April 2014

Program Evaluation of the Colorado Consortium on Differential Response

Final Report

Social Work Research Center

Research for Results

Colorado State University

College of Health and Human Sciences

School of Social Work

110 Education
Fort Collins, CO 80523
(970) 491-0885
http://www.ssw.chhs.colostate.edu/research/swrc/index.aspx

Program Evaluation of the Colorado Consortium on Differential Response

Final Report

Prepared by:

Marc Winokur
Rebecca Orsi
Helen Holmquist-Johnson

Social Work Research Center
School of Social Work
Colorado State University

Raquel Ellis
George Gabel
John Rogers
Summer Brenwald
Marneena Evans

Westat®

Prepared for:

National Quality Improvement Center
on Differential Response in Child Protective Services

The recommended citation for this report is:

Winokur, M., Ellis, R., Orsi, R., Rogers, J., Gabel, G., Brenwald, S., Holmquist-Johnson, H., & Evans, M. (2014). *Program evaluation of the Colorado Consortium on Differential Response: Final report.* Fort Collins, CO: Social Work Research Center, School of Social Work, Colorado State University.

Acknowledgements

The evaluation team would like to acknowledge the important contributions of all of the individuals and organizations that made this evaluation possible. We first thank the families who graciously shared their experiences and voices with us through the family exit survey. We also thank the Parent Partners from Jefferson County for helping us pilot test the family exit survey. The evaluation was further enriched by the involvement of the caseworkers, supervisors, and administrators from Arapahoe, Fremont, Garfield, Jefferson, and Larimer counties. We are so appreciative of their time and effort in participating in workgroups, following the random assignment protocol, and completing case-specific reports and general surveys. Thank you as well to the community stakeholders in these counties for sharing their perspectives in the focus groups during our two site visits. We are indebted to the practice leads in each county, Angela Lytle, Stacie Kwitek, Sharon Longhurst-Pritt, Mary Berg, and Jim Drendel for their commitment to the evaluation process, which permeated their agencies and communities. Furthermore, we are thankful for Graig Crawford, Senior Data Analyst from Jefferson County, who worked with the state to provide us with the best possible SACWIS data.

We thank the Colorado Department of Human Services (CDHS), under the leadership of Executive Director Reggie Bicha, for placing their trust in the evaluation team and supporting us every step of the way. Special thanks to Dan Makelky for his integral role in helping Colorado apply for and receive the grant to implement this project. The evaluation flourished because of the amazing partnership we had with Ida Drury, who served as the Project Director for CDHS.

We are grateful for the guidance provided by Catherine Nolan, Dori Sneddon, and Jean Blankenship from the Children's Bureau. We owe much of the success for the evaluation to the shared vision of Lisa Merkel-Holguin, Amy Hahn, and John Fluke from the Kempe Center for the Prevention and Treatment of Child Abuse and Neglect at the University of Colorado Denver School of Medicine.

The evaluation was truly a team effort and we are so thankful to Ying-Ying Yuan from the cross-site evaluator, Walter R. McDonald & Associates (WRMA), and Brett Brown, formerly of WRMA, for sharing their research expertise and evaluation wisdom with us. We also acknowledge Gary Siegel and Tony Loman from the Institute of Applied Research for sharing the research design and survey instrumentation that informed our evaluation. In addition, we greatly appreciate our evaluation partners, Tamara Fuller from the Children and Family Research Center at the University of Illinois at Urbana-Champaign, and Madeleine Kimmich, Julie Murphy, and Linda Newton-Curtis from the Human Services Research Institute for their

friendship and willingness to collaborate with us over the past four years. Lastly, we would like to acknowledge our colleagues at Colorado State University and Westat for all of the support they provided us during the evaluation.

The Colorado Consortium on Differential Response evaluation was supported by a contract with the Colorado Department of Human Services through a grant from the National Quality Improvement Center on Differential Response in Child Protective Services (QIC-DR), which was funded by the Children's Bureau, Administration for Children and Families (ACF), U.S. Department of Health and Human Services. The findings and conclusions expressed in this report do not necessarily represent those of CDHS, QIC-DR, or ACF.

Table of Contents

Program Evaluation of the Colorado Consortium on Differential Response

Executive Summary

Background

In 2007, a series of high-profile child fatalities in Colorado prompted numerous examinations of the state's child welfare system. Then-Governor Bill Ritter formed the Governor's Child Welfare Action Committee to conduct a full systemic evaluation. Comprised of various local, state, and national experts on child welfare, this committee met over the course of several years and developed 35 recommendations for improvement. One of the recommendations was Differential Response (DR), which is an innovative system reform that allows child protective services (CPS) to address screened-in allegations of child maltreatment in different ways. Although there is some variation in how CPS systems implement DR, the typical approach includes responding to high-risk cases with a traditional investigation while responding to low- and moderate-risk cases with an assessment of the family's needs and strengths. Safety and risk assessments are completed for all cases.

