 shows that the percentage of participants with clinically significant alcohol abuse decreases over the six-month intervention period in both female and male caregivers. The change from Baseline to Six months was statistically significant ($p < 0.05$) for female caregivers, but the smaller numbers of male caregivers possibly precluded significance in male caregivers where a similar trend was observed.

Figure 41 shows that the percentage of participants with clinically significant drug abuse decreases over the six-month intervention period with large decreases in male caregivers, less so in female caregivers. The change from Baseline to Stepdown (4-6 weeks) and then to Six months was statistically significant ($p < 0.05$) for male caregivers with most of the decrease occurring from baseline to stepdown. Decrease was less in female caregivers and did not reach significance.

Figure 39 & Figure 40 (for alcohol abuse) and Figure 42 & Figure 43 (for drug abuse) are river-plots to show the progression of participants, hopefully, from clinically significant to not clinically significant abuse. To review, to interpret the river-plots, imagine all of the participants stacked in each column categorized by whether they are clinically significant or not clinically significant. Beginning from the left-side (Baseline) column, for each participant a line is drawn from their right-side to their left-side in the next column to the right in their clinical category in that next time-point. Ordering the participants, within their category, to group those with similar progressions from left-to-right then forms the ribbons (or 'rivers'). Those that do not have a line (ribbon) flowing from their right-side did not appear in the next time-point, possibly having exited the program (for whatever reason). Since some may have simply missed the Stepdown visit (and therefore have no line), an additional river-plot is added to the right showing progress from Baseline to Six months regardless of whether they completed the Stepdown visit.

[25] Brown, R.L., and Rounds, L.A. Conjoint screening questionnaires for alcohol and drug abuse. Wisconsin Medical Journal 94:135-140, 1995.

[26] Steinweg, D. L., and Worth, H. "Alcoholism: The Keys to the CAGE" American Journal of Medicine 94: 520-523, May 1993.

These river-plots suggest that the decrease in the proportion of participants with clinically significant abuse may be a result of a disproportionate number of clinically significant participants exiting the program rather than their abuse progressing to no longer being clinically significant.

Alcohol

The results for the analyses on Alcohol are presented in Figure 38 to Figure 40.

FIGURE 38: COUNTS OF ALCOHOL ABUSE IN FEMALE AND MALE CAREGIVERS.
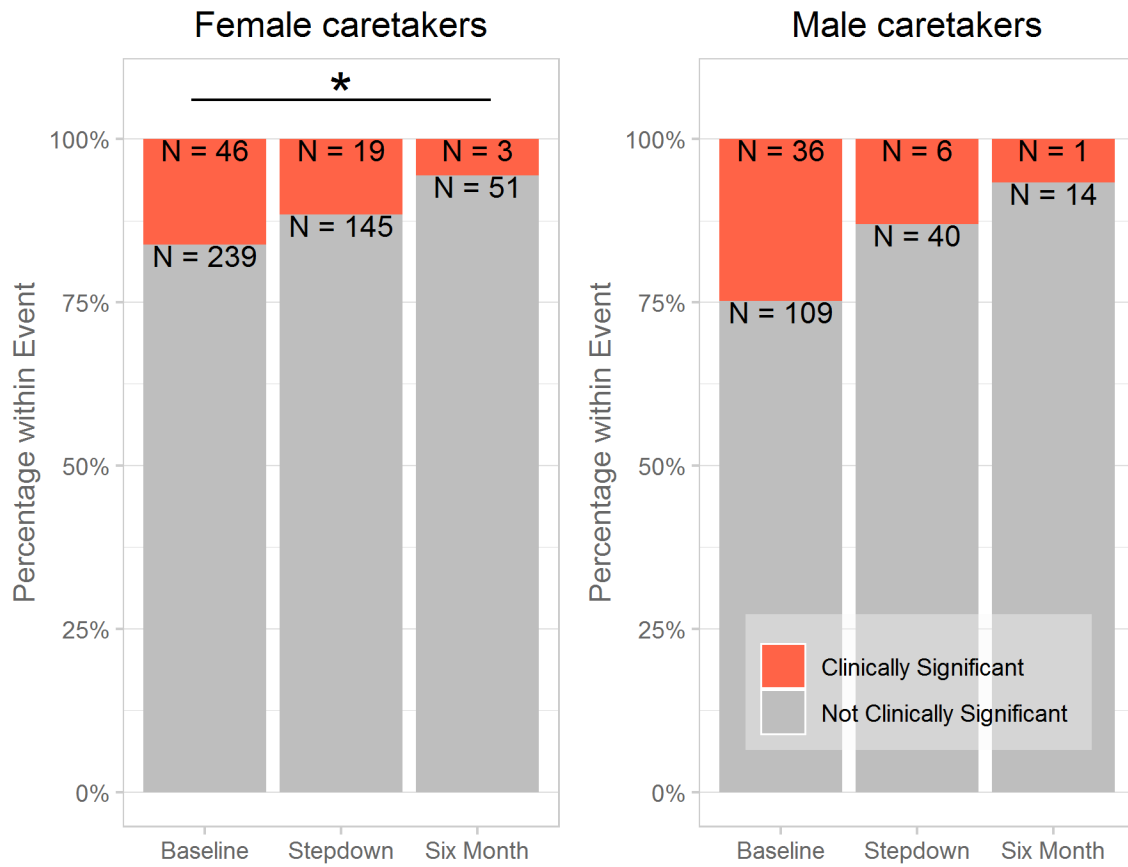
At baseline, the proportion of female caregivers with clinically significant alcohol issues (16.1%) is significantly (p = 0.030) less than the proportion of male caregivers with clinically significant alcohol issues (24.8%).

141

FIGURE 39: PROGRESSION OF ALCOHOL ABUSE IN FEMALE CAREGIVERS.
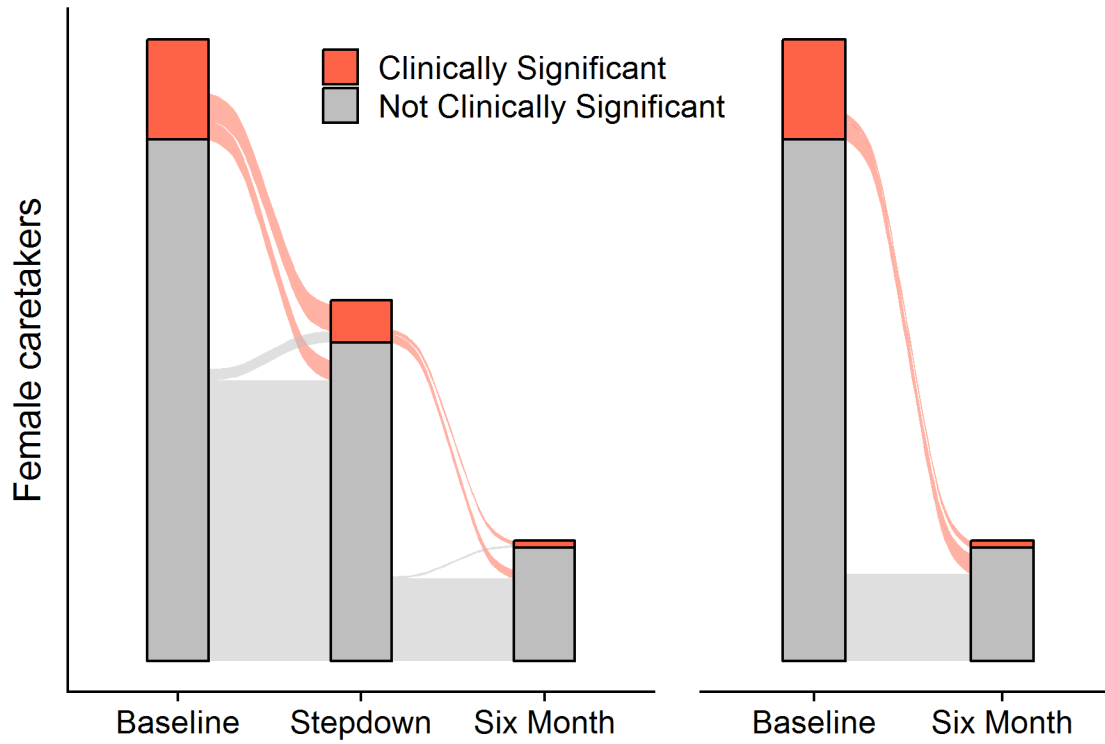
FIGURE 40: PROGRESSION OF ALCOHOL ABUSE IN MALE CAREGIVERS.
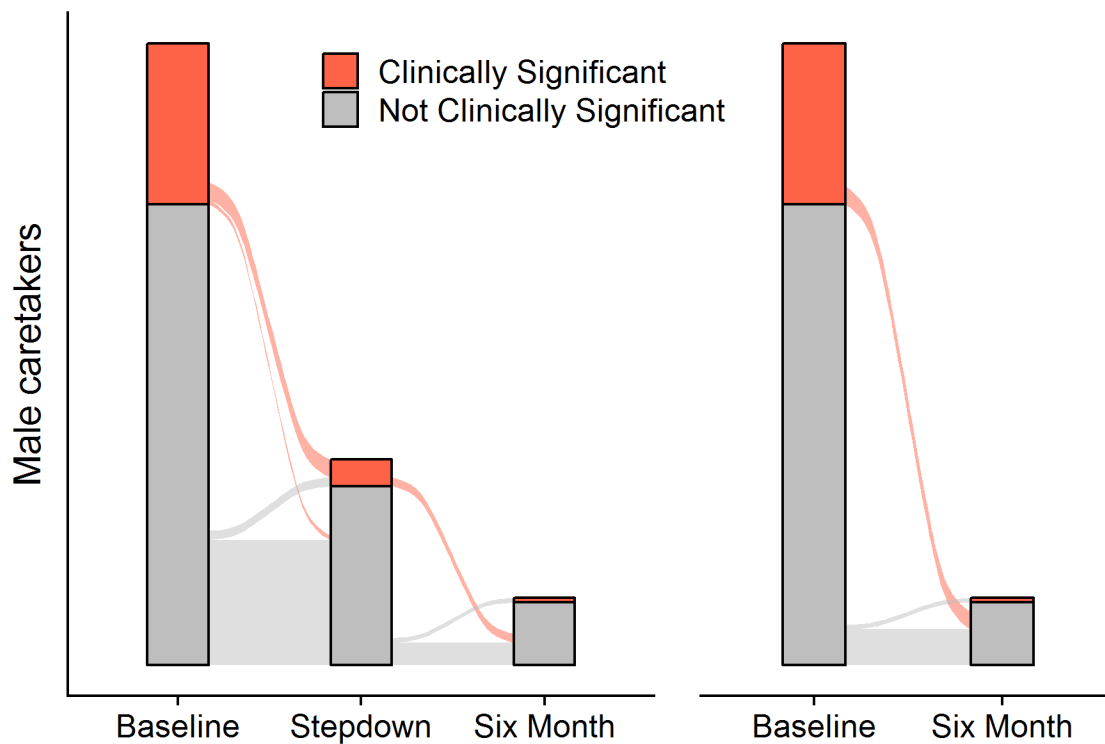
Drugs

The results for the analyses on Drugs are presented in Figure 41 to Figure 43 and Table 138 to Table 142.

At baseline, the proportion of female caregivers with clinically significant drug issues (36.1%) is not significantly (p = 0.29) different from the proportion of male caregivers with clinically significant drug issues (31.0%).

FIGURE 42: PROGRESSION OF DRUG ABUSE IN FEMALE CAREGIVERS.
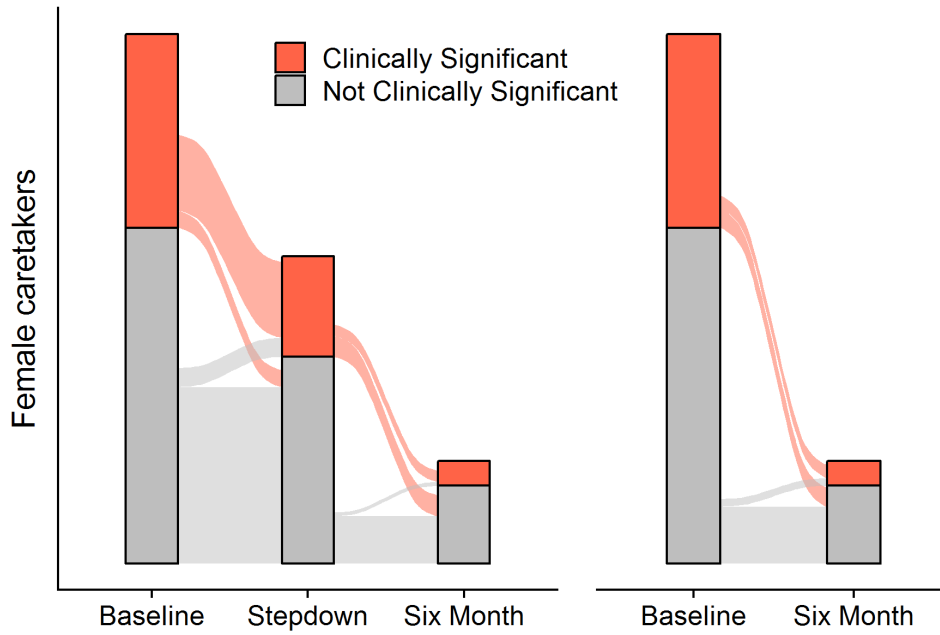
FIGURE 43: PROGRESSION OF DRUG ABUSE IN MALE CAREGIVERS.
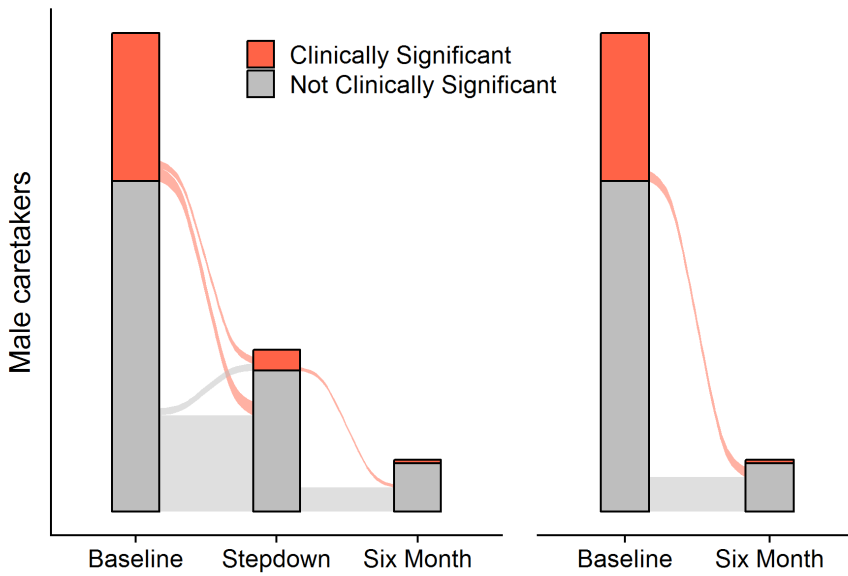
	Female caregivers			Male caregivers		
CAGE	Baseline	Stepdown	6 months	Baseline	Stepdown	6 months
Alcohol						
Cut-down	53/285 (18.6%)	19/164 (11.6%)	5/54 (9.3%)	43/145 (29.7%)	6/47 (12.8%)	2/15 (13.3%)
Annoyed	32/285 (11.2%)	19/164 (11.6%)	3/54 (5.6%)	25/145 (17.2%)	3/47 (6.4%)	0/15 (0.0%)
Guilty	44/285 (15.4%)	19/164 (11.6%)	4/54 (7.4%)	34/145 (23.4%)	7/47 (14.9%)	2/15 (13.3%)
Eye-opener	25/285 (8.8%)	9/164 (5.5%)	2/54 (3.7%)	21/145 (14.5%)	1/47 (2.1%)	0/15 (0.0%)
Clin.signif	46/285 (16.1%)	19/164 (11.6%)	3/54 (5.6%)	36/145 (24.8%)	6/47 (12.8%)	1/15 (6.7%)
Drugs						
Cut-down	106/285 (37.2%)	46/164 (28.0%)	11/54 (20.4%)	51/145 (35.2%)	7/47 (14.9%)	1/15 (6.7%)
Annoyed	65/285 (22.8%)	38/164 (23.2%)	12/54 (22.2%)	25/145 (17.2%)	4/47 (8.5%)	1/15 (6.7%)
Guilty	113/285 (39.6%)	54/164 (32.9%)	13/54 (24.1%)	46/145 (31.7%)	9/47 (19.1%)	1/15 (6.7%)
Eye-opener	53/285 (18.6%)	28/164 (17.1%)	10/54 (18.5%)	17/145 (11.7%)	2/47 (4.3%)	1/15 (6.7%)
Clin.signif	103/285 (36.1%)	53/164 (32.3%)	13/54 (24.1%)	45/145 (31.0%)	6/47 (12.8%)	1/15 (6.7%)

	Female caregivers			Male caregivers		
CAGE	Baseline	Stepdown	6 months	Baseline	Stepdown	6 months
Alcohol						
Cut-down	15/65 (23.1%)	3/35 (8.6%)	0/9 (0.0%)	11/35 (31.4%)	1/6 (16.7%)	0/4 (0.0%)
Annoyed	12/65 (18.5%)	3/35 (8.6%)	0/9 (0.0%)	7/35 (20.0%)	0/6 (0.0%)	0/4 (0.0%)
Guilty	17/65 (26.2%)	3/35 (8.6%)	0/9 (0.0%)	10/35 (28.6%)	2/6 (33.3%)	1/4 (25.0%)
Eye-opener	5/65 (7.7%)	1/35 (2.9%)	0/9 (0.0%)	7/35 (20.0%)	0/6 (0.0%)	0/4 (0.0%)
Clin.signif	15/65 (23.1%)	3/35 (8.6%)	0/9 (0.0%)	11/35 (31.4%)	1/6 (16.7%)	0/4 (0.0%)
Drugs						
Cut-down	28/65 (43.1%)	10/35 (28.6%)	1/9 (11.1%)	21/35 (60.0%)	1/6 (16.7%)	0/4 (0.0%)
Annoyed	15/65 (23.1%)	8/35 (22.9%)	1/9 (11.1%)	10/35 (28.6%)	1/6 (16.7%)	0/4 (0.0%)
Guilty	31/65 (47.7%)	10/35 (28.6%)	2/9 (22.2%)	18/35 (51.4%)	1/6 (16.7%)	0/4 (0.0%)
Eye-opener	11/65 (16.9%)	4/35 (11.4%)	2/9 (22.2%)	8/35 (22.9%)	0/6 (0.0%)	0/4 (0.0%)
Clin.signif	26/65 (40.0%)	12/35 (34.3%)	2/9 (22.2%)	18/35 (51.4%)	1/6 (16.7%)	0/4 (0.0%)

	Female caregivers			Male caregivers		
CAGE	Baseline	Stepdown	6 months	Baseline	Stepdown	6 months
Alcohol						
Cut-down	7/51 (13.7%)	2/32 (6.2%)	0/8 (0.0%)	6/27 (22.2%)	2/11 (18.2%)	0/3 (0.0%)
Annoyed	3/51 (5.9%)	3/32 (9.4%)	0/8 (0.0%)	4/27 (14.8%)	1/11 (9.1%)	0/3 (0.0%)
Guilty	7/51 (13.7%)	5/32 (15.6%)	0/8 (0.0%)	8/27 (29.6%)	2/11 (18.2%)	0/3 (0.0%)
Eye-opener	3/51 (5.9%)	2/32 (6.2%)	0/8 (0.0%)	4/27 (14.8%)	0/11 (0.0%)	0/3 (0.0%)
Clin.signif	6/51 (11.8%)	4/32 (12.5%)	0/8 (0.0%)	7/27 (25.9%)	2/11 (18.2%)	0/3 (0.0%)
Drugs						
Cut-down	19/51 (37.3%)	6/32 (18.8%)	0/8 (0.0%)	10/27 (37.0%)	2/11 (18.2%)	0/3 (0.0%)
Annoyed	11/51 (21.6%)	7/32 (21.9%)	0/8 (0.0%)	5/27 (18.5%)	1/11 (9.1%)	0/3 (0.0%)
Guilty	18/51 (35.3%)	10/32 (31.2%)	0/8 (0.0%)	12/27 (44.4%)	2/11 (18.2%)	0/3 (0.0%)
Eye-opener	10/51 (19.6%)	6/32 (18.8%)	0/8 (0.0%)	5/27 (18.5%)	1/11 (9.1%)	0/3 (0.0%)
Clin.signif	18/51 (35.3%)	8/32 (25.0%)	0/8 (0.0%)	10/27 (37.0%)	2/11 (18.2%)	0/3 (0.0%)

	Female caregivers			Male caregivers		
CAGE	Baseline	Stepdown	6 months	Baseline	Stepdown	6 months
Alcohol						
Cut-down	21/116 (18.1%)	11/74 (14.9%)	3/22 (13.6%)	20/61 (32.8%)	3/23 (13.0%)	2/6 (33.3%)
Annoyed	10/116 (8.6%)	11/74 (14.9%)	2/22 (9.1%)	11/61 (18.0%)	2/23 (8.7%)	0/6 (0.0%)
Guilty	10/116 (8.6%)	9/74 (12.2%)	3/22 (13.6%)	13/61 (21.3%)	3/23 (13.0%)	1/6 (16.7%)
Eye-opener	10/116 (8.6%)	5/74 (6.8%)	2/22 (9.1%)	8/61 (13.1%)	1/23 (4.3%)	0/6 (0.0%)
Clin.signif	16/116 (13.8%)	9/74 (12.2%)	2/22 (9.1%)	15/61 (24.6%)	3/23 (13.0%)	1/6 (16.7%)
Drugs						
Cut-down	39/116 (33.6%)	22/74 (29.7%)	7/22 (31.8%)	15/61 (24.6%)	2/23 (8.7%)	0/6 (0.0%)
Annoyed	26/116 (22.4%)	16/74 (21.6%)	7/22 (31.8%)	7/61 (11.5%)	1/23 (4.3%)	0/6 (0.0%)
Guilty	44/116 (37.9%)	24/74 (32.4%)	7/22 (31.8%)	11/61 (18.0%)	4/23 (17.4%)	0/6 (0.0%)
Eye-opener	19/116 (16.4%)	13/74 (17.6%)	5/22 (22.7%)	2/61 (3.3%)	0/23 (0.0%)	0/6 (0.0%)
Clin.signif	40/116 (34.5%)	23/74 (31.1%)	7/22 (31.8%)	12/61 (19.7%)	1/23 (4.3%)	0/6 (0.0%)

	Female caregivers			Male caregivers		
CAGE	Baseline	Stepdown	6 months	Baseline	Stepdown	6 months
Alcohol						
Cut-down	8/43 (18.6%)	2/19 (10.5%)	1/13 (7.7%)	5/15 (33.3%)	0/6 (0.0%)	0/1 (0.0%)
Annoyed	7/43 (16.3%)	2/19 (10.5%)	1/13 (7.7%)	2/15 (13.3%)	0/6 (0.0%)	0/1 (0.0%)
Guilty	8/43 (18.6%)	2/19 (10.5%)	1/13 (7.7%)	2/15 (13.3%)	0/6 (0.0%)	0/1 (0.0%)
Eye-opener	5/43 (11.6%)	0/19 (0.0%)	0/13 (0.0%)	1/15 (6.7%)	0/6 (0.0%)	0/1 (0.0%)
Clin.signif	7/43 (16.3%)	2/19 (10.5%)	1/13 (7.7%)	2/15 (13.3%)	0/6 (0.0%)	0/1 (0.0%)
Drugs						
Cut-down	18/43 (41.9%)	7/19 (36.8%)	3/13 (23.1%)	5/15 (33.3%)	1/6 (16.7%)	1/1 (100.0%)
Annoyed	12/43 (27.9%)	6/19 (31.6%)	3/13 (23.1%)	2/15 (13.3%)	1/6 (16.7%)	1/1 (100.0%)
Guilty	18/43 (41.9%)	9/19 (47.4%)	3/13 (23.1%)	4/15 (26.7%)	1/6 (16.7%)	1/1 (100.0%)
Eye-opener	11/43 (25.6%)	5/19 (26.3%)	2/13 (15.4%)	2/15 (13.3%)	1/6 (16.7%)	1/1 (100.0%)
Clin.signif	17/43 (39.5%)	9/19 (47.4%)	3/13 (23.1%)	5/15 (33.3%)	1/6 (16.7%)	1/1 (100.0%)

Question 4: Does ISS result in better child functioning outcomes as measured on the OK-TASCC child health measure compared to the SAU condition?

The key takeaways for this question include:

- There is little difference between SAU-assigned and ISS-Assigned conditions, whether using raw measurement scores, or the probability of a measurement score meeting or exceeding a critical cutoff value.
- There is little difference between those who received SAU and those who received ISS; One Treatment on the Treated model found that those who received ISS show higher measures of symptoms. Those who were assigned to ISS but did not receive it often show higher scores on measures of symptoms and lower scores on our measure of development.
- The only consistently significant main effect on any of the four measures was time. The influence of time seems to be nonlinear, and often interacts with Treatment Condition.
- ANOVA tests show that different regions often have significantly different distributions of PSC, BPSC, and Milestones Scores. Mixed Effects models specifically found that those who receive ISS in Region 2 show significantly lower BPSC scores than those who received ISS in Region 3.

Methods

The data used to address this question come from a combination of the Kids and BEST data systems and were collected from both ISS and SAU conditions. These data were intended to be collected at 1-month intervals, but in practice the interval was variable. This analysis has a sample size of 2,286 unique measurement occasions (after filtering down to the first twelve months of measurements for each family). The analytical strategy used is a series of mixed-effects models using ISS Group Assignment (or Group Received), PREMISS Score, Time, and a random intercept of child ID to model scores on the OK-TASCC Measures.

Almost half of all families in the Waiver project have received verified screening (48.08%). The Child Behavioral Health screener which, until March of 2017, was stored in the BEST system separate from the DHS KIDS system contains an ID variable not equal to the KIDS client ID. All of the children and families in the Waiver project have child and family ID variables from KIDS, but not all of them can be linked to the BEST screener data because (1) there is imperfect compliance in administering the screeners, and (2) there is imperfect record keeping that breaks the linking process from Waiver to KIDS to BEST. Nonetheless, there is a high match rate between BEST and KIDS. Thus, a lack of compliance in administering the screener likely explains the majority of missing screeners for families in the Waiver project.

We also hypothesized that the special needs of the child influenced whether they have a screener score on record. We fit a mixed-effects logistic model regressing the presence of a screener on whether the child had specific special needs. We found that those children with Physical or Visual/Auditory impairments were significantly less likely to have an associated CBHS screener than those without. Those with an emotional disability on record are significantly more likely to have a screener on record. The full results are displayed below.

TABLE **143** - PREDICTING PRESENCE OF **CBHS** SCREENER

	Estimate	Std. Error	z value	p-value
Intercept	9.07	0.23	40.13	< 0.01
Mental Impairment	-1.01	1.52	-0.66	0.51
Visual/Audio Impairment	-20.29	2.96	-6.85	< 0.01
Physical Impairment	-21.63	4.19	-5.16	< 0.01
Emotional Impairment	2.14	0.79	2.71	<0.01
Medical/Other Impairment	0.03	0.71	0.04	0.97

As noted previously, since March 1, 2017 the data have been kept within the KIDS system. For purposes of this report, data were pulled in from both the BEST and KIDS data systems. Whenever this resulted in duplicate rows of screening information (as determined by Victim ID and Contact Date), the row that contained the Total PSC score was kept. In instances where the Total PSC Score was present (or absent) in rows from both data systems for the same child and the same day, the row from the KIDS data set received priority. The result is that 48.08% of families in the Waiver project were connected to verified screening data. It is these clients that are the subject of this chapter.

The Child Behavioral Health Screener is not a single measurement instrument. Rather it is a set of 11 partially overlapping instruments that vary in the domain measured and the appropriate age range for administration. Plainly, the indicators of behavioral health for a 5-month-old are often quite distinct from those of a 5-year-old. The diversity of screening instruments is a consequence of that. However, for all children between 4 and 12 years of age a single measure is used: the Pediatric Symptom Checklist (PSC). A little over half of the total sample (57.52%) is in this age range. For the present purposes, primarily the PSC results are reported here. Broadly, the other measures are (1) the Baby Pediatric Symptom Checklist (BPSC, ages 0 to 17 months), (2) the Preschool Pediatric Symptom Checklist (PPSC, ages 18 to 48 months), and a measure of child development (Milestones, ages 0 to 48 months). Each of these measures consists of different versions depending on the age of the child being administered the measure. Collectively, these three measures compose the Survey of Well-Being of Young Children (SWYC).

It is also worth noting that for this sample, cases were filtered out where a court intervened with a case prior to that case being assigned to Intensive Safety Services (ISS) or Services as Usual (SAU) by means of the PREMISS algorithm. The reasoning here is that children with court involvement likely reflect greater risks for safety and possibly more severe behavioral problems. We do not want to introduce this bias into our SAU condition.

We also narrowed our focus down to the first year of measurements of each child. The reasoning is that a family can only receive ISS for six months, while there is no analogous limit to the amount of time a family can receive SAU. To create a more equitable comparison, for each family (and for each measurement instrument), we filtered down to the first measurement occasion for each family and all subsequent measurement occasions within the next 365 days. This time frame will cover the maximum

of 6 months that a family can receive ISS plus the subsequent 6 months. These measurements will be compared to the first year of measurements for all those in the SAU condition.

Region 1 In region 1, the sample consisted of 112 children (64 in SAU, 48 in ISS) from 98 families (54 in SAU, 44 in ISS); that is, 1.14 children per family. The mean number of screeners administered per child is 4.5 (SD = 3.26 Range = 1 to 13). The mean number of screeners per family is 5.14 (SD = 4.64, Range = 1 to 27). The sample was 34% male and 54% female (the remainder of unspecified gender). Child ages ranged from 4 years to 12 years (mean = 7.8, median = 7.44, SD = 2.54).

Region 2 In region 2, the sample consisted of 72 children (42 in SAU, 30 in ISS) from 72 families (42 in SAU, 30 in ISS); that is, 1 child per family. The mean number of screeners administered per child is 4.76 (SD = 3.95 Range = 1 to 15). The mean number of screeners per family is 4.76 (SD = 3.95, Range = 1 to 15). The sample was 60% male and 40% female (the remainder of unspecified gender). Child ages ranged from 4.01 years to 11.9 years (mean = 7.52, median = 7.21, SD = 2.52).

Region 3 In region 3, the sample consisted of 218 children (29 in SAU, 189 in ISS) from 143 families (23 in SAU, 120 in ISS); that is, 1.52 children per family. The mean number of screeners administered per child is 5.6 (SD = 3.86 Range = 1 to 13). The mean number of screeners per family is 8.53 (SD = 7.83, Range = 1 to 43). The sample was 39% male and 39% female (the remainder of unspecified gender). Child ages ranged from 4 years to 11.98 years (mean = 7.1, median = 6.86, SD = 2.37).

Region 4 In region 4, the sample consisted of 43 children (24 in SAU, 19 in ISS) from 43 families (24 in SAU, 19 in ISS); that is, 1 child per family. The mean number of screeners administered per child is 3.88 (SD = 2.64 Range = 1 to 10). The mean number of screeners per family is 3.88 (SD = 2.64, Range = 1 to 10). The sample was 51% male and 49% female (the remainder of unspecified gender). Child ages ranged from 4.04 years to 11.91 years (mean = 7.28, median =7 .02, SD = 2.38).

Region 5 In region 5, the sample consisted of 75 children (41 in SAU, 34 in ISS) from 72 families (39 in SAU, 33 in ISS); that is, 1.04 children per family. The mean number of screeners administered per child is 4.15 (SD = 2.9 Range = 1 to 13). The mean number of screeners per family is 4.32 (SD = 2.93, Range = 1 to 13). The sample was 51% male and 43% female (the remainder of unspecified gender). Child ages ranged from 4.02 years to 11.94 years (mean = 7.6, median = 7.45, SD = 2.31).

Figure 44 below shows the total number of measurement periods per child. That is, it shows the number of children with exactly that number of occasions. For example, 58 children have exactly two occasions of measurement. Those 58 children appear exclusively in the bin corresponding to 2 measurement occasions.

FIGURE 44 THE NUMBER OF CHILDREN WITH EACH NUMBER OF MEASUREMENTS
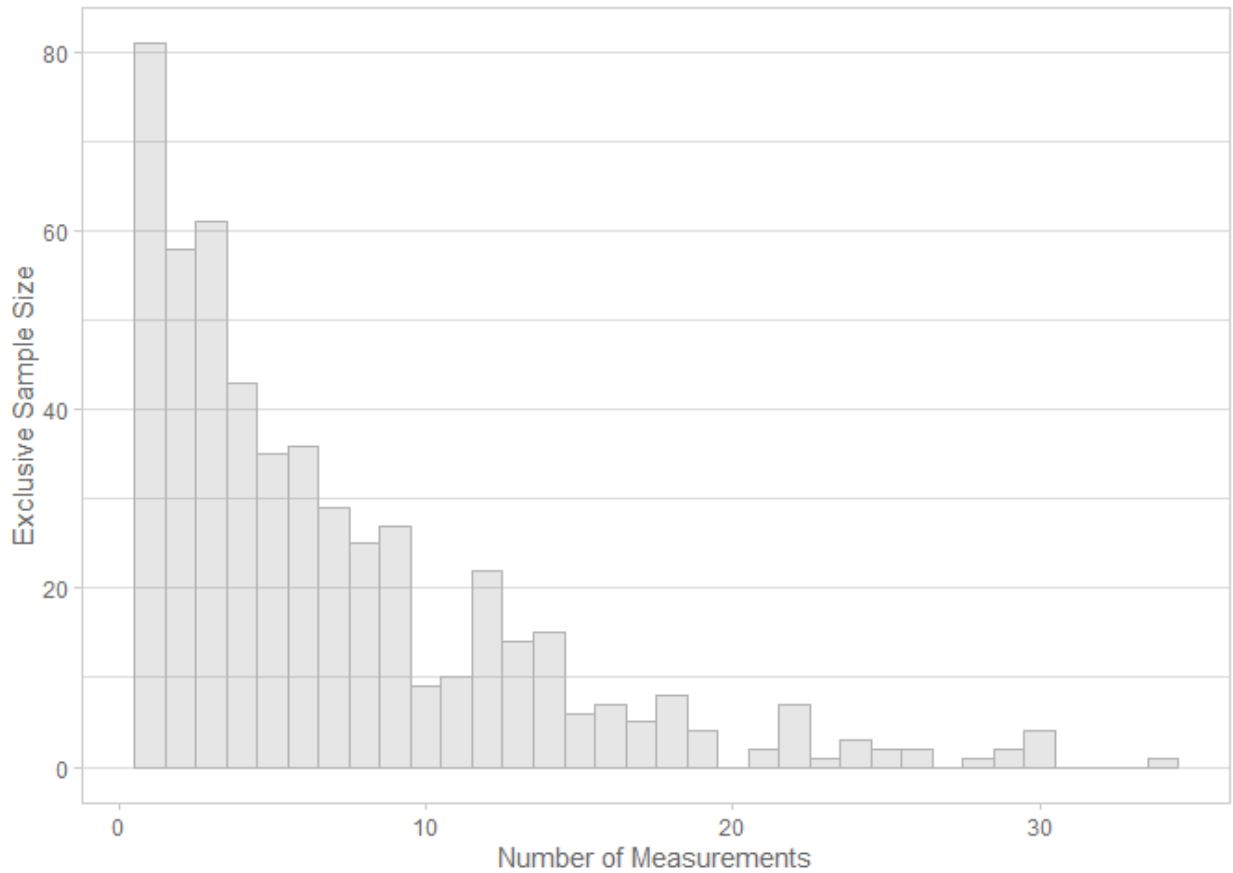
The next figure shows the *cumulative* number of measurement occasions for each child. That is, it shows the number of children with at least that number of occasions. For example, 379 children have two or more measurements. The 58 children who have exactly two 2 measurements will appear in the first two bins below, corresponding to having at least one measurement and having at least two measurements respectively.

The next set of results shows how the PSC score changes over time. The total PSC score is the summed score across 20 items, each one of which can take the values 0 ("never" or "not at all"), 1 ("sometimes" or "somewhat"), or 2 ("very much", or "often"). The original version of the PSC has three subscales: five items assessing attention problems (e.g., has trouble paying attention), five items assessing internalizing behavior problems (e.g., worries a lot), and 7 items assessing externalizing behavior problems (e.g., teases others). Specific to this population, three additional items were added to assess any trauma symptoms (e.g., gets very upset when reminded of traumatic events). Thus, the PSC as administered in this sample has four subscales that relate to a common overall factor. The total score is the sum of all the items with higher scores indicating more problems. The table below shows how the PSC scores change over time. A "Wave" is defined for each child relative to their own occasions of measurement. Wave 1 is always the child's first measurement occasion, no matter when that occurred. Ideally, Wave 2 would be 30 days after Wave 1; however, the actual amount of time varies from person to person with a mean of 39.25 days (median = 31 days, SD=30.13 days, range 1 to 362 days). Those in the ISS Received and ISS Not Workable condition appear to have higher PSC scores, and to maintain them longer over time.

TABLE 144 MEAN AND STANDARD DEVIATION OF THE PSC TOTAL SCORES AT EACH WAVE

	Mean (SD) PSC Total Score		
Wave	SAU	ISS Received	ISS Not Workable
1	9.6 (7.9)	9.7 (7.3)	9.9 (7.0)
2	9.1 (7.8)	8.9 (7.0)	9.2 (6.4)
3	9.0 (8.4)	9.3 (6.7)	9.4 (5.9)
4	8.3 (7.5)	9.6 (6.8)	7.7 (6.0)
5	8.4 (7.1)	9.0 (6.7)	7.7 (6.5)
6	7.4 (6.4)	8.8 (7.3)	8.7 (7.3)
7	7.8 (6.4)	7.6 (6.7)	8.6 (6.7)
8	6.5 (5.3)	7.7 (6.4)	10.0 (8.4)
9	6.8 (5.1)	8.9 (8.0)	10.2 (8.3)
10	7.7 (5.6)	5.9 (4.6)	9.7 (7.2)
11	6.7 (5.3)	8.6 (9.3)	10.6 (8.8)
12	6.7 (6.1)	4.8 (4.9)	10.4 (8.8)
13	5.8 (2.9)	6.8 (6.9)	10.1 (8.5)
14	6.3 (5.5)	5.2 (3.6)	6.7 (6.3)
15	9.5 (7.9)	6.8 (4.6)	6.8 (5.5)
16	10.0 (14.1)	9.0 (4.7)	7.0 (5.4)
17	7.0 (9.9)	8.0 (3.6)	5.3 (4.6)
18	5.0 (7.1)	10.5 (3.9)	9.0 (8.0)
19	0.0 (NA)	6.0 (6.0)	8.3 (6.0)
20	0.0 (NA)	8.0 (8.7)	12.0 (NA)
21	0.0 (NA)	4.7 (2.9)	11.0 (NA)
22	NA	5.0 (4.6)	NA
23	NA	7.3 (4.0)	NA
24	NA	6.0 (0.0)	NA
25	NA	6.0 (NA)	NA

The figure below shows the PSC total scores for the sample in terms of the number of days since their first screener. The thin lines are the PSC scores for individual children; the lines are color-coded according to the key in the upper right corner of the Figure (purple = ISS Received, yellow = SAU, blue = ISS Not Workable). The thicker lines of the same colors are LOESS lines that are smooth local regression lines that can be thought of like a running mean. The shaded grey regions around the LOESS lines are the confidence intervals around the LOESS lines. The three lines seem to show very similar trajectories for approximately the first 150 days, at which point they seem to diverge. Those in the ISS Not Workable group seem to show the highest scores (i.e. most pediatric symptoms) after that point, suggesting that the individuals who are selected for ISS and then removed represent a unique population. Those in ISS and SAU seem to show similar declines in scores (i.e. fewer pediatric symptoms) over time. Those in the SAU condition may be showing the fewest symptoms.

Table 145 that gives the number of measurement occasions within each condition.

FIGURE 46 - PSC TOTAL SCORES FOR SAMPLE VS DAYS SINCE FIRST SCREENER IN ACTUAL GROUP
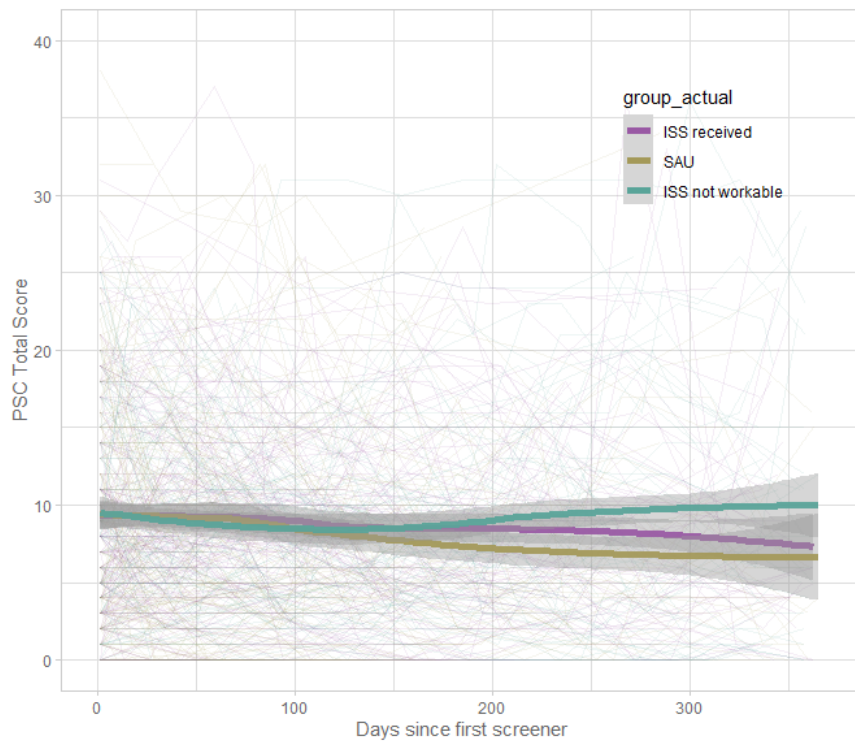
TABLE 145 COUNT OF MEASUREMENTS WITHIN ACTUAL GROUP CONDITION

Treatment Received	Count
ISS received	874
SAU	865
ISS not workable	806

156

The figure below presents the same data as the figure above but combines the two ISS groups into a single *assigned* group. These lines provide a view of the impact of Intent-to-Treat. Here, we see a similar pattern as above. Both groups show a decline over time, with little distinction between groups until day 150. After that point, those assigned to the SAU condition seem to shower lower PSC scores.

FIGURE 47 - PSC TOTAL SCORES FOR SAMPLE VS DAYS SINCE FIRST SCREENER IN ASSIGNED GROUP
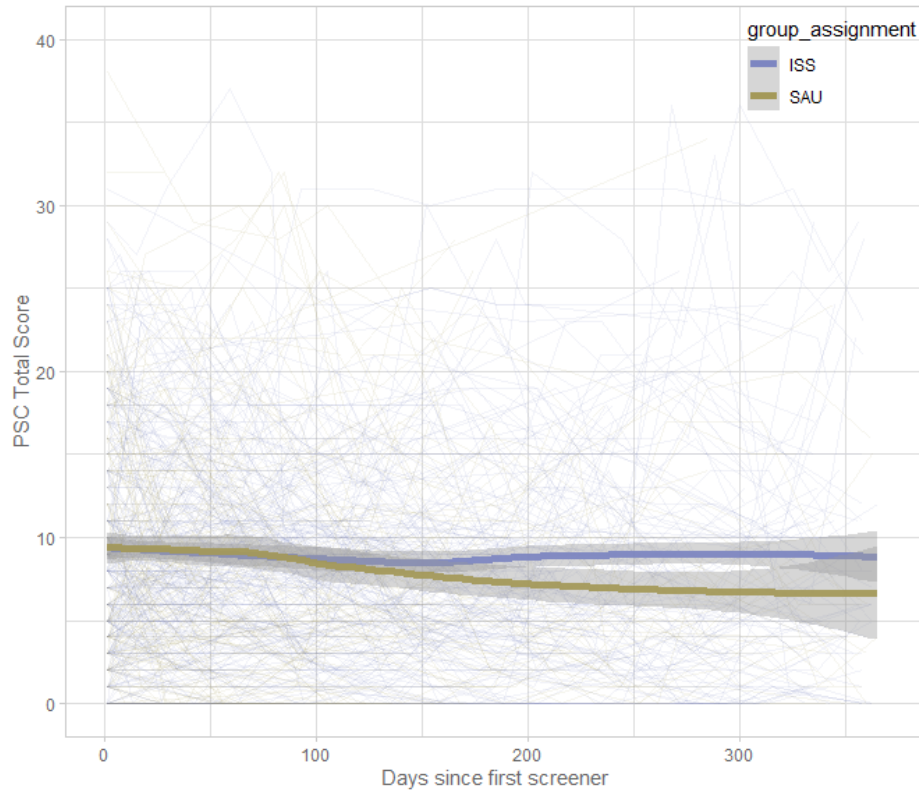
TABLE 146 COUNT OF MEASUREMENTS WITHIN ASSIGNED GROUP CONDITION

Condition Assigned	Count
ISS	1680
SAU	865

Linear mixed effects models echo the same findings as the above Figure. A mixed effects model was run using R and the lme4 package with the PSC total score as the outcome and the child ID variable providing the nesting structure. The model allowed for fixed effects of time, treatment group, and their interaction. The model also had a random effect of child ID. We also included the child's PREMISS score as a covariate. This is a measure of the level of safety in a given family situation and is instrumental in identifying children who are eligible for ISS. As such, it is important to account for child's PREMISS score in the model. Higher scores on the PREMISS indicate a safer home.

The Table below details the results of this model. These results suggest there is a non-zero PSC score at the first measurement with a mean of 8.11 in the ISS group and a mean of 8.11 + 0.61 = 8.72 in the SAU assigned group, but this difference in not quite significant (t-value 0.9). The Fixed Effect for the Intercept

157

term gives the mean of the SAU group at the first time point. There is a large amount of variability in this starting point (SD=5.43). This variability even exceeds the residual variability (SD=4.68). The time effect implies a general decline in PSC scores (lower scores are better, indicating fewer problems); this change does seem significant (estimate = -1.34, t = -2.78). Those in the SAU condition seem to show a greater decline across time that those assigned to ISS (estimate = -3.91, t = -4.38).

TABLE 147 RESULTS FROM AN INTENT-TO-TREAT LINEAR MIXED EFFECTS MODEL ON PSC

	Estimate	Std. Error	t value	Random SD
Intercept	8.11	0.75	10.79	5.4
PREMISS	-0.78	0.44	-1.78	NA
Days Since 1st Measurement	-1.34	0.48	-2.78	NA
SAU	0.61	0.68	0.9	NA
Days Since 1st Measurement*SAU	-3.91	0.89	-4.38	NA
Residual	NA	NA	NA	4.7

We were also interested in exploring the influence of regional effects. To start, we ran an ANOVA to test for the significance of the differences in the distribution of PSC scores across the 5 regions, and within the SAU and ISS-Assigned conditions. The results are displayed in the graphs and tables below. It seems that PSC scores do differ between assignment groups and between regions.

TABLE 148 PSC SCORES ACROSS REGIONS AND ISS VERSUS SAU CONDITIONS

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Assigned Group	1	150.16	150.16	3.01	0.08
Region	4	859.85	214.96	4.31	<0.01
Assigned Group *Region	4	689.65	172.41	3.46	0.01
Residuals	2535	126451.5	49.88	NA	NA
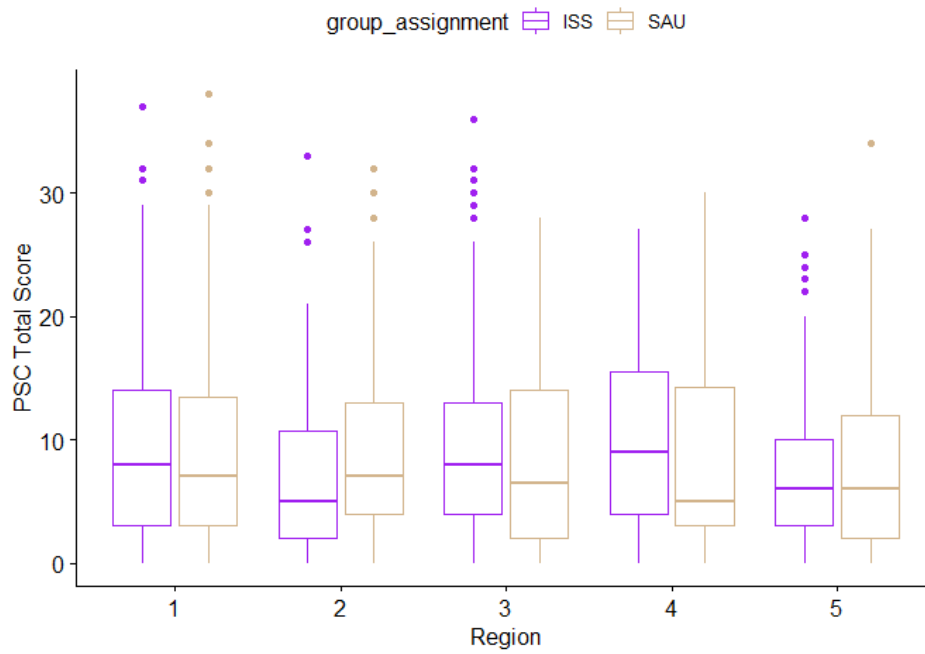
We followed this up by doing the same thing, but dividing the ISS assigned into its two component conditions: those who were assigned to ISS and actually received ISS, and those assigned to the ISS condition, but ISS was found to be unworkable. The results are shown in the graph and table below. Again, it seems that PSC scores do show different distributions across regions and assignment conditions.

FIGURE **49** PSC SCORES ACROSS REGIONS AND ISS-RECEIVED, ISS-NOT WORKABLE AND SAU CONDITIONS

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Actual Group	2	188.89	94.44	1.90	0.15
Region	4	842.67	210.67	4.23	0.00
Actual Group * Region	8	1242.36	155.30	3.12	0.00
Residuals	2530	125877.27	49.75	NA	NA

FIGURE 50 PSC SCORES ACROSS REGIONS AND ISS-RECEIVED, ISS-NOT WORKABLE AND SAU CONDITIONS
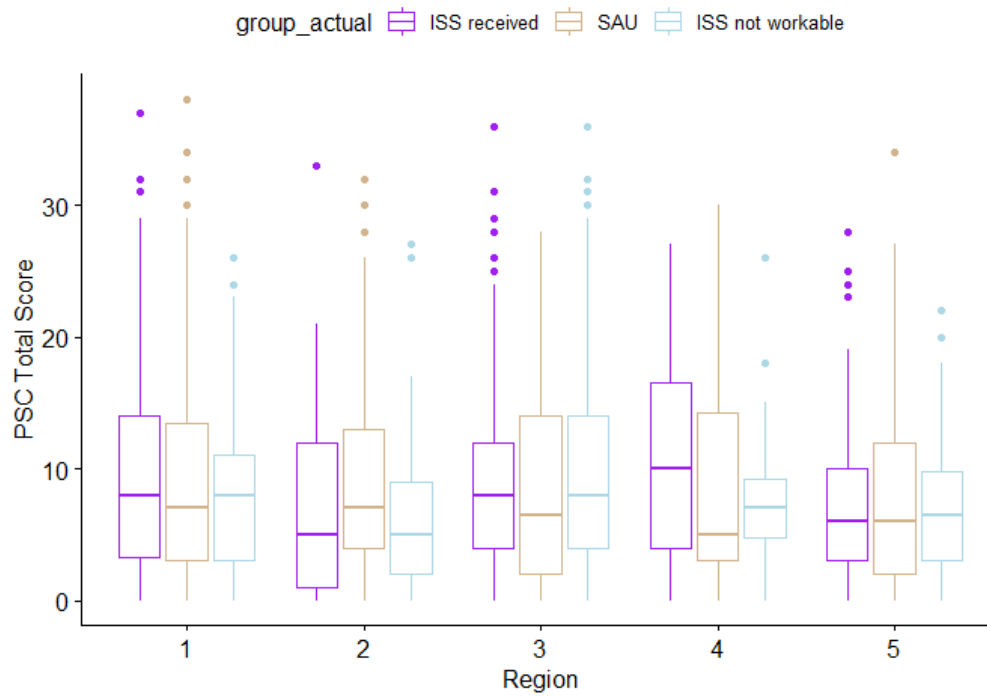
On the basis of these results, we fit a mixed effects model regressing a client's total PSC score onto a new variable that indicates the combination of a client's ISS Assignment condition and their region. This model also retains the random effects of PREMISS Score and the time since the first occasion of PSC Measurement, and the fixed effect of client ID. The idea is to fit a model that gives us some detail about the influence of Intent-to-Treat in a unique region.

The following table shows the results of this model. Note that for purposes of this model, the referent group is those individuals from Region 3 (Oklahoma County) who were assigned to ISS. The assignment-by-region categories do not seem to show significant change from the referent category, although those assigned ISS in Region 2 and Region 5 come close. There is no clear pattern as to whether those assigned to ISS or SAU show lower or higher scores within a region. The only variable with a significant influence seems to be time since first measurement, with PSC scores tending to decline as time passes (estimate = -2.53, t = -6.16).

TABLE 149 REGRESSING PSC ONTO CONDITION ASSIGNMENT-BY-REGION VARIABLES

	Estimate	Std. Error	t value	Random SD
Intercept	9.57	0.88	10.87	5.39
Region 1 – ISS Assigned	-0.8	1.07	-0.74	NA
Region 1 – SAU Assigned	-0.98	1	-0.98	NA
Region 2 – ISS Assigned	-2.66	1.4	-1.9	NA
Region 2 – SAU Assigned	-0.54	1.17	-0.46	NA
Region 3 – SAU Assigned	-1.69	1.42	-1.19	NA
Region 4 – ISS Assigned	-0.18	1.77	-0.1	NA
Region 4 – SAU Assigned	-1.58	1.55	-1.02	NA
Region 5 – ISS Assigned	-3.03	1.56	-1.94	NA
Region 5 – SAU Assigned	-1.47	1.22	-1.2	NA
PREMISS	-0.54	0.46	-1.17	NA
Days Since 1st Measurement	-2.53	0.41	-6.16	NA
Residual	NA	NA	NA	4.72

The following table shows the results of the same model, but examining actual ISS condition instead of assigned condition. Again, none of the Treatment-Received-by-Region categories show significant difference. No Treatment-Received condition seems to be associated with higher or lower PSC scores within a region. The only significant effect is that of time, such that as time passes, PSC scores tend to decrease (estimate = -2.53, t = -6.16).

TABLE 150 REGRESSING PSC ONTO TREATMENT RECEIVED-BY-REGION

	Estimate	Std. Error	t value	Random SD
Intercept	9.14	1	9.09	5.4
Region 1 – ISS Not Workable	-1.1	1.5	-0.7	NA
Region 1 – ISS Received	0.34	1.5	0.22	NA
Region 1 - SAU	-0.6	1.1	-0.5	NA
Region 2 – ISS Not Workable	-4.0	2.4	-1.7	NA
Region 2 – ISS Received	-1.4	1.8	-0.8	NA
Region 2 - SAU	-0.1	1.3	-0.1	NA
Region 3 – ISS Not Workable	0.84	1.1	0.79	NA
Region 3 - SAU	-1.3	1.5	-0.8	NA
Region 4 – ISS Not Workable	-2.9	3.1	-0.9	NA
Region 4 – ISS Received	1.63	2.2	0.75	NA
Region 4 - SAU	-1.2	1.6	-0.7	NA
Region 5 – ISS Not Workable	-2.6	2.5	-1.0	NA
Region 5 – ISS Received	-2.6	2.0	-1.3	NA
Region 5 - SAU	-1.0	1.3	-0.8	NA
PREMISS	-0.5	0.5	-1.2	NA
Days Since 1st Measurement	-2.5	0.4	-6.2	NA
Residual	NA	NA	NA	4.7

Based on examination of the LOESS plots shown earlier, we may also want to consider nonlinear patterns of change for the treatment trajectories. One simple family of nonlinear growth trajectories are piecewise linear curves. Piecewise linear curves have linear patterns of change everywhere but allow one region to have a different linear slope from another. The critical feature of piecewise linear curves is selection of the change points. Based on inspection of the graphs, we selected two change points: one at 45 days from the first screener, and one at 90 days from the first screener. Thus, we have a slope for days 0 through 45, a different slope for days 45 through 90, and a third slope for days 90 through their last measurement.

Based on the results of the Intent-to-Treat piecewise model below, we do see some indication that during the 45-to-90 day period, both groups show a decrease in PSC total relative to the first time interval (estimate = -17.97, t = -3.64); however, this change is not significantly different for the ISS-assigned and SAU-assigned conditions (estimate = 11.41, t = 1.49). We do, however, see that those

assigned to SAU do show a significantly more pronounced decline over time after the 90-day mark (estimate = -7.92, t = -5.22).