In October 2008, the American Humane Association (AHA), in partnership with Walter R. McDonald & Associates, Inc. and the Institute of Applied Research, received a five-year federal award from the Children's Bureau of the Administration on Children, Youth and Families, U.S. Department of Health and Human Services to develop the National Quality Improvement Center on Differential Response in Child Protective Services (QIC-DR). The major catalyst for the Colorado Consortium on Differential Response (CCDR) was the opportunity to apply for a grant from the Children's Bureau through the QIC-DR. The Colorado Department of Human Services (CDHS) solicited letters of interest from all 64 counties in Colorado and received five firm commitments. Administrators from the five committed counties saw DR as a good fit with or a logical extension of their current practice. Additionally, these five counties self-assessed that they were poised for a system reform of this magnitude.

In 2009, the QIC-DR selected CDHS–in conjunction with the Colorado counties of Arapahoe, Fremont, Garfield, Jefferson, and Larimer–as one of three Research & Development sites, along with Illinois and Ohio. In Colorado, child protective services (CPS) is state-supervised and county-administered with oversight by the Division of Child Welfare Services under the Office of Children, Youth, and Families in the CDHS. CDHS selected the Social Work Research

Center (SWRC) in the School of Social Work at Colorado State University (Fort Collins, CO) and Westat (Rockville, MD) as the evaluation team.

Methodology

The required evaluation approach addressed three main focus areas: (1) an outcome evaluation that rigorously examined the impact of DR on key child welfare outcomes including child safety and family engagement; (2) a process evaluation that explored the effect of DR on caseworker satisfaction and community buy-in; and (3) a cost evaluation that documented start-up costs for DR implementation, and compared initial and follow-up costs for caseworker contact, service provision, and out-of-home placement. The evaluation team used administrative outcome and cost data from Trails, caseworker case-specific reports and caseworker general surveys, family exit surveys, and focus groups and interviews to conduct the process, outcome, and cost evaluations of the DR project in Colorado.

Outcome Evaluation

The evaluation team implemented an experimental design featuring a randomized controlled trial (RCT) to produce net impact estimates for the introduction of differential response in Colorado. Eligible families were randomly assigned to either family assessment response (FAR) or investigation response (IR). The control group received a traditional investigative response including a maltreatment finding and the possible provision of services (after opening a child welfare case). The treatment group received a comprehensive assessment of family needs and strengths. Families assigned to the FAR track did not receive a maltreatment determination. Caseworkers used a menu of family engagement strategies in the delivery of services and ongoing assessment of safety and risk for families in both tracks. These main tenants of the RCT did not change for either track during the course of the study. The outcome evaluation addressed two fundamental questions:

1. Are children whose families are assigned to a family assessment response as safe as or safer than children whose families are assigned to an investigation response?
2. Are there differences between the FAR and IR tracks in terms of family well-being, family engagement, and other factors that may affect child welfare outcomes?

All referrals that were screened out and for which there would be no further involvement by county CPS were not eligible for inclusion in the study. To decide on the eligibility criteria for assignment to the FAR track, the Agency Response Guide was developed to assist in defining the mandatory or discretionary reasons for the traditional investigation of high-risk cases. The following referral reasons were not FAR-eligible: (1) Allegation of serious

harm, (2) Allegation of sexual abuse, (3) Suspicious child fatality or homicide, (4) Institutional referral, and (5) Immediate investigation. Additionally, a referral was ineligible for FAR based on discretionary reasons, which were selected after a thorough review of presenting danger/harm, complicating/risk factors, strengths, cultural considerations, history, and child vulnerability. All remaining referrals were considered to be low- or moderate-risk and were eligible for assignment to the FAR track.

A randomizer tool was used during the study period to determine the track, either FAR or IR, for screened-in referrals that were deemed to be FAR-eligible. A second-stage random assignment identified cases for survey data collection. During the main study period from December, 2010, through February, 2012, there were 10,799 accepted referrals in the five participating counties, of which 5,391 or 50 percent were identified as FAR eligible. Of these referrals, 1,963 (36%) were randomly assigned to the IR track and 3,428 (64%) were randomly assigned to the FAR track. After the removal of a small percentage of ineligible referrals, the final IR sample size was 1,802 cases, the final FAR sample size was 3,194 cases, and the overall sample size was 4,996 cases. Almost all study families (90%) were assessed to be at low or medium risk for future maltreatment. The abuse, neglect, and total risk scores for study families also indicated lower risk.

Process Evaluation

For the process evaluation, the evaluation team used a qualitative research approach with a grounded theory design featuring focus groups, structured interviews, and document review. The process evaluation addressed one fundamental question:

1. What are the perceptions of caseworkers, supervisors, and community stakeholders in regard to caseworker satisfaction and community buy-in related to the implementation of DR in the five counties?

The process evaluation employed a non-random sampling design that was purposeful and convenient to recruit the identified constituent groups, including caseworkers, supervisors, and community stakeholders. The evaluation team conducted two site visits to the participating counties to collect process data through focus groups. The evaluation team conducted 31 focus groups during the Year 1 site visits and 16 focus groups during the Year 2 site visits.