TABLE 151 - RESULTS FROM AN INTENT-TO-TREAT PIECEWISE LINEAR MIXED EFFECTS MODEL ON PSC

	Estimate	Std. Error	t value	Random SD
Intercept	8.19	0.8	10.5	5.5
PREMISS	-0.79	0.4	-1.8	NA
Time Since 1st Measurement	3.25	5	0.65	NA
Time Between Days 45 and 90	-18	4.9	-3.64	NA
Time Since Day 90	1.59	0.8	1.89	NA
SAU	0.17	0.7	0.23	NA
Time Since 1st Measurement * SAU	-4.68	7.6	-0.61	NA
Time Between Days 45 and 90 * SAU	11.41	7.6	1.49	NA
Time Since Day 90 * SAU	-7.92	1.5	-5.22	NA
Residual	NA	NA	NA	4.7

This trend is also evident in the Piecewise Model of Actual Treatment group. Among those in the ISS condition, we see a significant decrease in slope of PSC scores across the 45-90 day interval (estimate = -21.48, t = -3.02). We again see that those in the SAU condition show a greater decline across time after the 90-day mark, relative to those who received ISS (estimate = -6.14, t = -3.31).

163

	Estimate	Std. Error	t value	Random SD
Intercept	8.3	0.88	9.47	5.5
PREMISS	-0.81	0.44	-1.84	NA
Time Since Measurement 1	7.08	6.86	1.03	NA
Time Between Day 45 and 90	-21.48	7.11	-3.02	NA
Time Since Day 90	-0.19	1.36	-0.14	NA
SAU	0.04	0.86	0.04	NA
ISS Not Workable	-0.29	0.96	-0.3	NA
Time Since 1st Measurement * SAU	-8.53	8.94	-0.95	NA
Time Since 1st Measurement * ISS Not Workable	-7.42	10.05	-0.74	NA
Time Between Day 45 and 90 * SAU	14.92	9.19	1.62	NA
Time Between Day 45 and 90 * ISS Not Workable	7.51	9.89	0.76	NA
Time Since Day 90 * SAU	-6.14	1.86	-3.31	NA
Time Since Day 90 * ISS Not Workable	2.72	1.73	1.57	NA
Residual	NA	NA	NA	4.7

The PSC items also all have unique cutoff values, such that getting a score at or above a certain value represents a case of clinical concern. As the cutoffs are all unique to specific subscales within the PSC, we fit three new models; each is a logistic regression predicting the odds of reaching or exceeding a cutoff value for a specific subscale. Each model uses ISS Assignment condition, an interaction between assignment condition and time since first measurement, a random intercept of child ID to account for repeated measures, and the family's PREMISS score.

For the Attention subscale, those assigned to SAU are not any more or less likely to cross the threshold designating a clinically significant case, (estimate = 0.15, p = 0.7).

There is, however, a significant interaction between assignment to the SAU condition and time (estimate = -1.39, p = 0.04). This suggests that, as time increases, the probability of getting a score on the PSC-Attention subscale that crosses the critical threshold decreases for those in the SAU-Assigned condition at a greater rate than for those in the ISS-Assigned condition.

TABLE 153 LOGISTIC MODEL OF PSC-ATTENTION SCORE ABOVE CUTOFF

	Estimate	Std. Error	t value	Pr(>|z|)	Random SD
Intercept	-3	0.48	-6.3	<0.01	2.66
SAU	0.15	0.39	0.38	0.7	NA
Days Since 1st Measurement	0.1	0.35	0.28	0.78	NA
PREMISS	-0.3	0.25	-1.1	0.26	NA
SAU * Days Since 1st Measurement	-1.4	0.68	-2	0.04	NA
Residual	NA	NA	NA	NA	1

For the Internalization subscale of the PSC measurement, those assigned to SAU are not any more or less likely to cross the threshold designating a clinically significant case, (estimate = -0.02, p = 0.95).

There is, however, a significant interaction between assignment to the SAU condition and time (estimate = -2.18, p = 0). This suggests that, as time increases, the probability of getting a score on the PSC-Internalization subscale decreases for those in the SAU-Assigned condition at a greater rate than for those in the ISS-Assigned condition.

TABLE 154 LOGISTIC MODEL OF PSC-INTERNALIZATION SCORE ABOVE CUTOFF

	Estimate	Std. Error	t value	Pr(>|z|)	Random SD
Intercept	-2.6	0.39	-6.62	<0.01	2
SAU	-0.02	0.33	-0.07	0.95	NA
Days Since 1st Measurement	-0.12	0.35	-0.34	0.73	NA
PREMISS	-0.07	0.21	-0.33	0.74	NA
SAU * Days Since 1st Measurement	-2.18	0.71	-3.07	<0.01	NA
Residual	NA	NA	NA	NA	1

For the Externalization subscale, those in assigned to SAU are not any more or less likely to cross the threshold designating a clinically significant case, (estimate = 0, p = 0.99).

There is, however, a significant interaction between assignment to the SAU condition and time (estimate = -1.5, p = 0.03). This suggests that, as time increases, the probability of getting a score on the PSC-Externalization subscale decreases for those in the SAU-Assigned condition at a greater rate than for those in the ISS-Assigned condition.

TABLE 155 LOGISTIC MODEL OF PSC-EXTERNALIZATION SCORE ABOVE CUTOFF

	Estimate	Std. Error	t value	Pr(>|z|)	Random SD
Intercept	-2.89	0.44	-6.55	<0.01	2.4
SAU	0.00	0.36	-0.01	0.99	NA
Days Since 1st Measurement	-0.38	0.32	-1.18	0.24	NA
PREMISS	-0.56	0.23	-2.41	0.02	NA
SAU * Days Since 1st Measurement	-1.5	0.69	-2.16	0.03	NA
Residual	NA	NA	NA	NA	1

The results reported above are mostly for Intent-to-Treat analyses. That is, in these analyses, all children that were assigned the ISS group were included in the ISS condition, regardless of whether or not they actually received ISS. However, as is common with many randomized controlled trials, not all those assigned to the experimental condition always receive it. As we have client-level data on whether the child actually received the ISS, we now wish to focus on the effect of Treatment on the Treated. This is accomplished by fitting a model that uses the client's PREMISS score and their assigned treatment to predict whether they actually receive treatment. A second form models the client's PSC score based on their actual receipt of ISS and their PREMISS score. Based on this model, the clients who receive ISS have an average PSC of 9.74, while those that did not receive ISS had an average PSC of 9.74 + -1.6 = 8.14, a difference that does reach statistical significance (p-value = 0.03).

TABLE 156 RESULTS FROM A TOT ANALYSIS – PSC SCORES

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	9.74	0.64	15.23	<0.01
SAU	-1.6	0.74	-2.16	0.03
PREMISS	-0.35	0.25	-1.41	0.16

Taking a slightly different approach, we also took a closer look at answering a question about specific time points. Starting with those individuals who start off in the PSC age range (those 4 years old and older), we wanted to explore whether the treatment and control groups show different trajectories from their first PSC measurement to their measurements taken after 3 and 6 months in their treatment condition. Using a similar approach to the two-stage least squares (TSLS) regression above, we fit one model that regresses an individual's Treatment Received variable onto their PREMISS score and their PSC score at their first time point. We then fit a second model regressing a client's PSC score at 3, or 6 months onto their predicted ISS-Received value from the first model, their PREMISS score, and their first PSC score. The results are shown in the tables below. We can see that those in the ISS Received group

do not show different PSC scores at the 3-month mark relative to the SAU group (estimate = -1.13, p = 0.6). The same is true of the comparisons made at the six-month mark (estimate = 0.6, p = 0.82).

TABLE 157 RESULTS FROM A TOT ANALYSIS – PSC TOTAL AT 3 MONTHS

	Estimate	Std. Error	t value	Random SD
Intercept	2.99	1.46	2.04	0.04
ISS	-1.13	2.16	-0.52	0.60
PSC Total - First	0.48	0.07	7.24	<0.01
PREMISS	-0.73	0.72	-1.02	0.31

TABLE 158 RESULTS FROM A TOT ANALYSIS – PSC TOTAL AT 6 MONTHS

	Estimate	Std. Error	t value	Random SD
Intercept	6.38	1.91	3.35	<0.01
ISS	0.60	2.59	0.23	03.82
PSC Total - First	0.26	0.08	3.40	<0.01
PREMISS	0.59	0.92	0.64	0.52

For those clients younger than four years we turn to the Baby Pediatric Symptom Checklist (BPSC), and the Preschool Pediatric Symptom Checklist (PPSC).

Specifically, BPSC scores are reported for 253 clients between 0 and 17 months of age (27.99% of total number of unique clients). The BPSC has 12 items with the same possible responses as the PSC scores, with the total BPSC score being the sum of each item response. Like the PSC, a higher score on the BPSC indicates more symptoms.

FIGURE 51 shows the BPSC total scores for the sample in terms of the number of days since their first screener. The thin lines are the BPSC scores for individual children; the lines are color-coded according to the key in the upper right corner of the Figure: purple for "ISS", yellow for "SAU". The thicker lines of the same colors are LOESS lines that are smooth local regression lines that can be thought of like a running mean. The shaded grey regions around the LOESS lines are the confidence intervals around the LOESS lines. The lines do not appear significantly different.
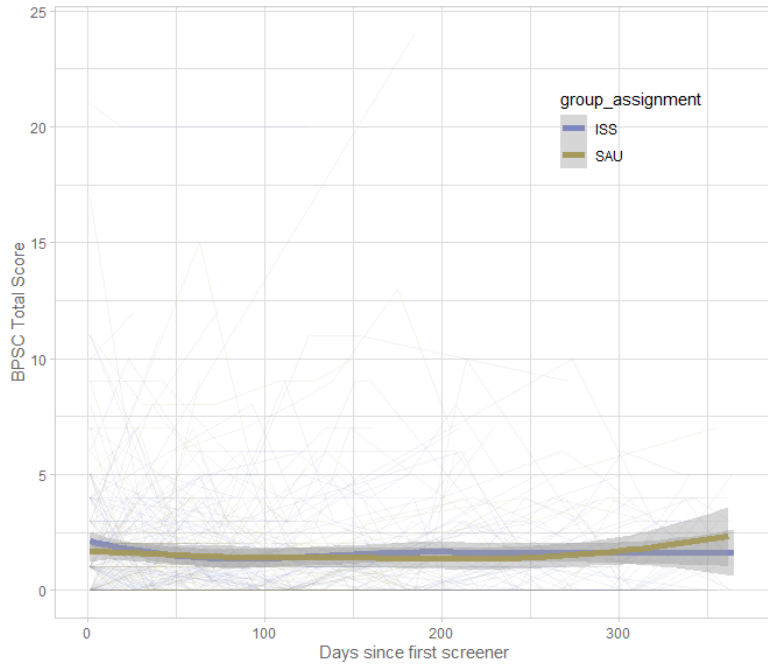
TABLE 159 - COUNT OF MEASUREMENTS WITHIN ASSIGNED GROUP CONDITION

Condition Assigned	Count
ISS	1680
SAU	865

The results of the linear mixed effects model for Treatment on the Treated show a similar story. Those in the SAU conditions do not show any significant differences in BPS scores relative to those assigned to ISS (estimate = -0.36, t = -1). Both conditions show a slow but statistically significant increase in symptoms over time (estimate = 0.67, t = 2.34), with no significant difference between SAU-Assigned and ISS-Assigned (estimate = -0.43, t = -0.84).

168

TABLE 160 RESULTS FROM A LINEAR MIXED EFFECTS MODEL ON BPSC

	Estimate	Std. Error	t value	Random SD
Intercept	1.22	0.49	2.48	2.3
SAU	-0.36	0.36	-1.00	NA
Time Since 1st Measurement	0.67	0.29	2.34	NA
PREMISS	-0.36	0.26	-1.38	NA
Time Since 1st Measurement * SAU	-0.43	0.51	-0.84	NA
Residual	NA	NA	NA	1.82

We were also interested in exploring the influence of regional effects. To start, we ran an ANOVA to test for the significance of the differences in the distribution of BPSC scores across the 5 regions, and within the SAU and ISS-Assigned conditions. The results are displayed in the graphs and tables below. The result is that it seems that BPSC scores do not quite differ between assignment groups and between regions.

TABLE 161 BPSC SCORES ACROSS REGIONS AND ISS ASSIGNMENT CONDITION

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Assigned Group	1	5.31	5.31	0.75	0.4
Region	4	64.05	16.01	2.26	0.1
Assigned Group *Region	4	28.26	7.06	1.00	0.4
Residuals	1178	8342.71	7.08	NA	NA

FIGURE 52 PSC SCORES ACROSS REGIONS AND ISS-ASSIGNED, AND SAU ASSIGNED CONDITIONS
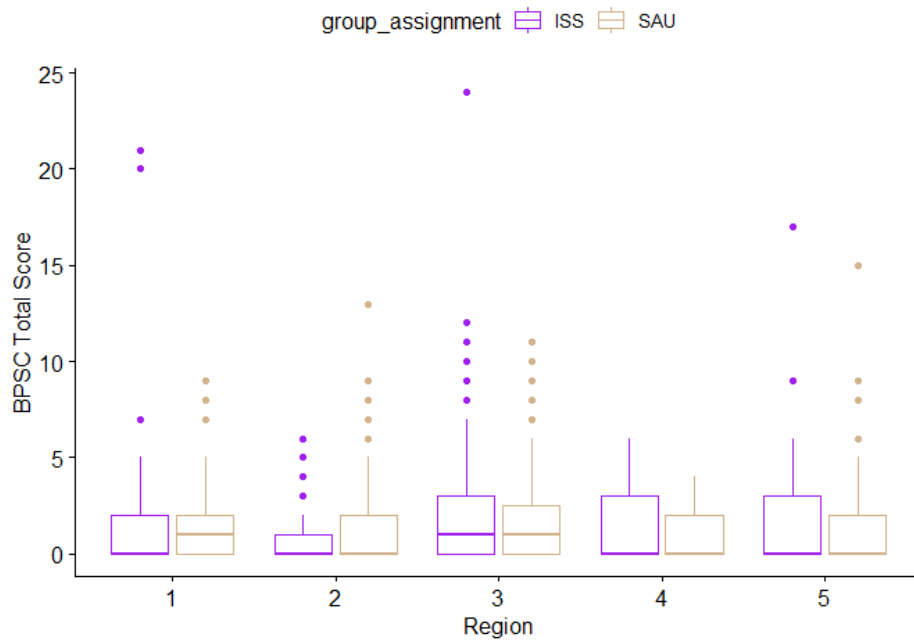
We followed this up by doing the same thing, but dividing the ISS assigned into its two component conditions: those who were assigned to ISS and actually received ISS, and those assigned to the ISS condition, but ISS was found to be unworkable. The results are shown in the graph and table below. Here, it seems that BPSC scores do show different distributions across regions and assignment conditions.

TABLE 162 BPSC SCORES ACROSS REGIONS AND ISS-RECEIVED, ISS-NOT WORKABLE AND SAU CONDITIONS

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Assigned Group	2	19.59	9.8	1.39	0.25
Region	4	62.78	15.7	2.23	0.06
Assigned Group *Region	8	118.03	14.75	2.1	0.03
Residuals	1173	8239.93	7.02	NA	NA

FIGURE 53 BPSC SCORES ACROSS REGIONS AND ISS-RECEIVED, ISS-NOT WORKABLE AND SAU CONDITIONS
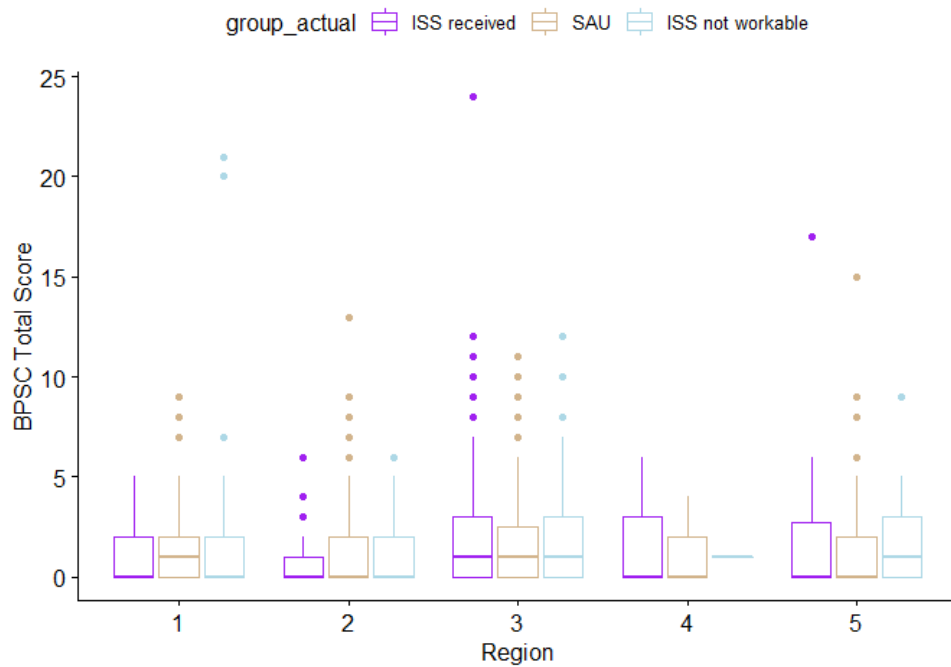
On the basis of these results, we fit a mixed-effects model regressing a client's total BPSC score onto a new variable that indicates the combination of a client's ISS Assignment condition and their region. This model also retains the random effects of PREMISS Score and the time since the first occasion of BPSC Measurement, and the fixed effect of client ID. The idea is to fit a model that gives us some detail about the influence of intention-to-treat in a unique region.

The following table shows the results of this model. Note that for purposes of this model, the referent group is those individuals from Region 3 (Oklahoma County) who were assigned to ISS. The one noteworthy find is that those in the ISS-Assigned condition in Region 2 show a significantly lower BPSC Score relative to the referent group (estimate = -1.42, t = -2.17), while those in Region 2 assigned to the SAU condition do not (estimate = -0.95, t = -1.28). Surprisingly, we see that, as time passes, BPSC scores tend to increase (estimate = 0.52, t = 2.19).

TABLE 163 REGRESSING BPSC ONTO CONDITION ASSIGNMENT-BY-REGION VARIABLES

	Estimate	Std. Error	t value	Random SD
Intercept	1.77	0.63	2.81	2.28
Region 1 – ISS Assigned	-0.24	0.62	-0.4	NA
Region 1 – SAU Assigned	-1.05	0.56	-1.86	NA
Region 2 – ISS Assigned	-1.42	0.66	-2.17	NA
Region 2 – SAU Assigned	-0.95	0.74	-1.28	NA
Region 3 – SAU Assigned	-0.56	0.6	-0.93	NA
Region 4 – ISS Assigned	-1.53	1.29	-1.19	NA
Region 4 – SAU Assigned	-1.37	1.31	-1.05	NA
Region 5 – ISS Assigned	-0.83	0.79	-1.05	NA
Region 5 – SAU Assigned	-0.58	0.69	-0.83	NA
PREMISS	-0.27	0.29	-0.93	NA
Days Since 1st Measurement	0.52	0.24	2.19	NA
Residual	NA	NA	NA	1.82

The following table shows the results of the same model, but examining actual ISS condition instead of assigned condition. We see that those who actually receive ISS in Region 2 show significantly lower BPSC Scores relative to the Referent group (estimate = -2.07, t = -2.67). Those in Region 2 that are assigned to SAU or who are assigned to ISS but do not receive it also may show lower BPSC scores relative to the referent group, but they do not reach significance.

We also continue to see a significant positive relationship between BPSC Scores and time, (estimate = 0.5, t = 2.11).

172

TABLE 164 REGRESSING BPSC ONTO TREATMENT RECEIVED-BY-REGION

	Estimate	Std. Error	t value	Random SD
Intercept	2.06	0.66	3.13	2.27
Region 1 – ISS Not Workable	1.45	0.96	1.50	NA
Region 1 – ISS Received	-1.48	0.76	-1.95	NA
Region 1 - SAU	-1.24	0.6	-2.07	NA
Region 2 – ISS Not Workable	-1.00	0.85	-1.17	NA
Region 2 – ISS Received	-2.07	0.77	-2.67	NA
Region 2 - SAU	-1.14	0.77	-1.49	NA
Region 3 – ISS Not Workable	-0.5	0.58	-0.86	NA
Region 3 - SAU	-0.74	0.64	-1.17	NA
Region 4 – ISS Not Workable	-1.72	1.3	-1.33	NA
Region 4 – ISS Received	-1.57	1.32	-1.19	NA
Region 4 - SAU	-1.01	1.09	-0.92	NA
Region 5 – ISS Not Workable	-1.11	1.04	-1.07	NA
Region 5 – ISS Received	-0.81	0.72	-1.12	NA
Region 5 - SAU	-0.22	0.29	-0.74	NA
PREMISS	0.5	0.24	2.11	NA
Days Since 1st Measurement	NA	NA	NA	1.82
Residual	2.06	0.66	3.13	2.27

To give more meaning to the changes in scores on the BPSC measure, we wanted to examine changes in BPSC measurements in comparison to clinical cutoff values. The BPSC scale has three subscales (Irritability, Inflexibility, and Routines), each with a unique cut-off value that identifies clinically significant cases; getting a subscale score at or above these cutoffs indicates a case of clinical concern. To examine how ISS Assignment condition influences scores relative to these cutoffs, we fit three new logistic models predicting the odds of a score reaching or exceeding a cutoff value. Each model uses ISS Assignment condition, an interaction between assignment condition and time since first measurement, a random intercept of child ID to account for repeated measures, and the family's PREMISS score as covariates.

For the Irritability subscale of the BPSC measurement, those assigned to SAU are not any more or less likely to cross the threshold designating a clinically significant case, (estimate = 1.01, p = 0.45).

There is, however, a significant interaction between assignment to the SAU condition and time (estimate = -5.87, p = 0.02). This suggests that, as time increases, the probability of getting a score on the BPSC-Irritability subscale above the critical threshold decreases for those in the SAU-Assigned condition more than for those in the ISS-Assigned condition.

TABLE 165 LOGISTIC MODEL OF BPSC-IRRITABILITY SCORE ABOVE CUTOFF

	Estimate	Std. Error	t value	Pr(>|z|)	Random SD
Intercept	-10.52	2.04	-5.16	<0.00	9.45
SAU	1.01	1.34	0.76	0.45	NA
Days Since 1st Measurement	1.83	1.14	1.61	0.11	NA
PREMISS	-0.4	0.89	-0.45	0.66	NA
SAU * Days Since 1st Measurement	-5.87	2.61	-2.25	0.02	NA
Residual	NA	NA	NA	NA	1

For the Inflexibility subscale of the BPSC measurement, those assigned to SAU are not any more or less likely to cross the threshold designating a clinically significant case (estimate = -0.93, p = 0.24).

The only significant main effect is that of the time since the first BPSC measure, (estimate = 2.76, p = <0.01). This suggests the probability of getting a score on the BPSC-Inflexibility subscale above the critical threshold decreases significantly over time. Since the interaction with being assigned to SAU is not significant, the effect of time seems to be the same for both those assigned to SAU and ISS conditions.

TABLE 166 LOGISTIC MODEL OF BPSC-INFLEXIBILITY SCORE ABOVE CUTOFF

	Estimate	Std. Error	t value	Pr(>|z|)	Random SD
Intercept	-7.13	1.35	-5.28	<0.01	4.1
SAU	-0.93	0.8	-1.16	0.24	NA
Days Since 1st Measurement	2.76	0.72	3.86	<0.01	NA
PREMISS	-0.86	0.54	-1.6	0.11	NA
SAU * Days Since 1st Measurement	1.12	1.26	0.88	0.38	NA
Residual	NA	NA	NA	NA	1.0

For the Routines subscale of the BPSC measurement, those in assigned to SAU are not any more or less likely to cross the threshold designating a clinically significant case, (estimate = 0.18, p = 0.84).

Indeed, none of the variables in this model show a significant relation to the probability of getting a clinically significant Routine Subscale score.

TABLE 167 LOGISTIC MODEL OF BPSC-ROUTINES SCORE ABOVE CUTOFF

	Estimate	Std. Error	t value	Pr(>|z|)	Random SD
Intercept	-7.74	1.49	-5.19	<0.01	5.4
SAU	0.18	0.89	0.21	0.84	NA
Days Since 1st Measurement	-0.79	1.03	-0.76	0.45	NA
PREMISS	-0.45	0.62	-0.72	0.47	NA
SAU * Days Since 1st Measurement	-0.32	1.84	-0.18	0.86	NA
Residual	NA	NA	NA	NA	1

To examine the effect of Treatment on the Treated in greater detail during the first few months of the intervention, we also ran a TSLS analysis to test for different trajectories between client's first BPSC scores and their BPSC scores at 3 and 6 months. The results are shown in the tables below. Unfortunately, we don't see any significant change between ISS Received and SAU groups at 3 or at 6 months. At 3 months, the only significant predictor of BPSC score is the child's first BPSC score (estimate = 0.53, t = 8.17); the first BPSC score drops in significance by 6 months (estimate = 0.24 t = 1.53).

TABLE 168 RESULTS FROM A TOT ANALYSIS – BPSC TOTAL AT 3 MONTHS

	Estimate	Std. Error	t value	Random SD
Intercept	-0.09	0.60	-0.15	0.88
ISS	-0.08	0.63	-0.13	0.90
PSC Total - First	0.53	0.07	8.17	<0.01
PREMISS	-0.37	0.30	-1.25	0.21

TABLE 169 RESULTS FROM A TOT ANALYSIS – BPSC TOTAL AT 6 MONTHS

	Estimate	Std. Error	t value	Random SD
Intercept	-0.02	1.48	-0.02	0.99
ISS	1.05	1.48	0.71	0.48
PSC Total - First	0.24	0.15	1.53	0.13
PREMISS	-0.63	0.68	-0.93	0.36

For those clients too young for the PSC and too old for the BPSC, we have the Preschool Pediatric Symptom Checklist (PPSC) The PPSC scores are reported for 72 clients between 18 and 48 months of age (7.96% of total number of unique clients). The PPSC has 18 items with the same possible responses as the PSC scores, with the total PPSC score being the sum of each item response. Like the PSC, a higher score on the PPSC indicates more symptoms.

The figure below shows the PPSC total scores for the sample in terms of the number of days since their first screener. Here, it appears that those in the SAU condition may start out with a higher PPSC score, while generally hovering around the same mean as those assigned to ISS. However, due to the relatively small number of individuals who received the PPSC screener, particularly in the SAU condition, it is difficult to draw strong conclusions.

FIGURE 54 PPSC TOTAL SCORES FOR SAMPLE VS DAYS SINCE FIRST SCREENER IN ASSIGNED GROUP
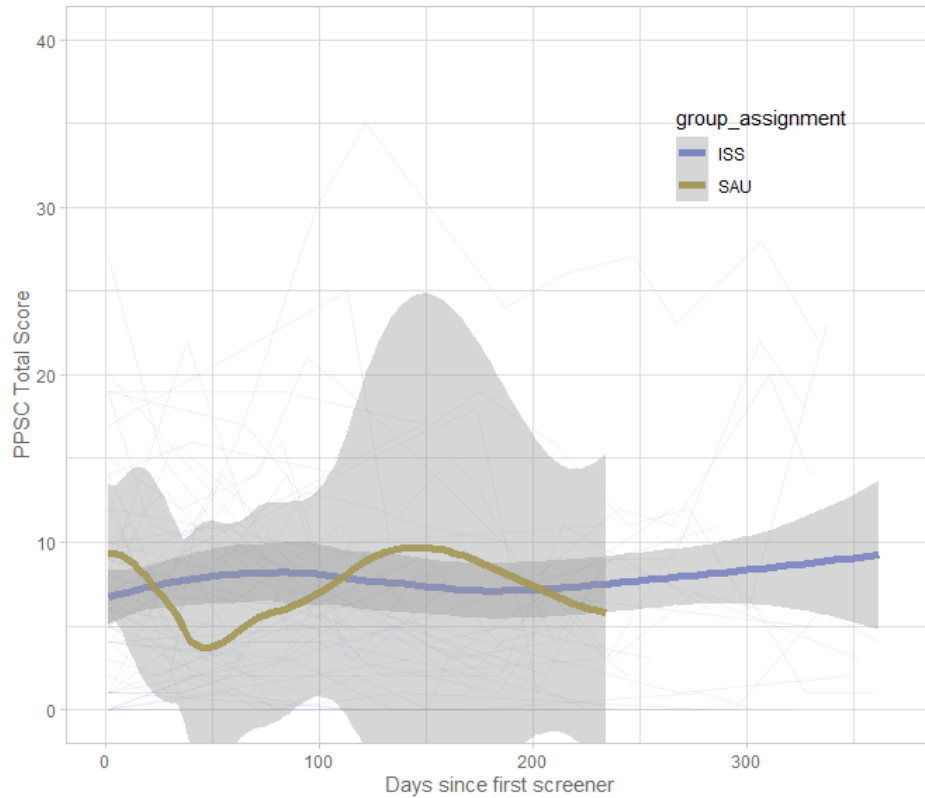
TABLE 170 COUNT OF MEASUREMENTS WITHIN ASSIGNED GROUP CONDITION

Condition Assigned	Count
ISS	254
SAU	23

Since this graph looks particularly non-linear over time, we fit a Linear Mixed Effects model showing PPSC scores across time, with time divided into bins (prior to 45 days, 45 to 90 days, and after 90 days); we again allowed a random slope for each bin, and a random intercept based on child ID. There is no significant main effect of assignment to SAU (estimate = 3.39 t = 1.22). We also see some unreasonably large estimates among the interaction effects involving SAU and the Time Bin variables. This is likely a symptom of having few children linked to PPSC scores, particularly in the SAU condition.

TABLE 171 RESULTS FROM A LINEAR MIXED EFFECTS MODEL ON PPSC

	Estimate	Std. Error	t value	Random SD
Intercept	8.11	2.35	3.46	5
PREMISS	0.83	1.11	0.74	NA
SAU	3.39	2.78	1.22	NA
Time Since 1st Measurement	7.52	11.3	0.67	NA
Time Between Days 45 and 90	3.02	11.6	0.26	NA
Time Since Day 90	-2.13	2.2	-1.00	NA
Time Since 1st Measurement * SAU	-87.6	35.4	-2.5	NA
Time Between Days 45 and 90 * SAU	49.61	48.3	1.03	NA
Time Since Day 90 * SAU	-14.1	18	-0.8	NA
Residual	NA	NA	NA	4.6

We were also interested in exploring the influence of regional effects. To start, we ran an ANOVA to test for the significance of the differences in the distribution PPSC scores across the 5 regions, and within the SAU and ISS-Assigned conditions. The results are displayed in the graphs and tables below. The result is that it seems that PPSC scores do NOT differ between assignment groups and between regions. However, we are again hampered by the small sample sizes among those who received the PPSC Measurement.

TABLE 172 PPSC SCORES ACROSS REGIONS AND ISS ASSIGNMENT CONDITION

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Assigned Group	1	0.18	0.18	0.00	1.00
Region	1	37.55	37.55	0.81	0.4
Assigned Group *Region	1	27.78	27.78	0.6	0.4
Residuals	273	12691.63	46.49	NA	NA

FIGURE 55 PPSC SCORES ACROSS REGIONS AND ISS-ASSIGNED, AND SAU ASSIGNED CONDITIONS
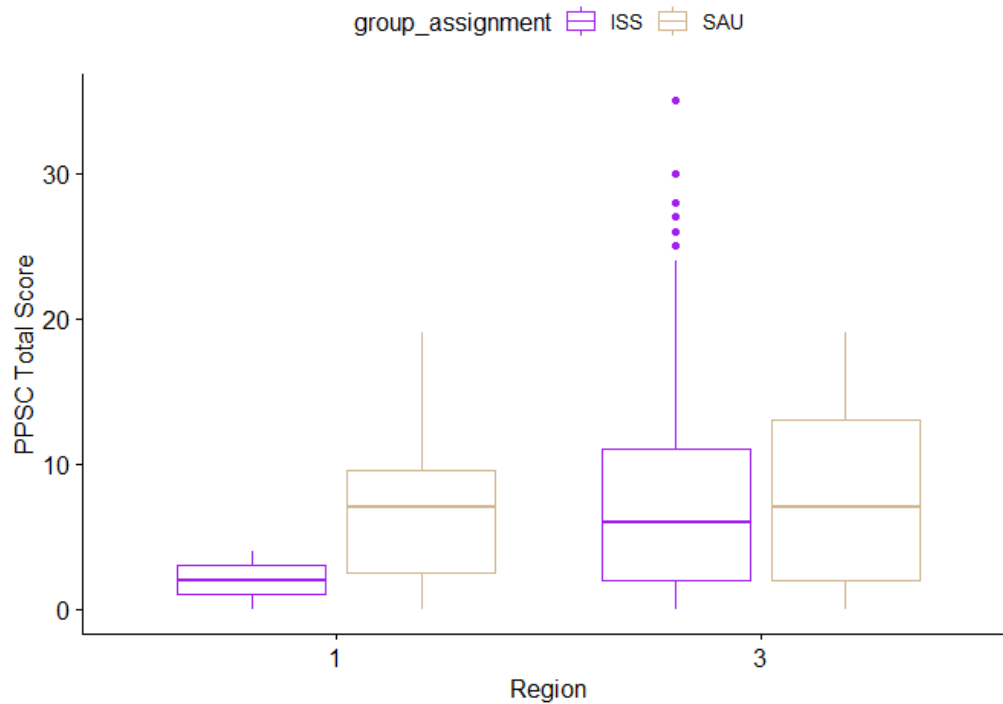
We followed this up by doing the same thing, but dividing the ISS assigned into its two component conditions: 1) those who were assigned to ISS and those who actually received ISS, and 2) those assigned to the ISS condition, but ISS was found to be unworkable. The results are shown in the graph and table below. We again see little differences in the distribution of PPSC Scores across regions and conditions.

179

TABLE 173 PPSC SCORES ACROSS REGIONS AND ISS-RECEIVED, ISS-NOT WORKABLE AND SAU CONDITIONS

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Assigned Group	2	50.87	25.4	0.6	0.6
Region	1	42.9	42.9	0.9	0.3
Assigned Group * Region	1	34.79	34.8	0.8	0.4
Residuals	272	12628.6	46.4	NA	NA

FIGURE 56 PPSC SCORES ACROSS REGIONS AND ISS-RECEIVED, ISS-NOT WORKABLE AND SAU CONDITIONS
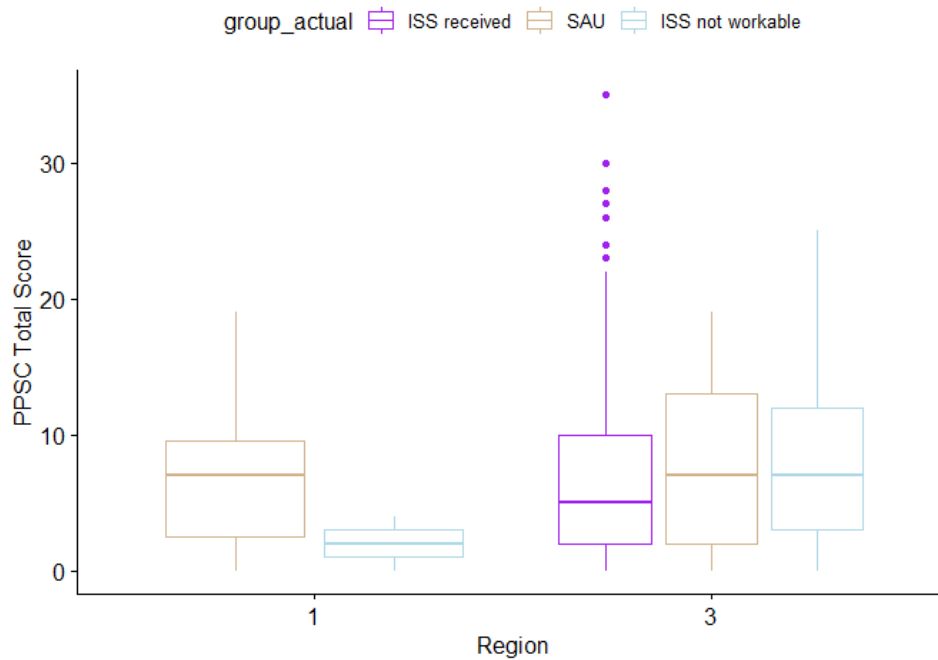
To give more meaning to the changes in the PPSC scores, we wanted to examine changes in PPSC measurements in comparison to clinical cutoff values. The PPSC Measure does not have subscales, and so has a single cutoff score that defines a case of clinical concern. To examine how ISS Assignment condition influences scores relative to this cutoff, we fit a logistic regression predicting the odds of reaching or exceeding a cutoff value. Like the previous logistic models, this model contains variables for ISS Assignment condition interacted with time since first measurement, a random intercept of child ID to account for repeated measures, as well as other covariates.

Those assigned to SAU are not any more or less likely to cross the threshold designating a clinically significant case, (estimate = 0.99, p = 0.99).

Indeed, none of the variables of in this model show a significant relation to the probability of getting a clinically significant PPSC score.

TABLE 174 LOGISTIC MODEL OF PPSC SCORE ABOVE CUTOFF

	Estimate	Std. Error	t value	Pr(>|z|)	Random SD
Intercept	-0.27	0.89	-0.3	0.76	1.63
SAU	0.99	118.49	0.01	0.99	NA
Days Since 1st Measurement	0.43	0.69	0.62	0.54	NA
PREMISS	0.38	0.43	0.89	0.38	NA
SAU * Days Since 1st Measurement	-215	43276.48	0	1	NA
Residual	NA	NA	NA	NA	1

We also ran a TSLS analysis comparing the trajectories from first PPSC score to their PPSC score at 3 months and 6 months. However, the high missingness in the PPSC scores leaves a very restricted sample. The results for the TSLS models are shown below, but the high effect sizes of the ISS Received variable are almost certainly an artefact of the small sample size.

TABLE 175 RESULTS FROM A TOT ANALYSIS – PPSC TOTAL AT 3 MONTHS

	Estimate	Std. Error	t value	Random SD
Intercept	8.98	6	1.5	0.15
ISS	17.11	20.7	0.83	0.42
PSC Total - First	0.69	0.35	1.96	0.07
PREMISS	4.97	4.68	1.06	0.30

TABLE 176 RESULTS FROM A TOT ANALYSIS – PPSC TOTAL AT 6 MONTHS

	Estimate	Std. Error	t value	Random SD
Intercept	27.93	14.39	1.94	0.10
ISS	-10.87	11.83	-0.92	0.39
PSC Total - First	-1.06	0.87	-1.23	0.27

Finally, we turn to the Milestones measurement. This tool has 10 items with the same possible responses as the PSC scores, with the total Milestones score being the sum of each item response. Unlike the previous measures, higher scores on the Milestones measure means greater physical and cognitive development. This measurement was meant for children 0 to 4 years of age, representing 437 individuals, or (48.34% of the total number of unique clients).

The graph below plots the Milestones Score for the different treatment groups across time. Here, it seems that the two groups start with similar average scores, with those in the SAU scores increasing to a slightly higher average Milestones score than those assigned to ISS. Those in the ISS score may catch up over time, around day 300 or beyond.

FIGURE 57 MILESTONES TOTAL SCORES FOR SAMPLE VS DAYS SINCE FIRST SCREENER IN ASSIGNED GROUP
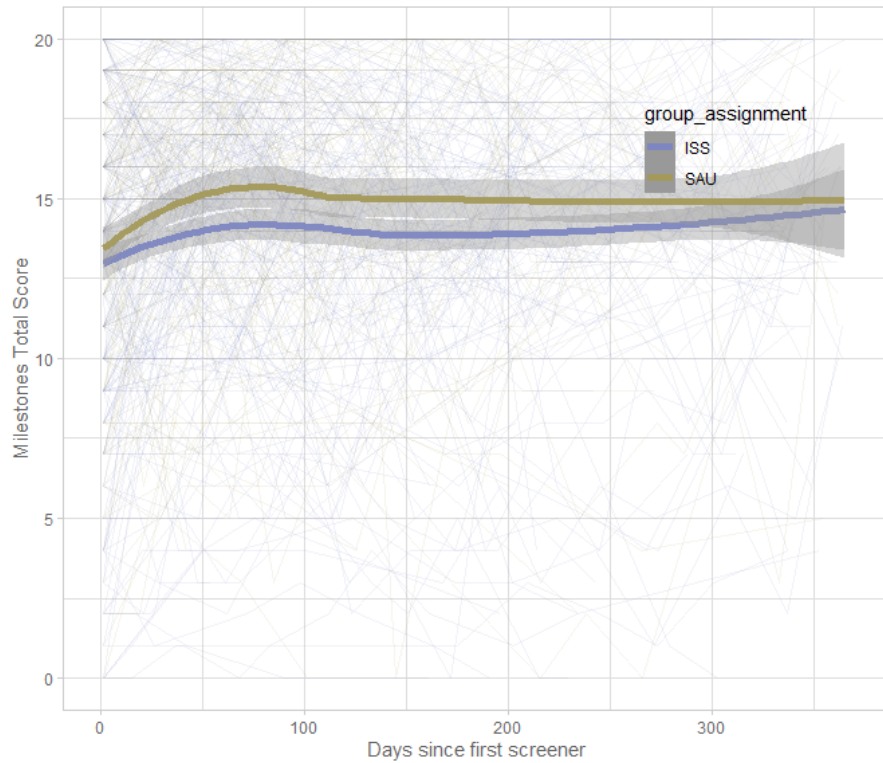
TABLE 177 COUNT OF MEASUREMENTS WITHIN ASSIGNED GROUP CONDITION

Condition Assigned	Count
ISS	1493
SAU	763

The Milestones measure involves several different sets of questions, each meant for children of a certain age group. To get a more detailed picture of the effect of ISS assignment on Milestones score, we calculated the mean score for all items meant for the same age group (for all assignment conditions), took the difference of each Milestones measure and the mean for the corresponding age range, and plotted the new scores for each assignment condition across time. The results are shown below. We see

that those in the SAU and ISS-Received conditions are comparable, with both possibly outperforming those in the 'ISS Not Workable' condition, until at least a year after the first measurement.

FIGURE 58 - MILESTONES DEVIATION FROM AVERAGE VS DAYS SINCE FIRST SCREENER IN ACTUAL GROUP
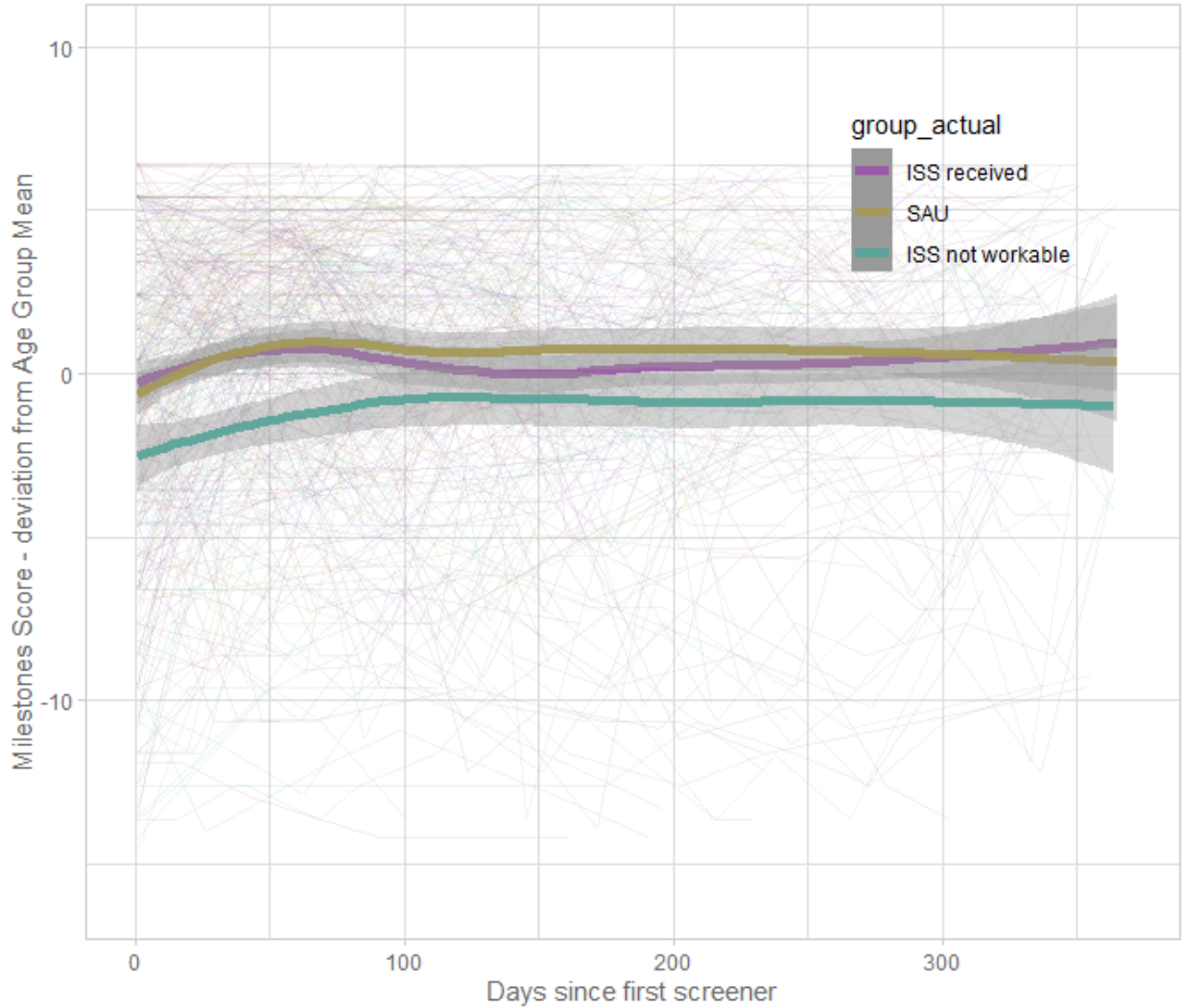
TABLE 178 - Count of Measurements within Actual Group Condition

Condition Assigned	Count
ISS received	888
SAU	763
ISS not workable	605

Again, we have fit a Linear Mixed Effects model with a random intercept based on child ID. The results are displayed below. We see that those in the SAU condition appear to have a slightly higher Milestones Score at the start, although the difference does not quite reach significance (t = 2.11). We also have

support that both groups trend upwards across time (estimate = 1.84, t = 5.21); the trajectories for these two conditions across time do not seem to be significantly different (estimate = 0.88, t = 1.38).

TABLE 179 - Count of Measurements within Actual Group Condition

	Estimate	Std. Error	t value	Random SD
Intercept	14.4	0.61	23.62	3.77
SAU	0.71	0.48	1.48	NA
Days Since 1st Measurement	1.84	0.35	5.21	NA
PREMISS	0.69	0.33	2.11	NA
SAU * Days Since 1st Measurement	0.88	0.64	1.38	NA
Residual	NA	NA	NA	3.06

We were also interested in exploring the influence of regional effects. To start, we ran an ANOVA to test for the significance of the differences in the distribution Milestones scores across the 5 regions, and within the SAU and ISS-Assigned conditions. The results are displayed in the graphs and tables below. The result is that it seems that Milestones scores do differ between assignment groups and between regions.

TABLE 180 - Milestones Scores Across Regions and ISS Assignment Conditions

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Assigned Group	1	345.72	345.72	15.27	<0.01
Region	4	515.36	128.84	5.69	<0.01
Assigned Group *Region	4	286.12	71.53	3.16	0.01
Residuals	2246	50843.6	22.64	NA	NA

FIGURE 181 - MILESTONES Scores Across Regions and ISS-Assigned, and SAU Assigned conditions
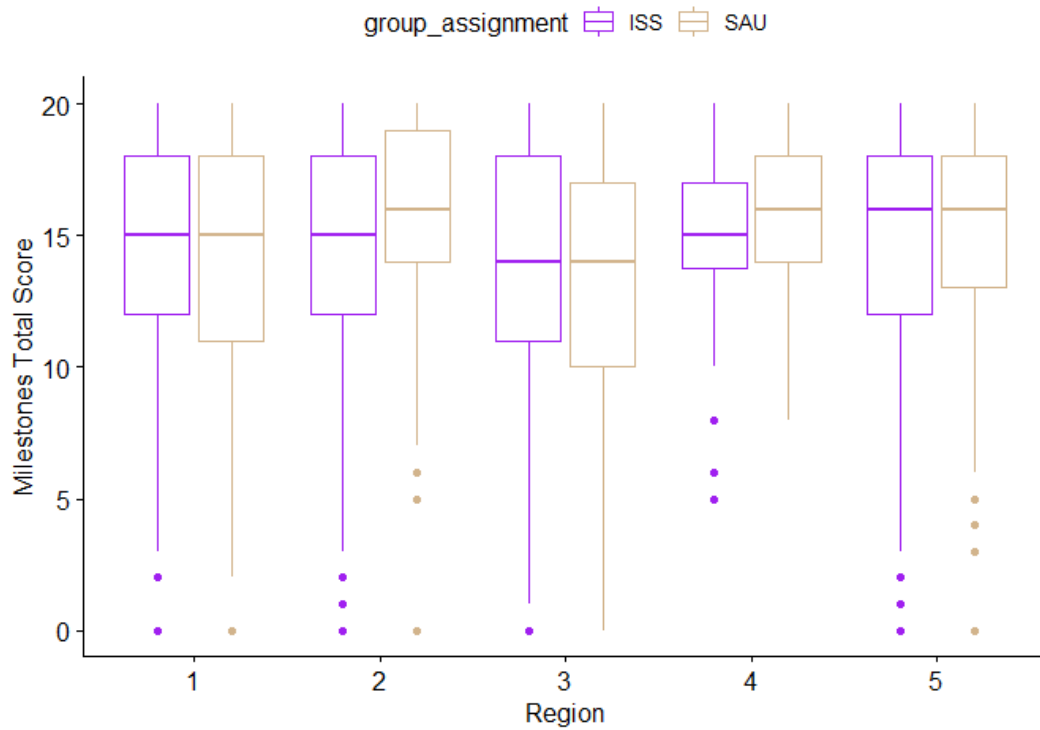
We followed this up by doing the same thing, but dividing the individuals assigned to ISS assigned into two component conditions: 1) those who were assigned to ISS and actually received ISS, and 2) those assigned to the ISS condition, but ISS was found to be unworkable. The results are shown in the table and graph below. Again, it seems that Milestones scores do show different distributions across regions and assignment conditions.

TABLE 182 - Milestones Scores Across Regions and ISS Assignment Conditions

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Assigned Group	2	1345.69	672.8	30.3	<0.01
Region	4	453.93	113.5	5.12	<0.01
Assigned Group *Region	8	496.62	62.08	2.8	<0.01
Residuals	2241	49694.6	22.18	NA	NA

FIGURE 183 - MILESTONES Scores Across Regions and ISS-Received, ISS-Not Workable and SAU conditions
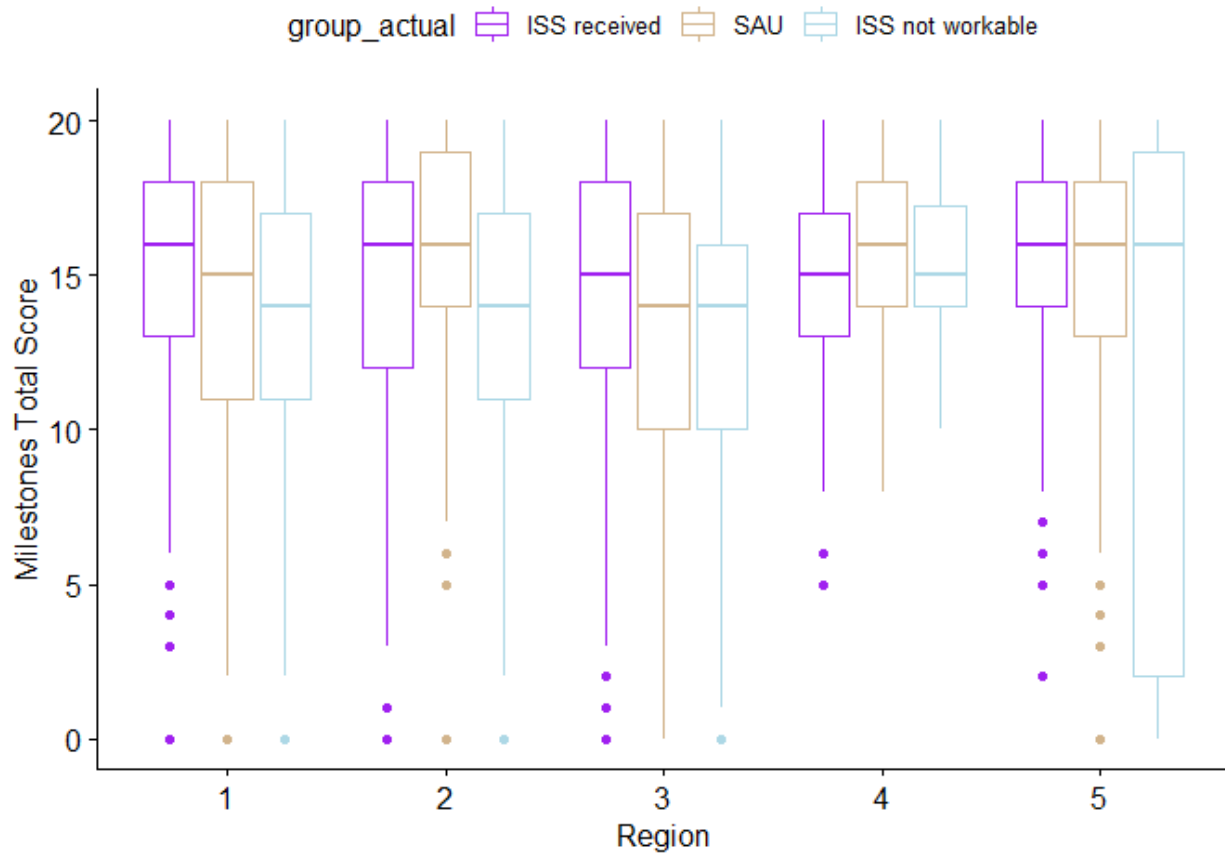
On the basis of these results, we fit a mixed effects model regressing a client's Milestones score onto a new variable that indicates the combination of a client's ISS Assignment condition and their region. This model also retains the random effects of PREMISS Score and the time since the first occasion of Milestones measurement, and the fixed effect of client ID. The idea is to fit a model that gives us some detail about the influence of intention-to-treat in a unique region.

The following table shows the results of this model. Note that for purposes of this model, the referent group is those individuals from Region 3 (Oklahoma County) who were assigned to ISS. We see that those assigned to SAU in Region 2 show significantly higher Milestones scores as a group relative to those in the referent group (estimate = 2.24, t = 2.33). The same is true of those in Region 4 (estimate = 2.97, t = 2.1). Though most differences do not seem to be significant, it seems that the general trend is for those assigned to the SAU condition to have higher Milestones scores than those assigned to the ISS condition.

TABLE 184 - Regressing Milestones onto Condition Assignment-by-Region Variables

	Estimate	Std. Error	t value	Random SD
Intercept	13.67	0.76	18.06	3.8
Region 1 – ISS Assigned	0.55	0.81	0.68	NA
Region 1 – SAU Assigned	0.77	0.74	1.05	NA
Region 2 – ISS Assigned	1.12	0.82	1.37	NA
Region 2 – SAU Assigned	2.24	0.96	2.33	NA
Region 3 – SAU Assigned	0.95	0.86	1.1	NA
Region 4 – ISS Assigned	1.7	1.58	1.08	NA
Region 4 – SAU Assigned	2.97	1.41	2.1	NA
Region 5 – ISS Assigned	1.14	1.05	1.08	NA
Region 5 – SAU Assigned	1.22	0.87	1.41	NA
PREMISS	0.53	0.36	1.47	NA
Days Since 1st Measurement	2.11	0.29	7.19	NA
Residual				3.06

The following table shows the results of the same model, but examining Treatment on the Treated. For this model, the referent group is those in Region 3 who actually received ISS. The only significant effect is that those in Region 3 who are assigned to ISS but do not receive it show significantly lower Milestones scores relative to the referent group (estimate = -1.95, t = -2.69). While the differences generally fail to be statistically significant, the general trend seems to be that those assigned to ISS but who do not receive it tend to have lower scores relative to the other Treatment Received conditions in the same region.

TABLE 185 - Regressing Milestones onto Treatment Received-by-Region

	Estimate	Std. Error	t value	Random SD
Intercept	14.49	0.82	17.69	3.77
Region 1 – ISS Not Workable	-0.47	1.16	-0.4	NA
Region 1 – ISS Received	-0.12	1.1	-0.11	NA
Region 1 - SAU	-0.06	0.8	-0.07	NA
Region 2 – ISS Not Workable	-0.58	1.18	-0.49	NA
Region 2 – ISS Received	0.76	0.97	0.78	NA
Region 2 - SAU	1.41	1.01	1.4	NA
Region 3 – ISS Not Workable	-1.95	0.72	-2.69	NA
Region 3 - SAU	0.12	0.92	0.13	NA
Region 4 – ISS Not Workable	0.41	2.84	0.15	NA
Region 4 – ISS Received	1.06	1.89	0.56	NA
Region 4 - SAU	2.13	1.44	1.48	NA
Region 5 – ISS Not Workable	0.25	1.64	0.15	NA
Region 5 – ISS Received	0.35	1.3	0.27	NA
Region 5 - SAU	0.39	0.92	0.43	NA
PREMISS	0.53	0.36	1.45	NA
Days Since 1st Measurement	2.13	0.29	7.28	NA
Residual	NA	NA	NA	3.05

We also fit a logistic regression predicting the probability of a Milestones Score meeting or exceeding a specific cutoff value. Note that for this analysis, being above a cutoff indicates a score in a 'healthy' range.

The results are in the table below. We see no main effect of ISS Assignment condition (estimate = 0.68, p = 0.14). However, there is a significant main effect of time, such that the more time passes since the first measurement, the more likely they are to reach the 'healthy' threshold (estimate = 1.09, p = 0.01).

TABLE 186 - Logistic Model of Milestones Score Above Cutoff

	Estimate	Std. Error	t value	Pr(>|z|)	Random SD
Intercept	1.59	0.56	2.84	<0.01	2.61
SAU	0.68	0.47	1.47	0.14	NA
Days Since 1st Measurement	1.09	0.42	2.61	0.01	NA
PREMISS	0.17	0.28	0.6	0.55	NA
SAU * Days Since 1st Measurement	-0.29	0.75	-0.38	0.7	NA
Residual	NA	NA	NA	NA	1

The following table shows the results of a TSLS analysis comparing the trajectory of the ISS Received and comparison groups from their first Milestones score to Milestones score at 3 months, and 6 months. It is worth reiterating that the Milestones score is a measure of physical development and so, unlike with the previous measures, a higher score indicates greater health. The results are shown below. Unexpectedly, being in the ISS-Received condition seems to be associated with a lower Milestones score at 6 months, (p = 0.05). The same is apparent, if not quite significant at 3 months, (p = 0.08).

TABLE 187 - RESULTS FROM A TOT ANALYSIS – MILESTONES TOTAL AT 3 MONTHS

	Estimate	Std. Error	t value	Random SD
Intercept	10.65	1.09	9.77	<0.01
ISS	-1.78	1.01	-1.77	0.08
PSC Total - First	0.39	0.05	7.26	<0.01
PREMISS	0.43	0.4	1.06	0.29

TABLE 188 - RESULTS FROM A TOT ANALYSIS – MILESTONES TOTAL AT 6 MONTHS

	Estimate	Std. Error	t value	Random SD
Intercept	9.57	1.37	6.96	<0.01
ISS	-2.77	1.41	-1.96	0.05
PSC Total - First	0.50	0.07	7.29	<0.01
PREMISS	0.68	0.56	1.22	0.23

The results reported here are preliminary and should be interpreted with caution because of their limitations. First, only about 50% of the Children in Waiver have any corresponding screening information. This could be because screeners were not done on these children, but could also result from an inability to accurately link the screener that was done on a child to that child. Another likely reason for the lack of screeners linked to certain children is special needs. A preliminary analysis found that some special needs significantly predicted lacking (or having) a screen on record. And the match rate remained high. However, it is still possible that that we systematically missed matches for Waiver data. Second, the model of change was usually linear. Such a model is implausible for long term change processes. Efforts were made to explore alternative time formats, but more should be done to adequately capture this nonlinear process. A growth model that asymptotes to zero should be used as more data are collected and matched. Future reports should investigate and correct these limitations. Additionally, we attempted to add more emphasis on the effect of Treatment on the Treated in this analysis. This let us test specific questions about changes in scores between two specific points in time. Further exploring this approach would likely be worthwhile. We also explored possible differences in ISS Assignment conditions between regions. In most cases, the region seemed to show some influence on our measurement outcomes. Future models should account for these regional differences.

Question 5: Does ISS result in improved parenting skills over the service period?

> The key takeaways for this question include:
>
> - ISS did improve parenting skills in general.
> - In particular, ISS lead to significant improvements in parental skills that were in the inadequate range at baseline.
> - The one notable exception was a significant deterioration in "Parental Cooperation with Case Planning/Services".

Methods

The analysis for this question utilizes CWBS data to evaluate change in constructs of parenting skills and knowledge. Descriptive statistics (mean (SD)) are reported for each item at each assessment point. For guidance, the specific responses scale options for each item are reported in each table to the right. Table 189 gives the Overall (all regions combined) results while Table 190– Table 193 give the results for Regions 1, 2, 3 and 5, respectively. Lastly, Table 194 again gives the overall results (all regions combined) but restricted to those participants that were seen at all four time-points.

For many items, the observed pre-post changes are small, which can be partly explained by floor effects occurring at the baseline assessment. The providers rated most clients as high functioning (top score 1.0) at their first assessment opportunity with all items averaging above 2.0 at Baseline for the Overall results.

Comparisons are made between successive pairwise time-points (significance indicated with *) with a final comparison between the first (Baseline) and last (6 month) time points (significance indicated with †) to access overall change.

Most measures, with some exceptions, showed improvement under both ISS (Baseline to Stepdown) and CHBS (Stepdown to 6 months), often significant improvements ($P < 0.05$). The one stand-out exception was parental cooperation with the program that was significantly worsening (increasing score) under both ISS and CHBS and overall. This also showed significant differences between ISS and CHBS at Stepdown (at 6 weeks).

Region 3, which accounts for 55% of the Overall count, showed similar results and trends to those described for the Overall. Regions 1, 2 and 5 showed somewhat similar results and trends, however, the smaller numbers of clients reported for these regions resulted in wider variations in the item averages and fluctuations from the overall trends noted. While caution is still needed before claiming all regions show similar results and trends, the general overall pattern does seem to be similar.

TABLE 189: AVERAGE CWBS SCORES FOR PARENTING SKILLS AND KNOWLEDGE (OVERALL)

	ISS		CHBS			Response Scale				
	Baseline (n = 389)	Stepdown (n = 285)	Stepdown (n = 228)	6 month (n = 221)		1	2	3	4	5
Overcrowding	1.57 (0.78)	1.50 (0.71)**	1.40 (0.70)	1.31 (0.68)**	†††	No overcrowding	Mild overcrowding	Moderate overcrowding	Serious overcrowding	
Household Sanitation	1.57 (0.83)	1.41 (0.64)***	1.42 (0.80)	1.20 (0.42)**	†††	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	Severely inadequate
Home Safety/ Access Hazards	1.74 (0.82)	1.62 (0.69)**	1.54 (0.77)	1.20 (0.45)***	†††	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	Severely inadequate
Clothing & Hygiene	1.08 (0.30)	1.08 (0.33)	1.09 (0.29)	1.04 (0.20)		Clean and adequate	Inadequate but no limitations	Limiting child function	Risk of illness	Causing illness
Food/ Nutrition	1.14 (0.58)	1.08 (0.46)	1.08 (0.28)	1.07 (0.40)		Regular meals	Irregular but no limitations	Limiting child function	Risk of malnutrition	Causing malnutrition
Physical Healthcare	1.23 (0.64)	1.11 (0.40)*	1.09 (0.37)	1.02 (0.19)*	††	Basic care provided	No preventive care provided	No care for illness/injury	Limiting child function	Risk severe impairment
Mental Healthcare	1.30 (0.60)	1.23 (0.44)	1.18 (0.43)	1.08 (0.29)***	††	Emotionally responsive	Inconsistent response	No care for problems	Unaddressed severe problems	
Development/ Education	1.19 (0.43)	1.13 (0.35)	1.14 (0.36)	1.05 (0.22)**	††	Needs met	Inconsistently meeting needs	No care for problems	Risk severe delays	
Money Management	1.67 (0.82)	1.76 (0.85)	1.47 (0.67)**	1.34 (0.62)	††	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Parental Cooperation	1.12 (0.38)	1.34 (0.70)***	1.23 (0.51)***	1.33 (0.68)**	†††	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Supervision	1.33 (0.66)	1.25 (0.53)	1.18 (0.42)	1.10 (0.36)*	†	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Positive interaction	1.37 (0.60)	1.34 (0.55)	1.35 (0.51)	1.23 (0.48)**	†	Very accepting & affectionate	Fairly accepting & affectionate	Not affectionate	Rejecting or hostile	
Discipline	1.48 (0.77)	1.30 (0.54)*	1.31 (0.58)	1.12 (0.39)***		Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Clear rules/ Limit Setting	1.64 (0.72)	1.70 (0.69)	1.53 (0.62)	1.21 (0.46)***	†††	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Parent Expectations	1.49 (0.64)	1.50 (0.59)	1.51 (0.63)	1.23 (0.51)***	†	Very realistic	Somewhat unrealistic	Unrealistic	Very unrealistic	
Discipline Consistency	1.79 (0.74)	1.85 (0.69)	1.71 (0.73)	1.34 (0.55)***	†††	High consistency	Moderate consistency	Marginal consistency	Low consistency	

Mean (SD); * p < 0.05, ** p < 0.01, *** p < 0.001; indications under ISS:Stepdown are for ISS:Baseline vs ISS:Stepdown, under CHBS:Stepdown are for ISS:Stepdown vs CHBS:Stepdown (comparison of raters), under CHBS:6 month are for CHBS:Stepdown vs CHBS:6 month† p < 0.05, †† p < 0.01, ††† p < 0.001; ISS:Baseline vs CHBS:6 month.

TABLE 190: AVERAGE CWBS SCORES FOR PARENTING SKILLS AND KNOWLEDGE (REGION 1)

	ISS		CHBS			Response Scale				
	Baseline (n = 68)	Stepdown (n = 56)	Stepdown (n = 37)	6 month (n = 35)		1	2	3	4	5
Overcrowding	1.48 (0.74)	1.33 (0.58)	1.29 (0.62)	1.25 (0.59)	†	No overcrowding	Mild overcrowding	Moderate overcrowding	Serious overcrowding	
Household Sanitation	1.63 (0.81)	1.61 (0.79)	1.77 (1.00)	1.21 (0.42)*	†	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	Severely inadequate
Home Safety/ Access Hazards	2.08 (0.84)	2.04 (0.82)	1.74 (0.89)	1.18 (0.39)*	†††	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	Severely inadequate
Clothing & Hygiene	1.08 (0.35)	1.11 (0.39)	1.15 (0.37)	1.04 (0.20)		Clean and adequate	Inadequate but no limitations	Limiting child function	Risk of illness	Causing illness
Food/ Nutrition	1.04 (0.21)	1.11 (0.40)	1.14 (0.36)	1.04 (0.20)		Regular meals	Irregular but no limitations	Limiting child function	Risk of malnutrition	Causing malnutrition
Physical Healthcare	1.22 (0.84)	1.00 (0.00)	1.04 (0.19)	1.00 (0.00)		Basic care provided	No preventive care provided	No care for illness/injury	Limiting child function	Risk severe impairment
Mental Healthcare	1.32 (0.65)	1.14 (0.36)	1.10 (0.31)	1.00 (0.00)		Emotionally responsive	Inconsistent response	No care for problems	Unaddressed severe problems	
Development/ Education	1.25 (0.49)	1.21 (0.41)	1.11 (0.31)	1.04 (0.20)	†	Needs met	Inconsistently meeting needs	No care for problems	Risk severe delays	
Money Management	1.37 (0.66)	1.58 (0.64)	1.49 (0.61)	1.34 (0.55)		Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Parental Cooperation	1.06 (0.30)	1.36 (0.77)**	1.16 (0.44)	1.14 (0.35)	†	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Supervision	1.42 (0.72)	1.29 (0.52)	1.17 (0.38)	1.04 (0.20)	†	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Positive interaction	1.37 (0.61)	1.42 (0.61)	1.30 (0.53)	1.22 (0.51)		Very accepting & affectionate	Fairly accepting & affectionate	Not affectionate	Rejecting or hostile	
Discipline	1.60 (0.99)	1.25 (0.55)	1.24 (0.44)	1.00 (0.00)**		Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Clear rules/ Limit Setting	1.84 (0.90)	1.83 (0.83)	1.41 (0.50)	1.19 (0.49)	†	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Parent Expectations	1.58 (0.64)	1.59 (0.70)	1.41 (0.57)	1.15 (0.37)	†	Very realistic	Somewhat unrealistic	Unrealistic	Very unrealistic	
Discipline Consistency	2.00 (0.65)	2.00 (0.67)	1.50 (0.69)	1.38 (0.64)	†	High consistency	Moderate consistency	Marginal consistency	Low consistency	

Mean (SD); * p < 0.05, ** p < 0.01, *** p < 0.001; indications under ISS:Stepdown are for ISS:Baseline vs ISS:Stepdown, under CHBS:Stepdown are for ISS:Stepdown vs CHBS:Stepdown (comparison of raters), under CHBS:6 month are for CHBS:Stepdown vs CHBS:6 month† p < 0.05, †† p < 0.01, ††† p < 0.001; ISS:Baseline vs CHBS:6 month.

TABLE 191: AVERAGE CWBS SCORES FOR PARENTING SKILLS AND KNOWLEDGE (REGION 2)

	ISS		CHBS			Response Scale				
	Baseline (n = 52)	Stepdown (n = 47)	Stepdown (n = 33)	6 month (n = 25)		1	2	3	4	5
Overcrowding	1.68 (0.88)	1.79 (0.91)	1.50 (0.80)*	1.26 (0.73)*		No overcrowding	Mild overcrowding	Moderate overcrowding	Serious overcrowding	
Household Sanitation	1.84 (0.90)	1.55 (0.69)*	1.58 (1.00)	1.26 (0.45)	†	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	Severely inadequate
Home Safety/ Access Hazards	2.08 (0.80)	1.78 (0.60)*	1.69 (0.90)	1.21 (0.42)*	†††	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	Severely inadequate
Clothing & Hygiene	1.03 (0.17)	1.10 (0.37)	1.10 (0.30)	1.00 (0.00)		Clean and adequate	Inadequate but no limitations	Limiting child function	Risk of illness	Causing illness
Food/ Nutrition	1.24 (0.96)	1.12 (0.64)	1.10 (0.30)	1.06 (0.24)		Regular meals	Irregular but no limitations	Limiting child function	Risk of malnutrition	Causing malnutrition
Physical Healthcare	1.26 (0.74)	1.12 (0.33)	1.03 (0.18)	1.00 (0.00)		Basic care provided	No preventive care provided	No care for illness/injury	Limiting child function	Risk severe impairment
Mental Healthcare	1.46 (0.65)	1.32 (0.47)	1.19 (0.48)	1.11 (0.32)		Emotionally responsive	Inconsistent response	No care for problems	Unaddressed severe problems	
Development/ Education	1.10 (0.31)	1.15 (0.37)	1.13 (0.34)	1.06 (0.24)		Needs met	Inconsistently meeting needs	No care for problems	Risk severe delays	
Money Management	1.80 (0.72)	1.78 (0.85)	1.28 (0.45)**	1.22 (0.55)	†††	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Parental Cooperation	1.07 (0.25)	1.36 (0.74)*	1.18 (0.46)*	1.26 (0.65)		Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Supervision	1.19 (0.40)	1.16 (0.37)	1.17 (0.46)	1.11 (0.47)		Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Positive interaction	1.55 (0.72)	1.35 (0.57)	1.30 (0.53)	1.11 (0.32)*		Very accepting & affectionate	Fairly accepting & affectionate	Not affectionate	Rejecting or hostile	
Discipline	1.70 (0.80)	1.32 (0.60)	1.20 (0.41)	1.24 (0.66)		Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Clear rules/ Limit Setting	1.64 (0.73)	1.73 (0.69)	1.31 (0.47)**	1.17 (0.51)		Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Parent Expectations	1.63 (0.74)	1.43 (0.50)	1.43 (0.79)	1.11 (0.47)*		Very realistic	Somewhat unrealistic	Unrealistic	Very unrealistic	
Discipline Consistency	1.73 (0.80)	1.88 (0.65)	1.48 (0.59)	1.22 (0.55)*		High consistency	Moderate consistency	Marginal consistency	Low consistency	

Mean (SD); * p < 0.05, ** p < 0.01, *** p < 0.001; indications under ISS:Stepdown are for ISS:Baseline vs ISS:Stepdown, under CHBS:Stepdown are for ISS:Stepdown vs CHBS:Stepdown (comparison of raters), under CHBS:6 month are for CHBS:Stepdown vs CHBS:6 month† p < 0.05, †† p < 0.01, ††† p < 0.001; ISS:Baseline vs CHBS:6 month.

	ISS		CHBS			Response Scale				
	Baseline (n = 206)	Stepdown (n = 158)	Stepdown (n = 130)	6 month (n = 128)		1	2	3	4	5
Overcrowding	1.65 (0.77)	1.46 (0.64)***	1.40 (0.66)	1.26 (0.58)**	††	No overcrowding	Mild overcrowding	Moderate overcrowding	Serious overcrowding	
Household Sanitation	1.49 (0.81)	1.30 (0.53)*	1.29 (0.64)	1.17 (0.41)	††	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	Severely inadequate
Home Safety/ Access Hazards	1.54 (0.76)	1.40 (0.58)	1.44 (0.71)	1.18 (0.44)***	†††	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	Severely inadequate
Clothing & Hygiene	1.05 (0.23)	1.08 (0.31)	1.07 (0.26)	1.05 (0.21)		Clean and adequate	Inadequate but no limitations	Limiting child function	Risk of illness	Causing illness
Food/ Nutrition	1.14 (0.52)	1.07 (0.42)*	1.04 (0.21)	1.08 (0.50)		Regular meals	Irregular but no limitations	Limiting child function	Risk of malnutrition	Causing malnutrition
Physical Healthcare	1.25 (0.58)	1.13 (0.47)	1.11 (0.46)	1.04 (0.24)	†	Basic care provided	No preventive care provided	No care for illness/injury	Limiting child function	Risk severe impairment
Mental Healthcare	1.27 (0.61)	1.24 (0.45)	1.16 (0.40)	1.11 (0.35)*		Emotionally responsive	Inconsistent response	No care for problems	Unaddressed severe problems	
Development/ Education	1.18 (0.43)	1.10 (0.34)	1.11 (0.35)	1.05 (0.22)	†	Needs met	Inconsistently meeting needs	No care for problems	Risk severe delays	
Money Management	1.71 (0.84)	1.79 (0.88)	1.50 (0.74)	1.34 (0.62)	†	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Parental Cooperation	1.11 (0.35)	1.34 (0.66)***	1.25 (0.53)*	1.41 (0.75)**	†††	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Supervision	1.35 (0.73)	1.29 (0.61)	1.22 (0.45)	1.14 (0.41)		Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Positive interaction	1.32 (0.57)	1.28 (0.50)	1.39 (0.51)*	1.29 (0.53)		Very accepting & affectionate	Fairly accepting & affectionate	Not affectionate	Rejecting or hostile	
Discipline	1.36 (0.69)	1.30 (0.53)	1.30 (0.51)	1.16 (0.40)*		Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Clear rules/ Limit Setting	1.59 (0.68)	1.66 (0.66)	1.63 (0.68)	1.24 (0.46)***	†	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Parent Expectations	1.40 (0.57)	1.48 (0.58)	1.55 (0.56)	1.30 (0.58)**		Very realistic	Somewhat unrealistic	Unrealistic	Very unrealistic	
Discipline Consistency	1.83 (0.77)	1.78 (0.71)	1.81 (0.75)	1.34 (0.53)***	††	High consistency	Moderate consistency	Marginal consistency	Low consistency	

Mean (SD); * p < 0.05, ** p < 0.01, *** p < 0.001; indications under ISS:Stepdown are for ISS:Baseline vs ISS:Stepdown, under CHBS:Stepdown are for ISS:Stepdown vs CHBS:Stepdown (comparison of raters), under CHBS:6 month are for CHBS:Stepdown vs CHBS:6 month† p < 0.05, †† p < 0.01, ††† p < 0.001; ISS:Baseline vs CHBS:6 month.

	ISS		CHBS		Response Scale				
	Baseline (n = 51)	Stepdown (n = 18)	Stepdown (n = 23)	6 month (n = 30)	1	2	3	4	5
Overcrowding	1.31 (0.64)	1.53 (0.74)	1.37 (0.83)	1.50 (1.04)	No overcrowding	Mild overcrowding	Moderate overcrowding	Serious overcrowding	
Household Sanitation	1.51 (0.84)	1.21 (0.58)	1.33 (0.73)	1.28 (0.46)	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	Severely inadequate
Home Safety/ Access Hazards	1.73 (0.82)	1.43 (0.51)	1.55 (0.69)	1.33 (0.59)*	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	Severely inadequate
Clothing & Hygiene	1.21 (0.50)	1.00 (0.00)	1.11 (0.32)	1.06 (0.25)	Clean and adequate	Inadequate but no limitations	Limiting child function	Risk of illness	Causing illness
Food/ Nutrition	1.20 (0.65)	1.00 (0.00)	1.18 (0.39)	1.06 (0.25)	Regular meals	Irregular but no limitations	Limiting child function	Risk of malnutrition	Causing malnutrition
Physical Healthcare	1.14 (0.36)	1.17 (0.58)	1.11 (0.32)	1.00 (0.00)	Basic care provided	No preventive care provided	No care for illness/injury	Limiting child function	Risk severe impairment
Mental Healthcare	1.20 (0.41)	1.20 (0.42)	1.39 (0.61)	1.00 (0.00)*	Emotionally responsive	Inconsistent response	No care for problems	Unaddressed severe problems	
Development/ Education	1.23 (0.43)	1.00 (0.00)	1.33 (0.49)	1.06 (0.25)*	Needs met	Inconsistently meeting needs	No care for problems	Risk severe delays	
Money Management	1.84 (0.99)	2.08 (1.12)	1.56 (0.63)	1.47 (0.83)	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Parental Cooperation	1.26 (0.49)	1.24 (0.75)	1.25 (0.55)	1.32 (0.67)	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Supervision	1.25 (0.53)	1.00 (0.00)	1.11 (0.32)	1.00 (0.00)	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Positive interaction	1.38 (0.55)	1.54 (0.66)	1.26 (0.45)	1.12 (0.33)*	Very accepting & affectionate	Fairly accepting & affectionate	Not affectionate	Rejecting or hostile	
Discipline	1.56 (0.70)	1.33 (0.50)	1.91 (1.22)	1.07 (0.27)*	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Clear rules/ Limit Setting	1.68 (0.72)	1.44 (0.53)	1.60 (0.70)	1.15 (0.38)* †	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Parent Expectations	1.64 (0.81)	1.55 (0.69)	1.50 (0.79)	1.13 (0.35)	Very realistic	Somewhat unrealistic	Unrealistic	Very unrealistic	
Discipline Consistency	1.50 (0.62)	2.00 (0.67)	2.00 (0.82)	1.36 (0.50)*	High consistency	Moderate consistency	Marginal consistency	Low consistency	

Mean (SD); * p < 0.05, ** p < 0.01, *** p < 0.001; indications under ISS:Stepdown are for ISS:Baseline vs ISS:Stepdown, under CHBS:Stepdown are for ISS:Stepdown vs CHBS:Stepdown (comparison of raters), under CHBS:6 month are for CHBS:Stepdown vs CHBS:6 month† p < 0.05, †† p < 0.01, ††† p < 0.001; ISS:Baseline vs CHBS:6 month.

TABLE 194: AVERAGE CWBS SCORES FOR PARENTING SKILLS AND KNOWLEDGE (OVERALL). INCLUDING ONLY THOSE CAREGIVERS SEEN AT ALL FOUR TIME-POINTS.

	ISS		CHBS			Response Scale				
	Baseline (n = 158)	Stepdown (n = 158)	Stepdown (n = 158)	6 month (n = 158)		1	2	3	4	5
Overcrowding	1.60 (0.78)	1.50 (0.69)	1.43 (0.74)	1.30 (0.69)**	††	No overcrowding	Mild overcrowding	Moderate overcrowding	Serious overcrowding	
Household Sanitation	1.56 (0.85)	1.32 (0.58)**	1.41 (0.75)	1.21 (0.41)*	†††	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	Severely inadequate
Home Safety/ Access Hazards	1.73 (0.81)	1.56 (0.63)*	1.51 (0.72)	1.19 (0.46)***	†††	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	Severely inadequate
Clothing & Hygiene	1.06 (0.28)	1.07 (0.29)	1.07 (0.26)	1.04 (0.19)		Clean and adequate	Inadequate but no limitations	Limiting child function	Risk of illness	Causing illness
Food/ Nutrition	1.15 (0.63)	1.06 (0.39)	1.09 (0.28)	1.06 (0.27)		Regular meals	Irregular but no limitations	Limiting child function	Risk of malnutrition	Causing malnutrition
Physical Healthcare	1.24 (0.71)	1.05 (0.22)**	1.07 (0.37)	1.03 (0.21)	†	Basic care provided	No preventive care provided	No care for illness/injury	Limiting child function	Risk severe impairment
Mental Healthcare	1.33 (0.67)	1.21 (0.41)	1.21 (0.47)	1.08 (0.31)***	††	Emotionally responsive	Inconsistent response	No care for problems	Unaddressed severe problems	
Development/ Education	1.17 (0.40)	1.06 (0.23)**	1.15 (0.38)*	1.06 (0.25)*	††	Needs met	Inconsistently meeting needs	No care for problems	Risk severe delays	
Money Management	1.66 (0.81)	1.63 (0.75)	1.47 (0.66)*	1.34 (0.63)*	††	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Parental Cooperation	1.03 (0.18)	1.08 (0.29)	1.18 (0.43)**	1.34 (0.69)***	†††	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Supervision	1.32 (0.62)	1.21 (0.48)*	1.21 (0.45)	1.11 (0.40)	†	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Positive interaction	1.43 (0.61)	1.30 (0.50)*	1.36 (0.51)	1.21 (0.47)**	†	Very accepting & affectionate	Fairly accepting & affectionate	Not affectionate	Rejecting or hostile	
Discipline	1.45 (0.74)	1.28 (0.53)*	1.35 (0.62)	1.14 (0.40)**		Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Clear rules/ Limit Setting	1.63 (0.73)	1.65 (0.65)	1.53 (0.62)	1.20 (0.45)***	††	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Parent Expectations	1.44 (0.58)	1.45 (0.54)	1.56 (0.66)	1.19 (0.48)***	†	Very realistic	Somewhat unrealistic	Unrealistic	Very unrealistic	
Discipline Consistency	1.93 (0.80)	1.80 (0.72)**	1.76 (0.75)	1.31 (0.53)***	†††	High consistency	Moderate consistency	Marginal consistency	Low consistency	

Mean (SD); * p < 0.05, ** p < 0.01, *** p < 0.001; indications under ISS:Stepdown are for ISS:Baseline vs ISS:Stepdown, under CHBS:Stepdown are for ISS:Stepdown vs CHBS:Stepdown (comparison of raters), under CHBS:6 month are for CHBS:Stepdown vs CHBS:6 month† p < 0.05, †† p < 0.01, ††† p < 0.001; ISS:Baseline vs CHBS:6 month.

	Baseline (n = 389) [Table 33]	Baseline – then not seen again (n = 30)	Baseline – those in between (n = 202)	Baseline - seen four times (n = 158) [Table 38]	Response Scale				
					1	2	3	4	5
Overcrowding	1.57 (0.78)	1.57 (0.77)	1.55 (0.79)	1.60 (0.78)	No overcrowding	Mild overcrowding	Moderate overcrowding	Serious overcrowding	
Household Sanitation	1.57 (0.83)	1.38 (0.78)	1.62 (0.82)	1.56 (0.85)	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	Severely inadequate
Home Safety/ Access Hazards	1.74 (0.82)	1.62 (0.86)	1.78 (0.83)	1.73 (0.81)	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	Severely inadequate
Clothing & Hygiene	1.08 (0.30)	1.15 (0.37)	1.08 (0.30)	1.06 (0.28)	Clean and adequate	Inadequate but no limitations	Limiting child function	Risk of illness	Causing illness
Food/ Nutrition	1.14 (0.58)	1.38 (0.81)	1.09 (0.47)	1.15 (0.63)	Regular meals	Irregular but no limitations	Limiting child function	Risk of malnutrition	Causing malnutrition
Physical Healthcare	1.23 (0.64)	1.33 (0.97)	1.19 (0.49)	1.24 (0.71)	Basic care provided	No preventive care provided	No care for illness/injury	Limiting child function	Risk severe impairment
Mental Healthcare	1.30 (0.60)	1.29 (0.47)	1.26 (0.54)	1.33 (0.67)	Emotionally responsive	Inconsistent response	No care for problems	Unaddressed severe problems	
Development/ Education	1.19 (0.43)	1.28 (0.46)	1.19 (0.45)	1.17 (0.40)	Needs met	Inconsistently meeting needs	No care for problems	Risk severe delays	
Money Management	1.67 (0.82)	1.89 (0.88)	1.64 (0.82)	1.66 (0.81)	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Parental Cooperation	1.12 (0.38)	1.48 (0.72)	1.14 (0.38)	1.03 (0.18)	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Supervision	1.33 (0.66)	1.26 (0.65)	1.35 (0.71)	1.32 (0.62)	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Positive interaction	1.37 (0.60)	1.52 (0.81)	1.28 (0.53)	1.43 (0.61)	Very accepting & affectionate	Fairly accepting & affectionate	Not affectionate	Rejecting or hostile	
Discipline	1.48 (0.77)	2.00 (1.18)	1.41 (0.66)	1.45 (0.74)	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Clear rules/ Limit Setting	1.64 (0.72)	1.69 (0.85)	1.65 (0.69)	1.63 (0.73)	Appropriate	Mildly inadequate	Moderately inadequate	Seriously inadequate	
Parent Expectations	1.49 (0.64)	1.78 (0.88)	1.50 (0.65)	1.44 (0.58)	Very realistic	Somewhat unrealistic	Unrealistic	Very unrealistic	
Discipline Consistency	1.79 (0.74)	1.73 (0.79)	1.63 (0.64)	1.93 (0.80)	High consistency	Moderate consistency	Marginal consistency	Low consistency	

Mean (SD)

The following Figures show the trajectories of caregiver circumstances through the program. To explain these figures, imagine the participants stacked in each column categorized (1 through 4 or 5) by their response to the question at that time-point. A ribbon connects each caregiver from the left-side of their response at one time-point to the right-side of their response at the next time-point. If there is no ribbon from the left-side of their response at one time-point then that caregiver did not give a response in the next time-point. For example, in "Overcrowding", the 100th caregiver in the "Baseline" column has a "No overcrowding" response followed by "No overcrowding" at Stepdown (ISS), as shown by the ribbon running straight across, "Moderate overcrowding" at "Stepdown (CHBS)", with the ribbon flowing up to the green category, "and "No overcrowding" at "Six Month", with the ribbon flowing back down to

the purple category. However, this figure is more useful at tracking overall trends rather than following individual trajectories. And, most families showed improvement overall, with just a few families experiencing worsening conditions.

FIGURE 59: OVERCROWDING
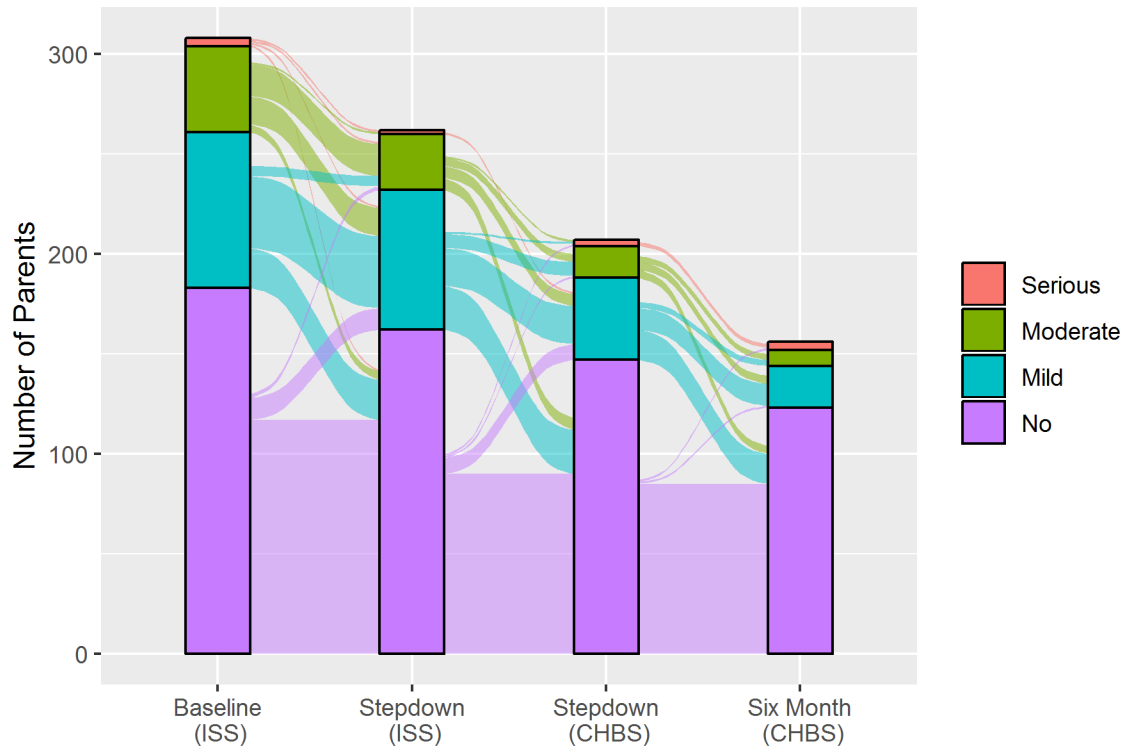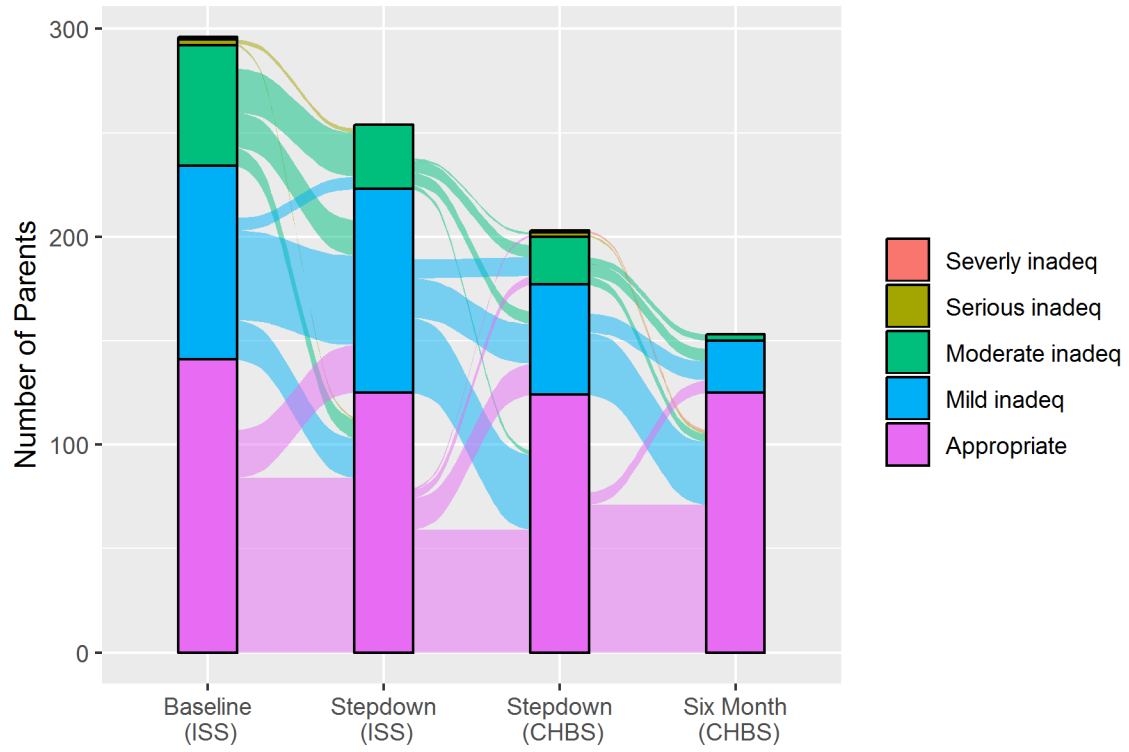
FIGURE 61: CLOTHING AND HYGIENE
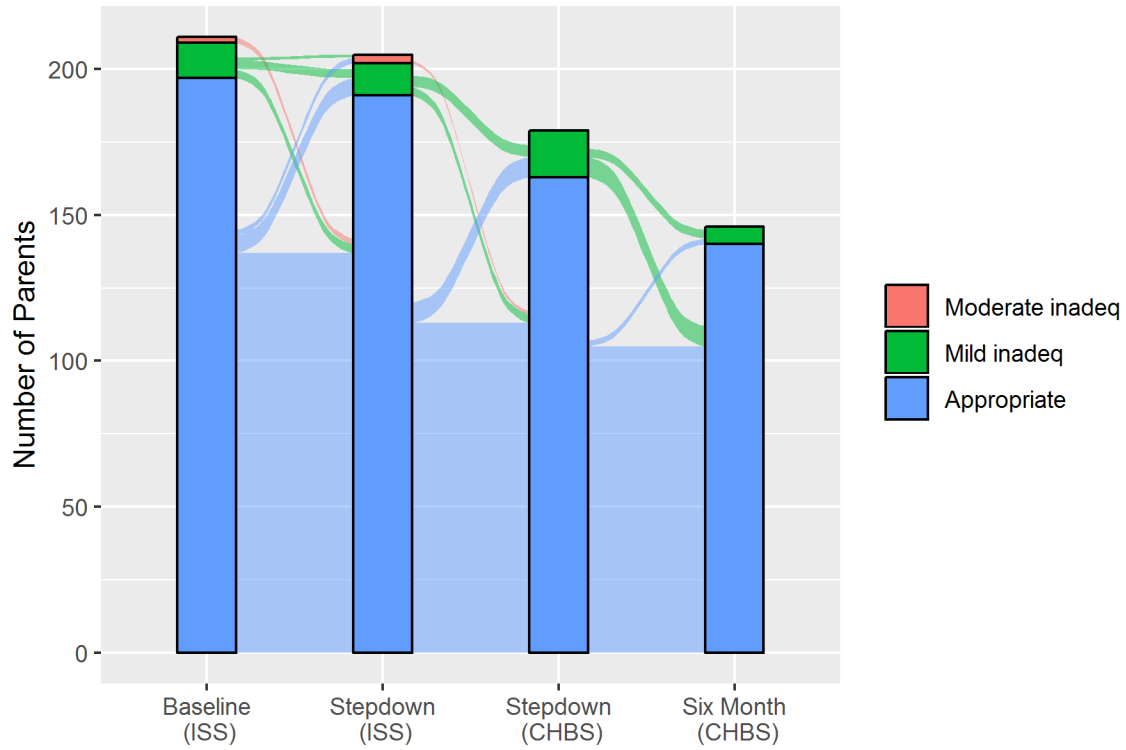
FIGURE 62: PARENTAL COOPERATION WITH CASE PLANNING/SERVICES
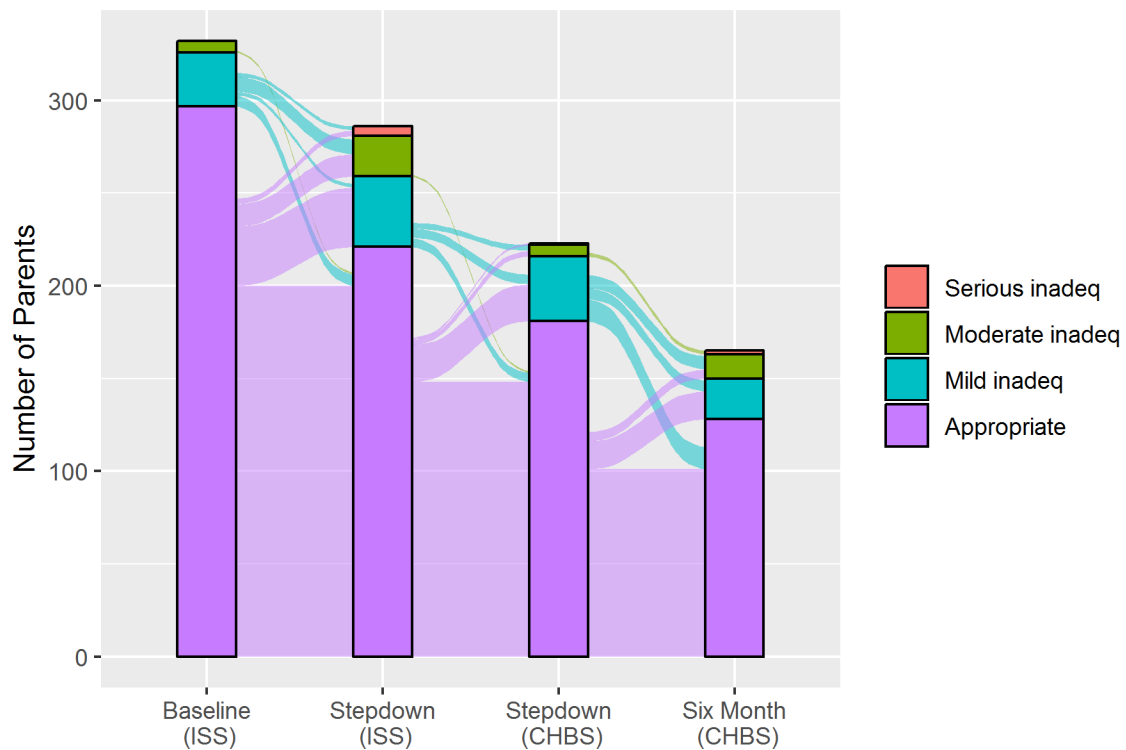
While for Figure 59 - Figure 61 there does not appear to be a bias with those with poorer scores leaving the program (for whatever reason), such a bias is suggested in Figure 62, especially in the changeover from ISS to CHBS. The latter is not surprising given that the measure is of parental cooperation.

Question 6: Does ISS result in fewer numbers of children entering out-of-home care and/or fewer days in placement relative to the SAU condition?

The key takeaways for this question include:

- The ISS group had a significant reduction in removals both on the first referral and on subsequent referrals.
- The duration of the child remaining in the home was longer for both the duration associated with the first referral and the duration associated with subsequent referrals.
- Although these results were more pronounced in TOT analyses, they were also observed in ITT.
- ISS did not reunify families quicker once a child was removed.

Question 6 can be conceptualized as eight hypotheses that are evaluated by different models (the eight are comprised of four pairs). In short, ISS was found to be significantly better for the first two pairs, while no differences were found on the third and fourth pair.

Within each pair, the first model addresses the child's *first* referral assigned to ISS/SAU. The pair's second model addresses *all* of the child's referrals since the assignment. The approximate pattern was consistent across regions, and is not shown below. Each model was duplicated (for a total of 4 x 2 x 2 analyses) to examine the effect of treatment received (*i.e.*, ISS received *vs* ISS not workable *vs* SAU) and the effect of intent-to-treat (ITT, *i.e.*, ISS *vs* SAU).
   a. Does ISS reduce the eventual likelihood of removal?
      1. On the first referral: Yes
      2. On all referrals: Yes
   b. Does ISS lengthen the duration in the home?
      1. associated with the first referral: Yes
      2. associated with all referrals (until the first removal): Yes
   c. If a child is removed, does ISS increase the likelihood of reunification?
      1. associated with the first referral: No
      2. On all referrals: No
   d. Once a child is removed, does ISS reunify families quicker?
      1. On the first referral: No -SAU is actually quicker.
      2. On all referrals: No -SAU is actually quicker.

Methods

The data used to address questions 6 through 8 come from KIDS (Oklahoma's SACWIS) data captured primarily at the time of a child abuse/neglect referral and were collected from both assigned (ISS, SAU) groups. This analysis has a sample size of 12,779 referrals on 9,019 unique children. Each child was associated with an "initial referral" that activated the PREMISS eligibility model, which implies that another 3,760 referrals (12,779 – 9,019) were observed during the study follow-up period. Across all 12,779 referrals, children were removed in 5,239 instances (sometimes the same child was removed on more than one occasion). Among the 9,109 children assigned to ISS or SAU, referral dates range from

2015-06-09 to 2019-10-02. Removal dates range from 2015-07-08 to 2019-10-08.[27] Removal end dates range from 2015-08-28 to 2019-10-8. There have been 5,107 children ever removed. Because there is such a high percentage of right-censored removals[28], the interpretation of 'Length of Removal' outcomes should be qualified with respect to this report's follow-up end-date (2019-10-08).

With one exception, the unit of analysis for all graphs and tables is a "child" (not a referral or a removal). The lone exception is the 2nd bivariate table for Q7. Fortunately for the sake of simplicity, the two Q7 tables have very similar proportions (the counts are halved b/c there are roughly two children per referral).

Binary outcomes were analyzed with logistic regression (e.g., proportion reunified), and duration outcomes were analyzed with Cox survival analysis (e.g., duration removed).

Region-specific values are displayed in the end of the section in TABLE **206** through

---

[27] This is the one set of analysis that differs from the rest of the analyses in this report with a cut-off date of October 2, 2019.

[28] Meaning the full duration of the removal period is unknown (censored), because at the time of the report end-date, the child was still experiencing separation from the parents.

TABLE **209**. Binary outcomes were analyzed with logistic regression (*e.g.*, proportion reunified) and duration outcomes were analyzed with Cox survival analysis (*e.g.*, duration removed). For models in Q6, Q7, and Q8, the child's first PREMISS risk score was included as a controlling covariate; it was always significantly correlated with the undesirable outcome (*i.e.*, removal, non-reunification, and subsequent referrals and removals).

Before answering Q6 directly, observe the gray monthly trend in **FIGURE 63** which shows how the assignment to ISS has changed. In short, the initial implementation had an undesirably low percentage of children assigned to ISS (*i.e.*, 23% of the 13 potential children). The gray trajectory indicates the percentage assigned to ISS per month. A child assigned to ISS is a purple point, while an SAU assignment is gold.[29] A removed child is an 'x', while a child remaining in the home is an open dot. The point locations are jittered to reduce overplotting at the midpoint of each month.

FIGURE 63 - ASSIGNMENTS TO CONDITION BY REFERRAL DATE
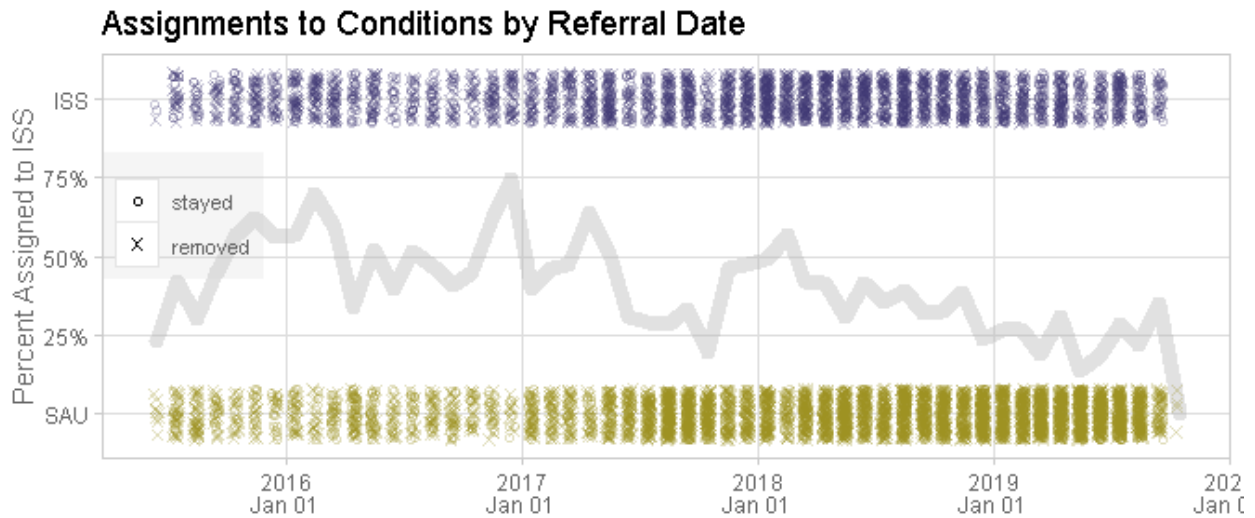
Q6a: Does ISS reduce the eventual likelihood of removal?
As seen in **FIGURE 64**, 18% of Oklahoma children receiving ISS services were removed, which is significantly better than (i) 63% of children assigned to ISS, but not receiving services and (ii) 57% of children assigned to SAU (both differences were significant, $p < .0001$). The proportions are almost identical when a child's later referrals were included, too (see the second row of **TABLE 196**). This pattern was also supported by a survival model, which incorporated the incomplete/censored cases. **FIGURE 64** shows the inferences from the *received* group, while **FIGURE 65** shows the inferences from the *assigned* group (ITT).

FIGURE 63, **FIGURE 64**, and **FIGURE 65** demonstrate the stark differences. In the marginal graph (for each condition), notice the "ISS received" bar has fewer removals than the other two bars. In the longitudinal

---

[29] As described above, the study design included an adaptive randomization procedure within each district/sub-district. The probability of ISS assignment fluctuated over time depending on capacity, with increasing ISS assignment probabilities for each additional open case slot. Also, rural regions (1, 2, 4) had lower total capacity (i.e., fewer providers), and as these came online, total case counts rose, but the overall ISS assignment percentage dropped.

graph, notice the purple "ISS received" line is regularly better/lower than the other two conditions. Each group's trajectory is represented by a line, and the points within each line represent the proportion of children removed among that month's referral.

Regarding ITT, the effect was still significant when considering if a child was only *assigned* to ISS (but not necessarily receiving services). When considering removals associated with the *first referral*, 47% of ISS children and 57% of SAU children were removed. When considering removals associated with *all referrals*, the percentage increased to 51% and 60%. Both differences were significant ($p < .0001$.)

FIGURE 64 – PROPORTION OF KIDS REMOVED ON INITIAL REFERRAL AFTER ASSIGNMENT (AS TREATED GROUPS)
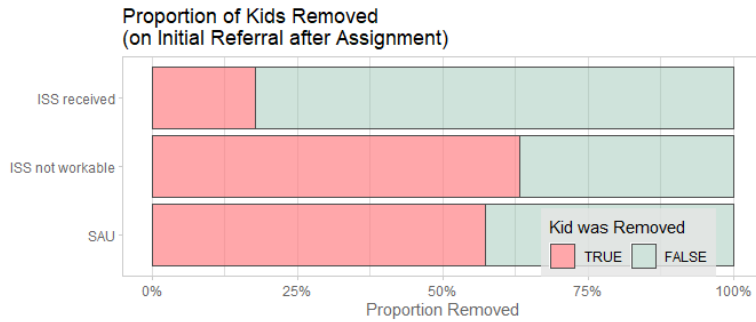
FIGURE 65 – PROPORTION OF KIDS REMOVED ON INITIAL REFERRAL AFTER ASSIGNMENT (ITT GROUPS)

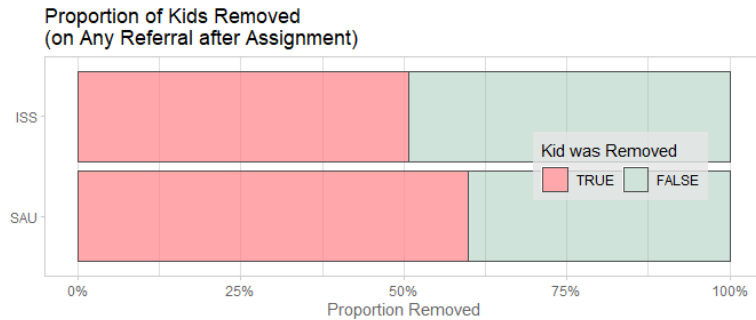Proportion of Kids Removed (on Initial Referral after Assignment)
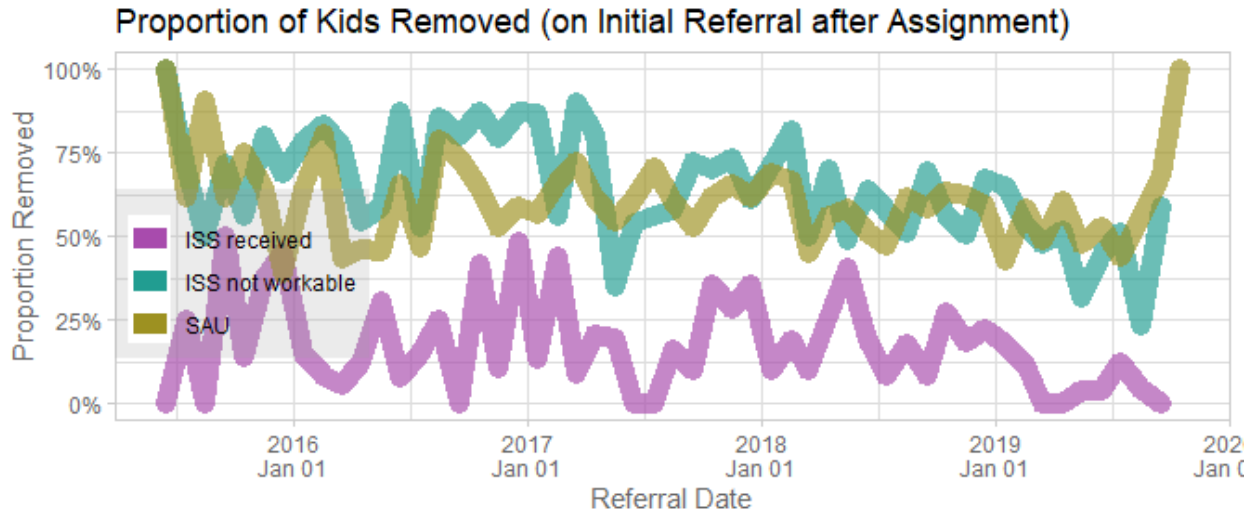
TABLE 196 - COUNT OF KIDS REMOVED

	ISS Received	ISS Not Workable	SAU
Removed on 1st Referral (*i.e.*, Q6a-1)	18% (211 of 1,193)	63% (1,315 of 2,079)	57% (3,297 of 5,747)
Removed on Any Referral (*i.e.*, Q6a-2)	24% (291 of 1,193)	66% (1,372 of 2,079)	60% (3,444 of 5,747)

TABLE 197 - LOGISTIC REGRESSION OF 1ST REMOVAL: COMPARING 'ISS RECEIVED' AGAINST 'ISS NOT WORKABLE' AND 'SAU'. THE FIRST ROW ESTIMATES THE INTERCEPT OF THE "ISS RECEIVED" GROUP, WHILE THE NEXT TWO ROWS ESTIMATE THE GROUP'S OFFSET FROM "ISS RECEIVED". THE BOTTOM ROW CONTROLS FOR THE PRE-TREATMENT RISK.

	Estimate	Std. Error	z value	p
Intercept	-2.33	0.11	-20.8	<.0001
Group Actual: ISS Not Workable	2.25	0.10	20.6	<.0001
Group Actual: SAU	2.06	0.09	20.6	<.0001
PREMISS Risk Score	-0.52	0.03	-14.0	<.0001

TABLE 198 - ITT LOGISTIC REGRESSION OF 1ST REMOVAL: COMPARING 'ISS' (IRRESPECTIVE IF SERVICES WERE RECEIVED) AGAINST AND 'SAU'. THE FIRST ROW ESTIMATES THE INTERCEPT OF THE "ISS ASSIGNED" GROUP, WHILE THE MIDDLE ROW ESTIMATES THE SAU'S OFFSET FROM "ISS ASSIGNED". THE BOTTOM ROW CONTROLS FOR THE PRE-TREATMENT RISK.

	Estimate	Std. Error	z value	p
Intercept	-0.75	0.06	-11.0	<.0001
Group Assignment SAU	0.48	0.05	9.3	<.0001
PREMISS Risk Score	-0.51	0.03	-14.4	<.0001

Q6b: Does ISS lengthen the duration in the home?
A survival analysis examines how long ISS-receiving clients remained in their home, compared to (i) children assigned to ISS, but not receiving services and (ii) children assigned to SAU. From this approach, "high survivability" is desirable, and represents longer duration remaining in their home. The ISS-receiving children were significantly better than the other two groups.

For a child's first referral and removal sequence, as seen in the survival plot (**FIGURE 67**), the percentage of remaining "ISS not workable" and "SAU" children quickly drops below 50% (20 days and 42 days for the two groups) while the ISS received children asymptote near 75%; recall from Q6a that 18% were removed (so the group never drops below 50%).

When considering all time spent in the home since a child's first referral (which includes after any reunifications they remain in the home), the pattern is similar but expectedly stretched longer. For example, compare TABLE 199 and TABLE 200.

Regarding ITT, the SAU Cox regression coefficient, an estimate of the log-transformed hazard ratio[30], was roughly 0.28 for both Q6b-1 and Q6b-2, which suggest significantly higher risk among SAU families relative to the ISS reference group (p < .0001).

Regarding TOT ([treatment on the treated]), [31] the SAU log hazard ratio was again unfavorable with estimates of 0.84 for Q6b-1 and .74 for Q6b-2 (p < .001 for both).

TABLE 199 - LENGTH IN HOME ASSOCIATED WITH FIRST REFERRAL

Group Actual	count	min	q25	median	mean	q75	sum
ISS Received	1,193	0	123	400	496	648	591,620
ISS Not Workable	2,079	0	4	20	217	206	451,181
SAU	5,747	0	4	42	209	252	1,199,383

TABLE 200 - LENGTH IN HOME ASSOCIATED WITH ALL REFERRALS

Group Actual	count	min	q25	median	mean	q75	sum
ISS Received	1,193	8	181	467	543	806	647,296
ISS Not norkable	2,079	0	32	303	379	596	787,274
SAU	5,747	0	27	198	317	509	1,824,482

---

[30] This is the ratio of hazard rates. A hazard rate is a measure of failure risk. Higher hazards reflect greater risk of failure. Log-tranformed hazard ratios comparing the SAU group to another group (e.g., ISS Received) reflect higher risk when positive (above zero) and lower risk when negative (below zero).

[31] For details of our instrumental variable survival model approach, see: Tchetgen, E.J., Walter, S., Vansteelandt, S., Martinussen, T., & Glymour, M. (2015). Instrumental variable estimation in a survival context. *Epidemiology, 26,* 402-10.
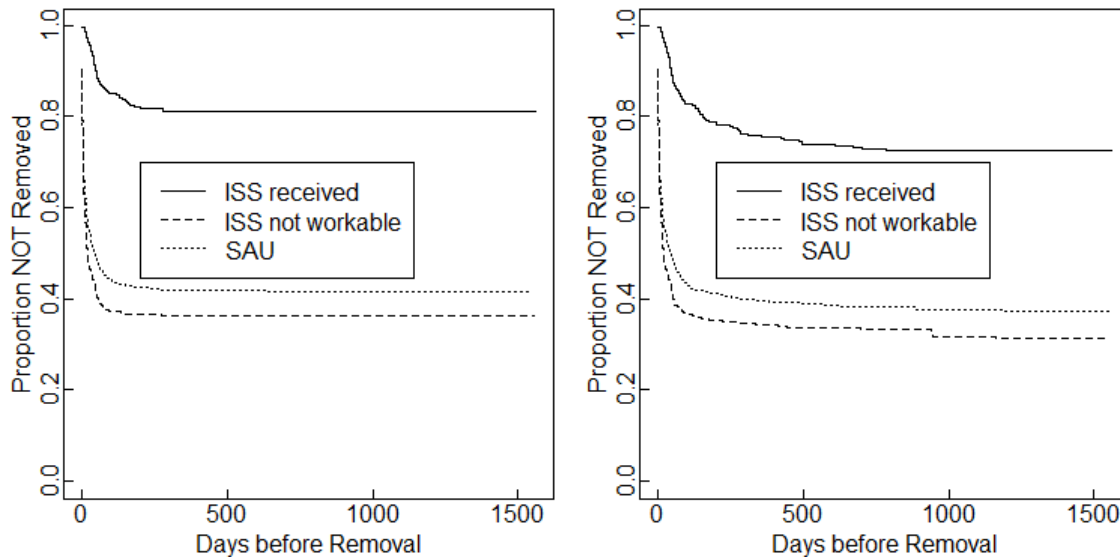
Q6c: If a child is removed, does ISS increase the likelihood of reunification?

Should a child experience a removal, there does not appear to be any extended benefits of prior ISS receipt. In Q6c (and subsequently in Q6d, Q7, and Q8) the three groups show no significant differences ($p > .24$). Restricting the samples to removed children only, 28% of the 211 "ISS received" children were reunified, 34% of the 1,315 "ISS not workable" children were reunified, and 29% of the 3,297 "SAU" children were reunified. **TABLE 201** summarizes the durations of the three groups (ignoring censoring and length of follow-up).

Regarding ITT, no significant differences existed between ISS and SAU for Q6c-1 and Q6c-2. SAU was roughly 11% (= 1 - $e^{-0.12}$) less likely to be reunified ($p >= .09$ for both).

TABLE 201 - COUNT OF KIDS REUNIFIED

	ISS Received	ISS Not Workable	SAU
Reunified on 1st Referral (*i.e.*, Q6c-1)	28% (59 of 211)	34% (441 of 1,315)	29% (969 of 3,297)
Reunified on Any Referral (*i.e.*, Q6c-2)	32% (92 of 291)	33% (458 of 1,372)	29% (1,013 of 3,444)

Q6d: Once a child is removed, does ISS reunify families quicker?

Q6d-1 involves the reunification of children whose removal was associated with the *initial* referral. The removal length of ISS received was significantly longer than SAU. A shorter duration is desirable, because it represents quicker reunification. ISS received was also longer than ISS not workable, but not significantly so.

Half of "ISS received" *removed* children were reunified by 923 days. In contrast, half of "SAU" children were reunified by 706 days (which is 7+ months sooner, $p<.001$). The "ISS not workable" group performed similarly to the "SAU" group; half were reunified by 753 days (almost 6 months sooner, $p<.001$). The longer ISS received days may be a function of underlying risk differences. Concentrating on removed children only, we ignore the randomized balance between ISS and SAU assigned cases. The same is true when isolating and comparing the ISS received subgroup, since only assignment, and not receipt, was randomized. It is possible that the removals experienced among the ISS received group occur for "riskier" (more complex) families compared to removals in the other two groups. In fact, this would be consistent with the foundational assumption of the PREMISS eligibility model: that a substantial portion of the SAU population will experience relatively quick reunification events and could possibly avoid removals with a more intensive front-end preventive service.

Q6d-2 involves the reunification of children whose removal was associated with *any* (not just the initial referral) referral. Q6d-2's pattern resembles Q6d-1's pattern, but is stretched longer. "SAU" is significantly different from "ISS received" ($p = .004$). "ISS not workable" survival is reasonably close to the "SAU" line. Yet, the survival model did not detect any significance between "ISS received" and "ISS not workable". We suspect this is because the "SAU" group was twice as large as the "ISS not workable" group (1,372 vs 3,444 children), so the model has greater statistical power to detect SAU differences from "ISS received" (which had only 291 removed children).

Regarding ITT, the SAU Cox regression coefficient[32] was roughly 0.14 for both Q6d-1 and Q6d-2, which suggests significantly quicker reunification compared to the ISS reference group ($p<.001$ for both comparisons).

TABLE 202 - DURATION REMOVED ASSOCIATED WITH FIRST REFERRAL

Group Actual	count	min	q25	median	mean	q75	sum
ISS Received	211	1	288	506	502	685	106,027
ISS Not Workable	1,315	0	271	429	456	626	599,776
SAU	3,297	0	131	300	346	503	1,139,928

TABLE 203 - DURATION REMOVED ASSOCIATED WITH ALL REFERRALS

Group Actual	count	min	q25	median	mean	q75	sum
ISS Received	291	1	259	459	475	684	138,211
ISS Not Workable	1,372	0	270	431	459	629	629,120
SAU	3,444	0	134	303	350	513	1,207,002

---

[32] In this analysis, the "hazard" rate term is a misnomer. Higher "risk" of reunification is desired, and can more aptly be characterized as higher likelihood of a successful reunification event.
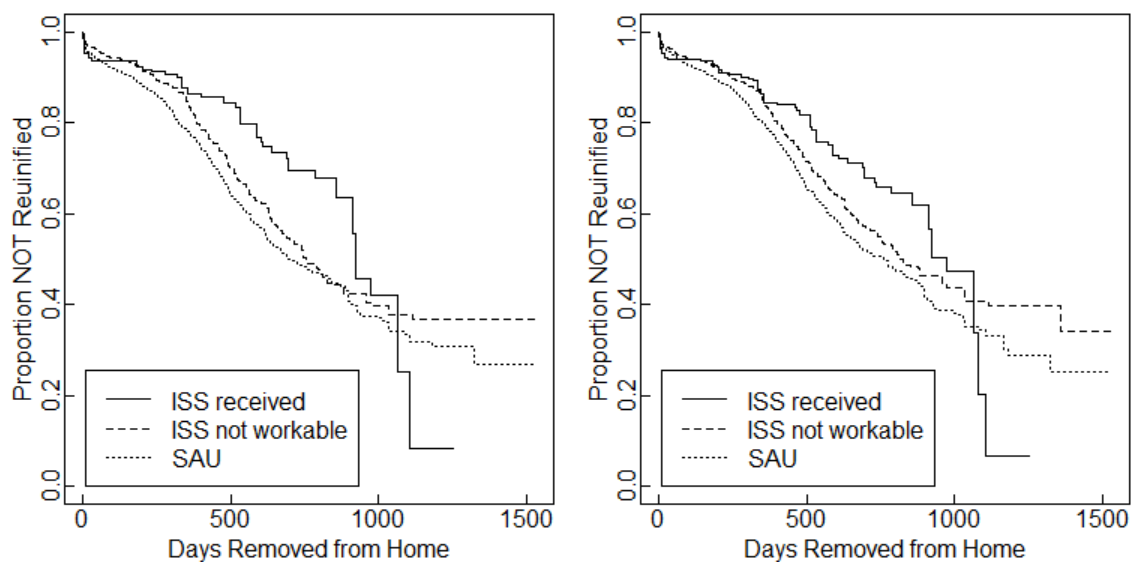
Question 7: Does ISS reduce the likelihood of an additional CPS referral?

The key takeaways for this question include:

- ISS was not found to reduce the likelihood of additional referrals.
- ISS was not found to reduce the likelihood of additional referrals following reunification.
- ISS was not found to increase the duration between initial and additional future referrals.

Future CPS referrals are examined from three related questions. Their results are also presented here.

7a. Does ISS reduce the likelihood of additional referrals? No, it actually is more likely.

7b. Does ISS reduce the likelihood of additional referrals after reunification? No, it actually is more likely.

7c. Does ISS increase the duration before additional referrals? No, it actually quickens the duration

Methods

The data and methods used to address question 7 were the same as those used in question 6. Please refer to Question 6 Method above for details.

Q7a: Does ISS reduce the likelihood of additional referrals?

For Q7a (and later Q8a), a client is *included* if they were either (in1) removed on the first referral and later reunited, or (in2) never removed at all; a client is *excluded* if they were (out1) removed on the first

212

referral and never reunited. For Q7b (and later Q8b), a client is *included* if they were (in1) removed on the first referral and later reunited; a client is *excluded* if they were either (out1) removed on the first referral and never reunited, or (out2) never removed at all.

The count of subsequent referrals (*i.e.*, referrals dated at least 30 days after the initially assigned referral) were compared between groups. As seen in

In the 48 months following the first ISS assignment, 2,364 assigned children were reunified and only 321 had a subsequent *referral* (14%) state-wide. 23% of 95 reunified ISS received children had a subsequent referral, which was marginally/significantly better than ISS not workable (14% of 753; p = .09) or SAU (13% of 1,516; p = .03). In the ITT analysis, the ISS advantage (compared to SAU) was not significant (p=.17).

FIGURE 69The goal of ISS is keeping families safely together in the home. The benefits of ISS will vary depending upon the interventions used to address the families' specific needs. Depending on the needs of the family and the targeted interventions used, expected results may include a reduction in parental depression; reduction in parental substance abuse; and/or an increase in the caregiver's communication and decision-making skills, thus, reducing the possibility of future domestic violence incidents and relationship conflict. The long-term benefit of ISS is the sustainable behavior changes in caregivers to eliminate or reduce the reoccurrence of abuse or neglect. The ultimate goal is the reduction of children being removed from the home and entering care as well as reduction of the total number of children in care. Figure 1 portrays the logic model for ISS.

Figure 1, the ISS received group had a significantly worse likelihood of subsequent referrals (32% of 1,077), compared to ISS not workable (29% of 1,517; p < .001) and SAU (26% of 1,027; p < .001). Regarding ITT, the ISS-assigned clients were more likely to receive an additional referral (30% vs 26%; p < .001). This gap does not necessarily reflect a failure of the ISS program. It could reflect that more ISS children remain in the home (as evident in Q6a and Q6b) and, therefore, have more exposure to (greater chance of) subsequent referrals (and to removals for Q8a).

Q7b: Does ISS reduce the likelihood of additional referrals after reunification?

In the 48 months following the first ISS assignment, 2,364 assigned children were reunified and only 321 had a subsequent *referral* (14%) state-wide. 23% of 95 reunified ISS received children had a subsequent referral, which was marginally/significantly better than ISS not workable (14% of 753; p = .09) or SAU (13% of 1,516; p = .03). In the ITT analysis, the ISS advantage (compared to SAU) was not significant (p=.17).
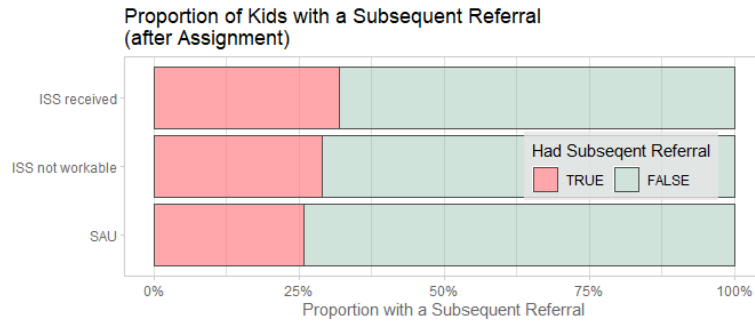
Q7c: Does ISS increase the duration before additional referrals?

Although the ISS received group experienced more subsequent referrals following services, it is still possible that the timing of these referrals takes longer, on average, than do those among their ISS not workable and SAU peers. To examine this hypothesis, we conducted another survival analyses from all the subsequent referrals noted in section Q7a.

FIGURE 70 below summarizes the findings from this investigation. Again, we see the ISS received group performing worse on this outcome than the ISS not workable ($p < .001$) and SAU ($p < .001$) groups, suggesting this subgroup experienced future referrals faster than the other two groups. However, as mentioned in Q7a, these groups are likely not equivalent on the dimension of risk. A disproportionate number of ISS not workable and SAU children were removed initially and were still in placements as a result of their initial referral. These individuals are not included in the analyses used to create the plots below, since they had no opportunity to experience a subsequent referral.

Regarding ITT, the SAU advantage (compared to ISS) was not significant ($p = .17$). When you combine the ISS not workable and ISS received groups, the survival curve for this group and SAU is virtually identical as shown in the right-hand panel of the figure below.
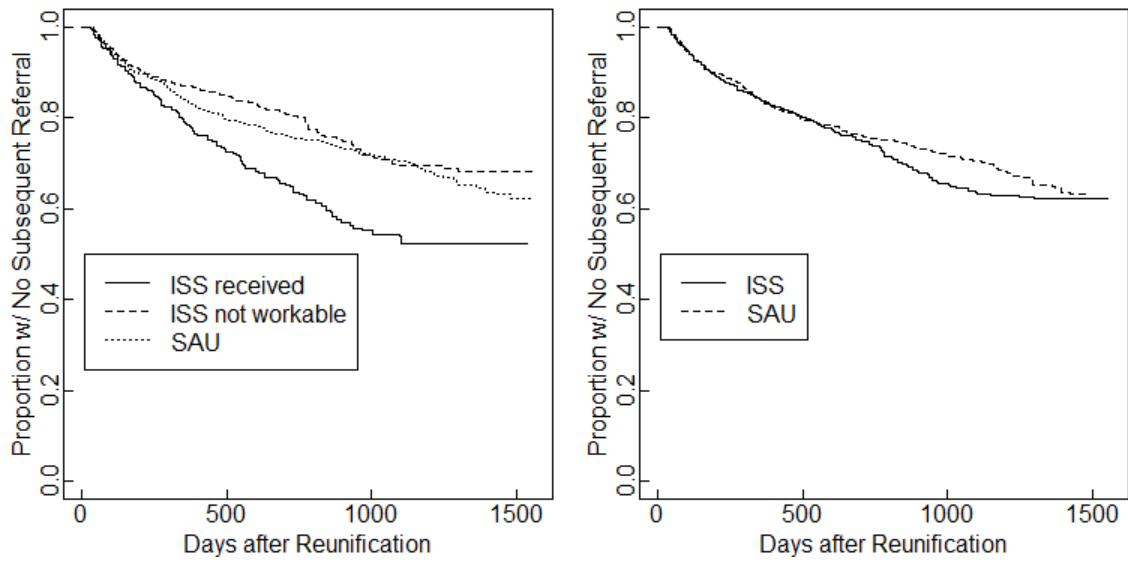
Question 8: Does ISS reduce the likelihood of an additional removal?

> The key takeaways for this question include:
>
> - Although we observed that the likelihood of additional removals was higher for ISS, this was not a significant result.
> - Due to low sample size, analysis could not determine if ISS reduces the likelihood of additional removals after reunification.
> - ISS did not increase the duration between initial and additional future removals.

Future *removals* are examined from three related perspectives.

a. Does ISS reduce the likelihood of additional removals: The likelihood is better, but not significantly.
b. Does ISS reduce the likelihood of additional removals *after* reunification: sample size too small.
c. Does ISS increase the duration before additional removals: No.

Methods
The data and methods for question 8 are described under the Question 6 Method section above.

Q8a: Does ISS reduce the likelihood of additional removals?

The count of subsequent removals (*i.e.*, removals unrelated to the initial referral that triggers a group assignment) were compared across the *As Treated* groups. Q8a includes two types of clients: those who were removed and reunited on the first referral, and those who were never removed on their first referral.

As seen in FIGURE 71, 8.4% (of 1,077) of the ISS received group had a removal, which is a significant increase compared to ISS not workable (5.9% of 1,517; $p = .04$), and to SAU (5.6% of 3,966; $p = .03$).

Regarding ITT, ISS was only slightly better, but not significant (6.9% vs 5.6%; $p = .34$).

Reducing (first time and subsequent) removals is a primary goal of ISS. While we see trends towards reductions in removals, we do not have significance.

FIGURE 71 - PROPORTION OF KIDS WITH SUBSEQUENT REMOVALS AFTER ASSIGNMENT
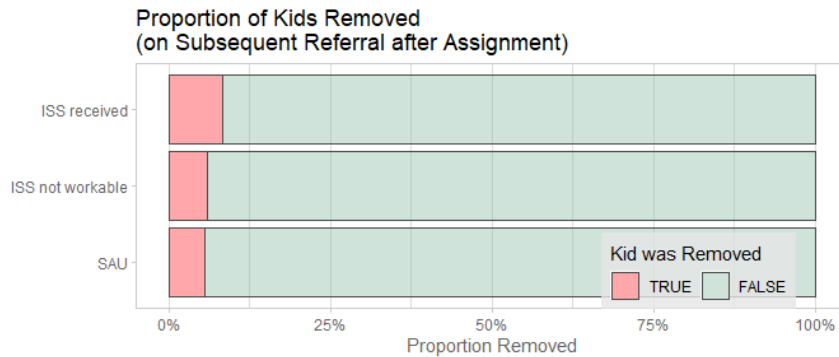
TABLE 204 - LOGISTIC REGRESSION OF SUBSEQUENT REMOVALS: COMPARING 'ISS RECEIVED' AGAINST 'ISS NOT WORKABLE' AND 'SAU'. THE FIRST ROW ESTIMATES THE INTERCEPT OF THE "ISS RECEIVED" GROUP, WHILE THE NEXT TWO ROWS ESTIMATE THE GROUP'S OFFSET FROM "ISS RECEIVED". THE BOTTOM ROW CONTROLS FOR THE PRE-TREATMENT RISK.

	Estimate	Std. Error	z value	p
Intercept	-2.58	0.18	-14.6	.0001
Group Actual: ISS Not Workable	-0.36	0.18	-2.0	.0427
Group Actual: SAU	-0.32	0.15	-2.1	.0328
PREMISS Risk Score	-0.11	0.08	-1.5	.1351

TABLE 205 - ITT LOGISTIC REGRESSION OF SUBSEQUENT REMOVAL: COMPARING 'ISS' (IRRESPECTIVE OF SERVICES RECEIVED) AND 'SAU'. THE FIRST ROW ESTIMATES THE INTERCEPT OF THE "ISS ASSIGNED" GROUP, WHILE THE MIDDLE ROW ESTIMATES THE SAU'S OFFSET FROM "ISS ASSIGNED". THE BOTTOM ROW CONTROLS FOR THE PRE-TREATMENT RISK.

	Estimate	Std. Error	z value	p
Intercept	-2.78	0.15	-18.8	<.001
Group Assignment SAU	-0.11	0.11	-1.0	.341
PREMISS Risk Score	-0.11	0.08	-1.4	.150

Q8b: Does ISS reduce the likelihood of additional removals after reunification?

In the 48 months following the first ISS assignment, only a few children were ever removed again. 1% of 95 ISS received kids had two removals. This is a nonsignificant improvement compared to 3% of 752 ISS

Not Workable kids ($p = .48$) and 4% of 1,516 SAU kids ($p = .30$). This absence of recurrence is mostly a product of the long period of time required to regain custody (reunify). See FIGURE 72.

FIGURE 72 - PROPORTION OF KIDS WITH SUBSEQUENT REMOVALS AFTER ASSIGNMENT AFTER REUNIFICATION
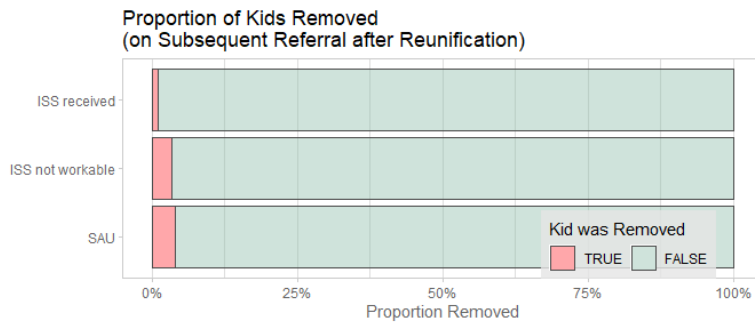
Q8c: Does ISS increase the duration before additional removals?

Although the ISS received group experienced slightly fewer subsequent removals following initial services, it is still possible that the timing of these removals takes longer, on average, than do those among their ISS not workable and SAU peers. To examine this hypothesis, we conducted another survival analyses for all the subsequent removals noted in section Q8a.

F<small>IGURE</small> **73** summarizes the findings from this investigation. Here we see the ISS received group performed marginally better than ISS not workable (p = .07) and comparably to SAU (p = .16) groups. As mentioned above, these groups are likely not equivalent on the dimension of risk. A disproportionate number of ISS not workable and SAU children were removed initially and were still in placements as a result of their initial referral. These individuals are not included in the analyses used to create the plots below, since they had no opportunity to experience a subsequent removal.

Regarding ITT, there was no significant difference between the two groups (p = .27). When the ISS not workable and ISS received groups are combined, the survival curve for this group is very similar to the overall SAU curve as shown in the right-hand panel of the figure below.
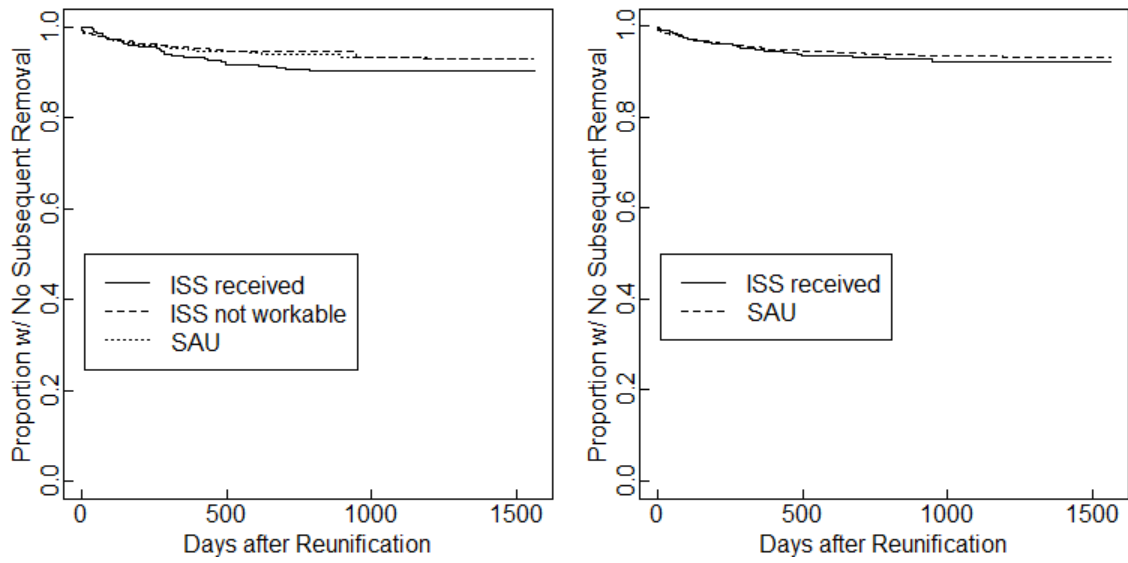
Region is determined by Staffed worker. Region 6 represents the state office.

TABLE 206 —BY REGION OF ASSIGNED WORKER: DURATION UNTIL REMOVAL FROM FIRST REFERRAL. ANALOG OF **TABLE 196** &

TABLE 199. WITHIN EACH COLUMN, THE FIRST VALUE IS THE DAYS UNTIL THE REMOVAL, THE SECOND VALUE IS THE PERCENTAGE THAT WAS REMOVED, AND THE THIRD VALUE IS THE COUNT OF KIDS IN THE CONDITION.

```
Region       ISS received    ISS not workable           SAU
1     470d; 18% of  208   248d; 60% of  480   187d; 59% of 1,591
2     417d; 13% of  179   180d; 64% of  258   178d; 53% of  918
3     603d; 20% of  451   211d; 76% of  769   280d; 61% of 1,385
4     256d; 13% of   78   187d; 45% of  130    51d; 55% of  725
5     350d; 16% of  201   235d; 46% of  393   200d; 55% of 1,046
6     746d; 22% of   76   140d; 80% of   49   394d; 56% of   82
```

TABLE 207—BY REGION OF ASSIGNED WORKER: DURATION UNTIL REMOVAL FROM ALL REFERRALS. ANALOG OF **TABLE 197** & **TABLE 200**. WITHIN EACH COLUMN, THE FIRST VALUE IS THE DAYS UNTIL THE REMOVAL, THE SECOND VALUE IS THE PERCENTAGE THAT WAS REMOVED, AND THE THIRD VALUE IS THE COUNT OF KIDS IN THE CONDITION.

```
Region       ISS received    ISS not workable           SAU
1     504d; 25% of  208   385d; 62% of  480   293d; 62% of 1,591
2     436d; 20% of  179   277d; 67% of  258   236d; 55% of  918
3     678d; 29% of  451   475d; 79% of  769   482d; 66% of 1,385
4     260d; 17% of   78   206d; 47% of  130   181d; 57% of  725
5     377d; 22% of  201   290d; 48% of  393   276d; 56% of 1,046
6     821d; 25% of   76   513d; 86% of   49   659d; 60% of   82
```

TABLE 208 —BY REGION OF ASSIGNED WORKER: REMOVAL DURATION FROM FIRST REFERRAL. ANALOG OF **TABLE 201** & **TABLE 202** WITHIN EACH COLUMN, THE FIRST VALUE IS THE DAYS REMOVED, THE SECOND VALUE IS THE PERCENTAGE THAT WAS REUNIFIED, AND THE THIRD VALUE IS THE COUNT OF ELIGIBLE KIDS IN THE CONDITION.

```
Region       ISS received    ISS not workable           SAU
1     429d; 19% of   37   406d; 37% of  286   348d; 37% of  945
2     490d; 22% of   23   412d; 24% of  166   290d; 18% of  485
3     603d; 30% of   92   532d; 34% of  586   461d; 31% of  850
4     372d;  0% of   10   322d; 24% of   58   205d; 17% of  399
5     257d; 41% of   32   345d; 40% of  180   296d; 33% of  572
6     674d; 35% of   17   582d; 31% of   39   613d; 28% of   46
```

```
Region        ISS received      ISS not workable             SAU
1      367d; 32% of  37   425d; 37% of  286   354d; 39% of  945
2      400d; 22% of  23   399d; 27% of  166   288d; 18% of  485
3      614d; 54% of  92   533d; 35% of  586   465d; 33% of  850
4      317d;  0% of  10   317d; 24% of   58   205d; 18% of  399
5      227d; 53% of  32   346d; 42% of  180   299d; 34% of  572
6      640d; 47% of  17   585d; 36% of   39   618d; 28% of   46
```

Cost Study

Title IV-E Cost Analysis: A Three Year Estimate

Our cost analysis estimates the costs of and any associated savings attributable to ISS from the perspective of DHS—the child welfare agency. To do that, we estimate the program costs in relationship to the cost savings from observed outcomes in the randomized control trial of ISS over the course of a 3 year period for families that we are able to track for that period of time. We follow the approach of estimating cost savings outlined in Maher et al. (2012)[33] and Annie E. Casey's brief on how to estimate cost savings stemming from reductions in foster care.[34] We take a similar approach by calculating the days in foster care and other service tracks and comparing the costs across evaluation groups – ISS Received, ISS Not Workable, Services As Usual. Like the Maher et al (2012) approach, we include personnel costs, as well as governance and administration costs to the agency, in addition to the foster care subsidy rate for foster care providers. Other than the governance and administrative costs, this information was provided to the researchers directly from OK DHS. In the final round of cost analyses, we will include any cost savings stemming from reductions in personnel costs from CPS investigations based on observed differences in subsequent reports of maltreatments.

Program Costs

Based on differences in available data specificity, we take two approaches to calculate the daily costs of different service tracks. These tracks include ISS, CHBS, FCS, and foster care. For ISS and CHBS, we calculate the daily rate by using the amount in the DHS contracts to providers for the service divided by the number of children served, weighted by the regional amounts and number of children served in each region. For CHBS, the ISS step-down program, we do the same and also weight the costs by number of cases served per condition in each region (since regional contract costs differed). We make the assumption that these costs are all-inclusive personnel and operating costs, including overhead, which we also refer to as governance and administration (G&A).

To calculate the daily cost of FCS and foster care, we take a different approach. For each, we calculate the personnel costs to provide these services based on average salaries by position (caseworkers, lead caseworkers, and supervisors), caseload size, and supervisory load (i.e., the number of caseworkers

---

[33] Maher, E. J., Corwin, T. W., Hodnett, R., & Faulk, K. (2012). A cost-savings analysis of a statewide parenting education program in child welfare. *Research on Social Work Practice*, *22*(6), 615-625.

[34] The Annie E. Casey Foundation (2019). *What's the Return on Investment? Using Placement Day Analysis to Measure Child Welfare Costs.*

supervised by one supervisor). These salaries, caseloads, and supervisory load were provided to the evaluation team by DHS. Based on this information, we convert these to an average daily personnel cost per child. To account for non-personnel G&A costs (overhead), which previous studies in child welfare have reported to be around 24-26% of the total costs (Corso & Filene, 2009; Foster et al., 2007), we have increased the personnel costs by 25%. Finally, we add the foster care stipend provided to foster care providers to the daily personal costs to estimate the total daily cost of foster care. Table 210 presents the daily rates and the components of each.

TABLE 210 DAILY COST FOR SERVICE TRACKS

Service Track	Daily Cost	What's Included
ISS	$130	Contract Amount/Number of Children Served
CHBS	$34.76 - $36.39	Contract Amount/Number of Children Served
Family Centered Services	$35.84	Caseworkers and supervisors salaries based on caseload and supervisors' caseloads + G&A
Permanency (Foster Care)	$37.94	Caseworkers and supervisors salaries based on caseload and supervisors' caseloads + G&A + foster parent stipend

One limitation in these estimations of daily costs for each service track is the difference in the data sources used to measure each. They might not be directly comparable, though we are treating them as such in this analysis. For example, the FCS and foster care daily rates are based on personnel expenditures, which is more of an ingredient approach, and CHBS and ISS costs are based on price paid for these services, which may or may not reflect true costs.

Program Cost Differentials

Based on observed outcomes, we calculate costs for an average family over a three-year period based on observed differences in days for each service track—ISS, CHBS, FCS, and foster care.

Estimates of the cost differentials between the three evaluation groups (ISS received, ISS not workable, and SAU) are based on mean days in each service track across all families in that track with the associated daily cost applied to that number of days. If a family spent no days in a particular service track, their zero value is still used in the calculation of the mean number of days. Thus, the mean days reflects the average number of days across all families in the evaluation group. We compute these costs separately for families across one year, two years, and three years so that we can see how the cost differentials change over time. The means for each year are based on those families whose randomization date allows for that follow-up period, which ends on October 2nd, 2019.

Table 211 presents the daily costs and the mean days by year and evaluation group for each service track. Obviously, mean ISS days are only relevant to the ISS received group.

		Daily Rate	1-Year Mean Days	2-Year Mean Days	3-Year Mean Days
ISS Not Workable	CHBS	34.76	31.8	36.9	13.3
	Foster Care	37.94	279.7	228.1	81.5
	FCS	35.84	31.8	36.9	13.3
Service as Usual	CHBS	36.39	38.2	36.2	10.7
	Foster Care	37.94	219.7	174.7	82.9
	FCS	35.84	38.2	36.2	10.7
ISS Received	ISS	130.54	37.6	--	--
	CHBS	34.76	173.8	20.8	17.1
	Foster Care	37.94	62	84.8	56.3
	FCS	35.84	211.4	20.8	17.1

It should be noted that we believe this is some duplication in days between the service tracks for the ISS groups, leading to an overestimate of costs for ISS. Given the different data sources for service end dates, we were not able to correct for this duplication.

Calculating Expenditures

Daily prices were multiplied by the mean days to estimate costs per year of service for each group. In our analysis, we use the mean number of days a group has experienced in each tract of service for each year as described above. These mean number of days are multiplied by the daily costs for each service to estimate the average annual cost of each tract for each group of cases. Summing across service tracts, we get the average total family cost for each period based on evaluation group assignment. These costs are estimated at the referral or family level and not the child level, so this may underestimate the costs for families with multiple children.

Results

Table 212 presents the average family costs per year across all service tracks for each evaluation group. For families receiving ISS, the first year costs are substantially higher than ISS not workable and SAU, as would be expected, given the high up-front costs for ISS. However, after three years, tracking all service days, the total costs between ISS Received and ISS Not Workable become almost indistinguishable and only slightly more than service as usual. In other words, in a short amount of time—three years, the more expensive and intensive intervention model gets closer to being no more expensive than not using it, which reflects the better outcomes that it achieves, offsetting the additional service delivery costs.

	ISS Received	ISS Not Workable	SAU
Year 1	$20,879	$12,858	$11,095
Year 2	$4,686	$11,260	$9244
Year 3	$3,344	$4,012	$3,919
Three-Year Average Costs	**$28,909**	**$28,150**	**$24,258**

The current estimates do not include any cost differentials based on investigations stemming from substantiated subsequent reports of maltreatment during the time periods for each service tract. Once these are included as a source of cost differentials, the relationship may change. We also plan to include some sensitivity analyses related to key assumptions in a later version of these cost analyses. Finally, these estimates do not include costs to other systems, or intangible costs, such as child and family well-being stemming from reducing the number of days, attributable to ISS, that youth spend in foster care.

Summary and Lessons Learned

Highlights of Findings to Date

Process Study

Through the reporting period covered herein, a total of 4,851 families have been identified as eligible for services. A total of 614 of these PREMISS eligible families were later discovered to meet exclusion criteria, thus, only 4,237 families were randomized to either the Intensive Safety Services (ISS) or Services As Usual (SAU) study conditions. Of these referrals, 1,522 were assigned to ISS- 526 belonging to the ISS received group and another 996 belonging to the ISS not workable group. The SAU condition to date has been assigned a total of 2715 referrals.

Through qualitative interviews and surveys of Demonstration project staff and stakeholders, the evaluators have observed the following barriers:

- Declining eligibility referrals in Region 3 since 2017
- Consistent difficulties getting eligibility data prior to a child safety meeting (CSM)- where often irreversible placement decisions occur. Delays of eligibility data are bad in all regions, but Regions 1, 2, and 4 are particularly troublesome with eligibility data arriving, on average, more than 9 days after a CSM.
- Poor documentation exchanges with the evaluation team. For example, there are more missing staffing forms than actual returned forms to date.
- Data suggest protocol violations are either common or are overrepresented in existing documentation.
- Communication and collaboration issues have been identified across DHS and contracted service agencies (e.g., ISS contractors communicating with CPS staff in a timely manner).
- Some worker concerns on the length of ISS- many wish it was longer than 6 weeks.

- Some concerns over the rushed assessments that currently inform eligibility. Some feel the PREMISS model occasionally selects inappropriate families.
- Delayed Family Centered Services (FCS) involvement (FCS workers are responsible for managing cases during ISS and CHBS service periods) following CPS investigation.

Process data from staff and stakeholders have also noted a number of project strengths to date. These include:

- Once ISS referral is received, case processing is closely following the protocol.
- Based on available data, there has been a high rate of accepted services among the ISS received group.
- When ordered properly, with eligibility data arriving ahead of the CSM, the Waiver and CSM initiatives have been working well together. Workers appreciate the additional option of keeping children in the home with more intensive services and find the CSM to be an essential component of strategic planning for ISS success.
- FCS and CPS have been extremely pleased with the responsiveness and competence of contracted ISS teams.
- FCS, ISS, and Core Waiver teams have been agile in responding to various unanticipated scenarios. That aspect of collaboration has been praised by all parties involved.
- The max ISS caseload sizes have been generally welcomed and applauded as suitable.
- Working alliance between ISS Received clients and providers appears to be higher than among workers and the SAU and ISS Not Workable clients. ISS received clients also report higher average satisfaction with services compared to these other two groups.

In the Outcome study portion of the evaluation, we are investigating the impact of ISS on a number of targeted constructs. Key outcome findings to date include:

- ISS recipients experienced larger reductions in safety threats (rated and monitored by DHS) at 6 months and greater growth in protective capacities (rated and monitored by DHS) at Stepdown (6 weeks after referral) and 6 months compared to SAU and ISS not workable groups. Only the protective capacities differences held up in the intent-to-treat (ITT) analyses comparing all ISS assigned clients to all SAU families.
- The ISS received group has evidenced significant improvements in parental depression and distress symptoms, rates of mental health and substance abuse treatment receipt, and rates of conflict among partners in the home.
- ISS received clients were also shown to significantly reduce concerning parenting behaviors during the CHBS stepdown service period.
- Our main measure of child well-being (involving items from the Pediatric Symptom Checklist) has not shown major differences across our study conditions. All children studied to date seem to improve gradually over time at comparable rates. There might be some evidence for greater ISS received benefits within a small window of roughly 45 to 90 days following the initial referral.
- In terms of the primary target, keeping kids out of care, ISS appears to be succeeding.
- Once a child is removed, there does not appear to be an ISS advantage. When examining future CPS involvement (both referrals and removals), again there is no apparent ISS

advantage, and in fact, ISS served cases may be more likely to interface with the system again. Reasons for this recidivism almost certainly include the fact that ISS cases have more future opportunities for CPS involvement, since fewer of these children are removed from the home initially. It is also worth mentioning that the recidivism rates are not large compared to the initial differences that favored the ISS received group. Also, the intent-to-treat differences in recidivism were near zero.

Finally in our cost study, estimated of costs differentials for the three service groups (ISS received, ISS not workable, and SAU) are provided in the full report. These cost differentials were roughly equivalent by the end of a third year of service.

In summary, the evaluation has proceeded in close accordance with the original plan. While there are a number of concerns being noted in our process study, the positive outcomes to date, including a significant reduction of kids in care when assigned to ISS, have brought enthusiasm and excitement to all who are involved. Further, although the costs differentials estimated to date do not favor the ISS condition, the differentials per family are relatively small and might reasonably be offset by the host of positive outcomes we are observing among the ISS served population.

Evaluation Lessons Learned and Recommendations

Evaluators have made some minor adaptations along the way which now serve as lessons learned for similar future evaluation designs. On the use of an automated eligibility algorithm, evaluators have scrapped initial plans to fully manage the operation of the algorithm in-house. Infrastructure requirements and staffing costs proved to be too great for this task, but also internal expertise has been shown to be an essential element of monitoring proper functioning and worker understanding of the processes involved. Evaluators also have had to adapt their use of randomization to meet the needs of contracted ISS agencies. An adaptive randomization algorithm is in place and minor changes to the probability were required during year 1 to ensure that the ISS agency maintained a financially viable service capacity during most months of the year. This meant change our initial assignment probabilities to heavily favor an ISS assignment when several ISS spots were open simultaneously. It took a bit of back-and-forth calibration to reach a happy medium. One striking lesson learned is that the data utilized is only as good as the source will allow. Evaluators are constantly reminded that much of the data being used for testing was not provided with scientific rigor in mind. These data are used to facilitate practice and suffer from a number of weakness, not the least of which is high levels of missingness. There is a delicate balance to manage on this front, as recommendations to either more strictly enforce data collection protocols or reallocate all data collection responsibilities to the evaluation team come with their own unique disadvantages. The former resolution threatens to disenfranchise and distract the service staff whose number one priority ought to be program implementation, while the latter requires a much higher financial investment in evaluation activities.

Future Plans

OK DHS has decided to continue with ISS beyond this Waiver demonstration period, in part because of the strong commitment to the goals of the program and the evaluation findings. The evaluation team is continuing to evaluate its effectiveness in a new study, while summarizing these five-year results for publication and review by the Family First Prevention Services Act clearinghouse.

Appendix A. Process Study Qualitative Guides

Administrators, Directors, and Supervisors

1. Please tell me your name and your current position at (name of institution).

2. Please describe the process of implementing ISS services at (name of agency).
    Probe: Positive experiences?

    Negative experiences?

3. What have been some of the barriers of implementing the ISS services?
    Probes: Getting referrals?
    Child welfare system?
    Caregiver engagement?
    Family willingness to participate?
    Inadequate training or skill?
    Funding issues?
    Judicial system?

    a. How is or has your agency dealt with these issues?

4. What are some administrative or political issues that have impacted the program?

5. How do various programs <u>within</u> and <u>between</u> agencies collaborate to make this program a success.

    a. If they are not collaborating:
        i. What could be done to make collaboration a success?
    b. What additional collaborations are needed?

6. Which decision-makers, or stakeholders, influence ISS services for youth and families?
    Probe: What does this person/agency/organization do and where are they located?
    Why are they important stakeholders?

7. What policies and practices (e.g., legislative, judicial system, interagency, program-specific) are in place to support ISS services?
    a. How do they support families?
    b. How do they prevent families from getting support?

8. How well were/are you able to support ISS services in connecting families to resources?
    a. What areas of need are readily addressed with outside services?
    b. What areas of need are there none or limited resources for families in this community? (e.g., housing, transportation, substance abuse services, etc.)
    c. What are some of the barriers to connecting families to outside services?

9. We know that FCS/ISS providers go through a lot of training to help ISS families.
   a. How accessible is training?
   b. How well do you think the training prepares workers to assist ISS families?
   c. What other supports do they receive?
      i. supervision
   d. What additional training and support do think they need?
10. What, if anything would make your job easier in supporting the FCS/ISS services in helping families?
   a. Changes in organization?
   b. Support?

11. How would you improve ISS services for families?
   a. Improve the effectiveness of the program?
   b. Improve the efficiency of the program?

12. What organizational changes have happened to make the program successful?
   a. What about cooperation with the court system?
   b. What needs to happen?

13. Please discuss any concerns you have had with the implementation (design, management, execution) of ISS services for families?
   a. What are some strengths to implementation of these services?
   b. What continued concerns do you have?

14. Tell me about any safety concerns that you have in helping families within the Family Centered and Intensive Safety Services framework?

15. What do you think we should have asked you that was not included in this interview?

Child Protective Services Workers
1. Please tell me your name and your current position at (name of institution).

2. Walk me through the process of identifying families for ISS services?
   a. Explain how you use the PREMISS model? (computer model to determine ISS eligibility)
   b. When do you obtain information that a family is ISS eligible?
   c. What are the barriers to making this protocol successful?
   d. What changes to the protocol are needed?
   e. When you determine a family is eligible for ISS services, what happens next?
      Probe: Child Safety Meeting, Case Transfer
         1. About how long does it take to conduct the CSM?
            a. What are the barriers to conducting the CSM?
         2. About how long does it take to complete the case transfer?
            a. What are the barriers to completing the case transfer?

3. How has your decision making practices changed since implementation of the ISS services? (on paper and in reality)

4. Now I am going to ask more questions about the child safety meeting for ISS cases.
   a. Please describe if information is shared with FCS prior to the meetings?
      i. When is this information shared?
      ii. Who shares information?
   b. Do you know when attending a child safety meeting if the family is ISS eligible?
      i. No or Sometimes: What are the barriers to having this information?
   c. How do you present the program to families? (give me your sales pitch)
      i. What is the reaction of the families?
      (If they don't present to families find out who does)

5. Please describe your collaboration with FCS workers?
   a. Do you feel that your collaboration is successful?
      i. Why or Why not?
   b. What could be done to make the collaboration more successful?

6. How do you feel about ISS services?
   a. Positive
   b. Negative
   c. Tell me about any safety concerns you have in serving families through ISS services?

7. What, if anything would have make your job easier in working with the ISS process? (i.e., paperwork, support, etc.)

8. What are some administrative or political issues that have impacted the program?

9. What organizational changes have happened to make the ISS process successful?
   a. What practice (protocol) changes have happened?
   b. What needs to happen to make the program a success?
      i. Leadership?

      ii.   Supervision?

     iii.   Support?

10. Please discuss any concerns you have with the implementation (design, management, execution) of ISS services for families?
    a. What are some strengths to implementation of these services?

11. What do you think we should have asked you that was not included in this interview?

Family Centered Services Workers
1. Please tell me your name and your current position at (name of institution).

2. When it is determined a family is eligible for ISS services, what happens next?
    Probe: Child Safety Meeting, Case Transfer
    i. About how long does it take to conduct the CSM?
        1. What are the barriers to conducting the CSM?
    ii. About how long does it take to complete the case transfer?
        1. What are the barriers to completing the case transfer?

3. Tell me about how working with ISS families is different than traditional FCS families.
    i. Job responsibilities
    ii. Decision making
        1. How has decision making practices changed since implementation of the ISS services? (on paper and reality)

4. Now I am going to ask more questions about the child safety meeting for ISS cases.
    a. Please describe if information is shared with FCS prior to the meetings?
        i. When is this information shared?
        ii. Who shares information?
    b. Do you know when attending a child safety meeting if the family is ISS eligible?
        i. No or Sometimes: What are the barriers to having this information?
    c. How do you present the program to families? (give me your sales pitch)
        i. What is the reaction of the families?
        (If they don't present to families find out who does)

5. How do you feel about ISS services?
    a. Positive
    b. Negative
6. We know you go through a lot of training to help ISS families.
    a. How accessible is training?
    b. How well did the training prepare you to assist ISS families?
    c. What other supports do you receive?
        i. supervision
    d. What additional training and support do you need?

7. Now we are going to ask about engagement of families in the ISS program.
    a. How do you keep families in the program?
    b. What are characteristics of families that are able to work through the treatment plan (FIP)?
    c. What are characteristics of families that are unable to work through the treatment plan (FIP)?
    d. How do these families compare to traditional FCS families?

8. What are some things that are going well with providing services to families when they have an ISS worker?
    a. What is not going well?

        b.   What are some of the barriers to helping families?

9.   What are some things that are going well with providing services to families when they are moved from an ISS worker to stepdown services (CHBS)?
        a.   What is not going well?
        b.   What are some barriers to helping families at this stage?
        c.   Tell me about the timeline of when CHBS enters.

10. Tell me about the collaborations <u>within</u> your agency that are going well.
        a.   Not going well?
               i.   What could be done to improve these collaborations?
        b.   Barriers to collaborating?
      (Probe: CPS worker, OCS liaison)

11. Tell me about the collaborations <u>between</u> the ISS agency that are going well.
        a.   Not going well?
               i.   What could be done to improve these collaborations?
        b.   Barriers to collaborating?
      (Probe: ISS Worker)

12. How well were/are you able to connect intensive safety service families to resources?
        a.   Where are some places that you commonly refer families to?
        b.   What areas of need are readily addressed with outside services?
        c.   What areas of need are there none or limited resources for families in this community?
        d.   What are some of the barriers to connecting families to outside services?

13. Tell me about any safety concerns that you have in helping families within the ISS framework?
        a.   Barriers families face in keeping their children safe?
               i.   How do you address those barriers?
        b.   What are some barriers you face in keeping families safe?
               i.   How do you address those barriers?
        c.   How do you help families reduce stress?

14. What, if anything would have make your job easier in helping families in ISS services?

15. What are some administrative or political issues that have impacted the program?

16. What organizational structural changes have happened to make the program successful (Show Flow Chart)?
        a.   What practice (protocol) changes have happened?
        b.   What needs to happen to make the program a success?
               i.   Leadership?
             ii.   Supervision?
           iii.   Support?

17. How would you improve the ISS services for families?

18. Please discuss any concerns you have with the implementation (design, management, execution) of ISS services for families?
    a. What are some strengths to implementation of these services?

19. What do you think we should have asked you that was not included in this interview?

Family Centered Services Supervisors
1. Please tell me your name and your current position at (name of institution).

2. What have been some of the barriers of implementing ISS services?
        Probes: Getting referrals?
                        Child welfare system?
                        Caregiver engagement?
                        Family willingness to participate?
                        Inadequate training or skill?
                        Funding issues?
                        Judicial system?

3. When it is determined a family is eligible for ISS services, what happens next?
        Probe: Child Safety Meeting, Case Transfer
            i. About how long does it take to conduct the CSM?
                    1. What are the barriers to conducting the CSM?
            ii. About how long does it take to complete the case transfer?
                    1. What are the barriers to completing the case transfer?

4. Tell me about how working with ISS families is different than traditional FCS families.
            i. Job responsibilities
            ii. Decision making
                    1. How has decision making practices changed since implementation of
                        the ISS services? (on paper and reality)

5. Now I am going to ask more questions about the child safety meeting for ISS cases.
        a. Please describe if information is shared with FCS prior to the meetings?
            i. When is this information shared?
            ii. Who shares information?
        b. Do you know when attending a child safety meeting if the family is ISS eligible?
            i. No or Sometimes: What are the barriers to having this information?
        c. How do you present the program to families? (give me your sales pitch)
            i. What is the reaction of the families?
                (If they don't present to families find out who does)

6. How do you feel about ISS services?
        a. Positive
        b. Negative

7. We know you and your workers go through a lot of training to supervise workers that help ISS
    families.
        a. How accessible is training?
        b. How well does the training prepare workers to assist ISS families?
        c. What other supports do workers you receive?
            i. supervision
        d. What additional training and support do workers need?

8. Now we are going to ask about engagement of families in the ISS program.
    a. How do you keep families in the program?
    b. What are characteristics of families that are able to work through the treatment plan (FIP)?
    c. What are characteristics of families that are unable to work through the treatment plan (FIP)?
    d. How do these families compare to traditional FCS families?
    e. How do you present the program to families? (give me your sales pitch)
        i. What is the reaction of the families?
           (If they don't present to families find out who does)

9. What are some things that are going well with providing services to families when they have an ISS worker?
    a. What is not going well?
    b. What are some of the barriers to helping families?

10. What are some things that are going well with providing services to families when they are moved from an ISS worker to stepdown services (CHBS)?
    a. What is not going well?
    b. What are some barriers to helping families at this stage?
    c. Tell me about the timeline of when CHBS enters.

11. Tell me about the collaborations <u>within</u> your agency that are going well.
    a. Not going well?
        i. What could be done to improve these collaborations?
    b. Barriers to collaborating?
    (Probe: CPS worker, OCS liaison)

12. Tell me about the collaborations <u>between</u> the ISS agency that are going well.
    a. Not going well?
        i. What could be done to improve these collaborations?
    b. Barriers to collaborating?
    (Probe: ISS Worker)

13. How well were/are your workers able to connect intensive safety service families to resources?
    a. Where are some places that workers commonly refer families to?
    b. What areas of need are readily addressed with outside services?
    c. What areas of need are there none or limited resources for families in this community?
    d. What are some of the barriers to connecting families to outside services?

14. Tell me about any safety concerns that you have in helping families within the ISS framework?
    a. Barriers families face in keeping their children safe?
        i. How do you address those barriers?
    b. What are some barriers you face in keeping families safe?
        i. How do you address those barriers?
    c. How do you help families reduce stress?

15. What, if anything would have make your job easier in supervising workers to help families in ISS services?

16. What are some administrative or political issues that have impacted the program?

17. What organizational structural changes have happened to make the program successful (Show Flow Chart)?
    a. What practice (protocol) changes have happened?
    b. What needs to happen to make the program a success?
        i. Leadership?
        ii. Supervision?
        iii. Support?

18. How would you improve ISS services for families?

19. Please discuss any concerns you have with the implementation (design, management, execution) of ISS services for families?
    a. What are some strengths to implementation of these services?

20. What do you think we should have asked you that was not included in this interview?

Intensive Safety Services Therapists
1. Please tell me your name and your current position at (name of institution).

2. We know you go through a lot of training to help ISS families.
   a. How accessible is training?
   b. How well did the training prepare you to assist ISS families?
      i. What was most helpful in training
      ii. What was the least helpful in training
   c. What other supports do you receive?
      i. supervision
   d. What additional training and support do you need?

3. Now we are going to ask about engagement of families in the ISS program.
   a. How do you keep families in the program?
   b. What are characteristics of families that are able to work through the treatment plan (FIP)?
   c. What are characteristics of families that are unable to work through the treatment plan (FIP)?
   d. How is the ISS program being presented to families?
      i. What is the reaction of the families?
         (If they don't present to families find out who does)
      ii. When you present the program to families – how do you present it? (give me your sales pitch)

4. What are some things that are going well with providing services to families within the ISS framework?
   a. What is not going well?
   b. What are some of the barriers to helping families?

5. Tell me about the collaborations with DHS that are going well.
   a. Not going well?
      i. What could be done to improve these collaborations?
   b. Barriers to collaborating?
   (Probe: FCS worker, CPS Worker, OCS liaison)

6. How well were/are you able to connect ISS families to services and resources?
   a. Where are some places that you commonly refer families to?
   b. What areas of need are readily addressed with outside services?
   c. What areas of need are there none or limited resources for families in this community?
   d. What are some of the barriers to connecting families to outside services?
   e. Tell me about any assistance the FCS worker provides in finding resources and services for families.

7. Tell me about any safety concerns that you have in helping families within the ISS framework?
   a. Barriers families face in keeping their children safe?
   b. How do you address those barriers?

  c. How do you help families reduce stress?

8. What, if anything would have make your job easier in helping families in ISS services?

9. What are some administrative or political issues that have impacted the program?

10. What organizational changes have happened to make the ISS program successful?
  a. What practice (protocol) changes have happened?
  b. What needs to happen to make the program a success?
    i. Leadership?
    ii. Supervision?
    iii. Support?

11. How would you improve ISS services for families?

12. Please discuss any concerns you have with the implementation (design, management, execution) of ISS services for families?
  a. What are some strengths to implementation of these services?

13. What do you think we should have asked you that was not included in this interview?

ISS Supervisors
1. Please tell me your name and your current position at (name of institution).

2. What have been some of the barriers of implementing ISS services?
   Probes: Getting referrals?
   Child welfare system?
   Caregiver engagement?
   Family willingness to participate?
   Inadequate training or skill?
   Access to other services?
   Funding issues?
   Judicial system?
   Anything else?

   a. How is or has your agency dealt with these issues?

3. We know you and your workers go through a lot of training to help ISS families.
   a. How accessible is training?
   b. How well did the training prepare you to assist workers in helping ISS families?
      i. What was most helpful in training
      ii. What was the least helpful in training
   c. What other supports do you provide to workers?
      i. supervision
   d. What additional training and support do workers need?

4. Now we are going to ask about engagement of families in the ISS program.
   a. How do you keep families in the program?
   b. What are characteristics of families that are able to work through the treatment plan (FIP)?
   c. What are characteristics of families that are unable to work through the treatment plan (FIP)?
   d. How is the ISS program being presented to families?
      i. What is the reaction of the families?
         (If they don't present to families find out who does)
      ii. When you present the program to families – how do you present it? (give me your sales pitch)
   e. As a supervisor, is there anything else that you to do help?
      i. When do you step in to help with a case?

5. What are some things that are going well with providing services to families within the ISS framework?
   a. What is not going well?
   b. What are some of the barriers to helping families?

6. Tell me about the collaborations with DHS that are going well.
   a. Not going well?

    i. What could be done to improve these collaborations?
  b. Barriers to collaborating?
  (Probe: FCS worker, CPS worker, OCS liaison)

7. How well is ISS able to connect families to services and resources?
  a. Where are some places that you commonly refer families to?
  b. What areas of need are readily addressed with outside services?
  c. What areas of need are there none or limited resources for families in this community?
  d. What are some of the barriers to connecting families to outside services?
  e. Tell me about any assistance the FCS worker provides in finding resources and services for families.

8. Tell me about any safety concerns that you have in helping families within the ISS framework?
  a. Barriers families face in keeping their children safe?
  b. How do you address those barriers?
  c. How do you help families reduce stress?

9. What, if anything would have make your job easier in supervising workers that help families in ISS services?

10. What are some administrative or political issues that have impacted the program?

11. What organizational changes have happened to make the program successful?
  a. What practice (protocol) changes have happened
  b. What needs to happen?
    i. Leadership?
    ii. Supervision?
    iii. Support?

12. How would you improve ISS services for families?

13. Please discuss any concerns you have with the implementation (design, management, execution) of ISS services for families?
  a. What are some strengths to implementation of these services?

14. What do you think we should have asked you that was not included in this interview?

Appendix B. Process Study Quantitative Survey Items
Select items from the following scales are presented to all qualitative participants at the completion of each interview.

- Children's Services Survey[35][36] sections VII and VIII
- Evidence-Based Practice Attitude Scale[37] Items 1-8
- Innovation Implementation Questionnaire[38]
- Turnover Intentions[39]
- Caseworker Skills & Job Experiences[40][41] Scales 2-4, 6-7

[35] Glisson C. (2002). The organizational context of children's mental health services. Clinical Child and Family Psychology Review, *5,* 233–253.

[36] Glisson C., and James L.R. (2002). The cross-level effects of culture and climate in human service teams. *Journal of Organizational Behavior, 23*, 767–794.

[37] Aarons, G. A. (2005). Measuring Provider Attitudes Toward Evidence-Based Practice: Consideration of Organizational Context and Individual Differences. *Child and Adolescent Psychiatric Clinics of North America*, *14*(2), 255–viii.

[38] Hurlburt, M., & Knapp, P. The Innovation Implementation Questionnaire measures 24 domains often affecting implementation of a new innovation.

[39] Knudsen H.K., Johnson J.A., and Roman P.M. (2003). Retaining counseling staff at substance abuse treatment centers: Effects of management practices. *Journal of Substance Abuse Treatment, 24,* 129–135.

[40] Detlaff, A.J., Graham, J.C., Holzman, J., Baumann, D. J., and Fluke, J.D. (2015). Factors that influence the removal decision in child protective services: development of an instrument to understand the decision-making process. *Child Abuse and Neglect, 49,* 24-34.

[41] Fluke, J., Parry, C., Shapiro, P., Hollinshead, D., Bollenbacher, V., Baumann, D. (2001). Davis-Brown, K., The Dynamics of Unsubstantiated Reports: A Multi-State Study –Final Report, Denver, Colorado: The American Humane Association.