Cost Evaluation

The cost study was designed to answer the following research question:

1. Are there differences in initial, follow-up, and overall costs between families assigned to FAR and IR?

The cost study also detailed the start-up costs for the project. The initial and follow-up costs included personnel costs associated with caseworker contacts, service costs, and OOH placement costs. The initial involvement represented an uninterrupted span beginning with the initial report, continuing through the assessment (for FAR cases) or investigation (for IR cases), and ending with the ongoing service period (if applicable). The follow-up period included subsequent involvement with a family for 365 days after the initial involvement period ended.

Intervention

The Colorado DR Model represents an organizational shift in child welfare agencies in Colorado that impacts all parts of the organization. It includes a series of infrastructure changes within the organizations as well as a deepened and enhanced social work practice set for all staff. Infrastructure refers to the daily processes within the organization (such as supervision and family meetings) that need to be aligned and congruent with new practices in order for any real uptake of organizational change to occur. These may include changes to policy, procedure, and practice.

Colorado implemented a dual-track response system with an investigation response track (now called high risk assessment) and a family assessment response track. The IR track involved a county agency response to higher-risk screened-in allegations of abuse or neglect. Investigative responses involved fact-finding to determine a preponderance of the evidence as to whether or not child maltreatment occurred as alleged. As the non-investigative track, FAR did not require or allow the worker to determine a preponderance of the evidence as to whether or not child maltreatment occurred as alleged. Only cases of lower and moderate risk screened-in allegations of abuse or neglect were eligible for the FAR track. For both the FAR and IR tracks, the services provided include engagement strategies that assist in the assessment of safety, risk, family needs, and family strengths.

The organizational processes of the Colorado DR Model are: enhanced screening, RED teams, group supervision, facilitated family meetings, group supervision, front-loaded services, and support planning. The social work practices of the Colorado DR Model are: rigorous and balanced assessment, strategies for including children, consultation and information sharing

network, evidence-based assessment tools, risk and goal statements, participation of extended networks, and behaviorally-based safety and support plans.

Colorado employed the Colorado Assessment Continuum in both the IR and FAR tracks, which involves conducting safety and risk assessments at or before 30 days from the date of assignment. When safety remained a concern, the caseworker could not end involvement with the family and had the discretion to re-track a referral randomly assigned to the FAR track to the IR track. This occurred in 3% of the FAR assessments. However, in the absence of safety concerns warranting a safety plan or removal, families receiving FAR had the choice of staying involved with the agency by participating in services. FAR assessments closed when the caseworker determined in partnership with the family that the agency and family goals were met and/or the family chose not to complete services. However, in FAR assessments where families needed or volunteered for services, the agency engaged in post-assessment service provision.

Based on findings from the case-specific report, families assigned to the FAR track received significantly more face-to-face and phone contacts than did families assigned to the IR track. Furthermore, families assigned to the FAR track were more likely to receive a service referral and an actual service than were families assigned to the IR track. Caseworkers also reported that families assigned to the FAR track were more likely to receive material needs and mental health services than were families assigned to the IR track. Although FAR families were more likely to report higher satisfaction with services received as compared to IR families, the results from the family exit survey indicate no differences between the tracks on the reported number of services received or the type of services received. According to caseworker perceptions, there was no difference between the tracks for level of service match or for level of service effectiveness. There was no difference between the tracks on dependency and neglect (D&N) filings during initial involvement, traditional child welfare case opened during initial involvement, or out-of-home placement during initial involvement.

Summary of Findings

Overall, the main findings from the outcome, process, and cost evaluations were positive in regard to child safety, family well-being, family engagement, caseworker satisfaction, community buy-in, and cost neutrality. The most promising finding was that there may be long-term child safety benefits and cost savings due to lower levels of re-involvement, over time, for families that receive a family assessment response. The lack of a statistically significant finding for the short-term safety outcomes was not surprising given that the DR system reform also enhanced traditional CPS practices in Colorado.

Child Safety

Based on administrative data findings, the outcome evaluation found no significant differences between the FAR and IR tracks on all of the safety outcomes examined. This includes referral within 365 days of initial referral, assessment within 365 days of initial referral, high risk assessment (HRA) within 365 days of initial referral, founded HRA within 365 days of initial referral, traditional child welfare case opening after initial involvement, and placement in out-of-home care after initial involvement. However, survival analysis findings suggest that FAR families were 18% less likely to have a first HRA, over time, than were IR families. Overall, the child safety findings suggest that family assessment response is as effective in preventing harm as a traditional investigation response for low- and moderate-risk cases. The family exit survey results support the findings from the administrative data, as there were no differences in perceived improvements in child safety between FAR and IR families as reported by caseworkers.

Family Well-being

According to caseworker perceptions, FAR families were more likely to have material needs and mental health needs met than were IR families. There were no statistically significant differences between the tracks in the improvement of family needs, given that a specified need was met. These findings were generally supported by results from the family exit survey, as there was no reported difference between the tracks on whether global family needs were met satisfactorily. When asked directly about improvement in family well-being as a result of their involvement with CPS, again there was no difference between the tracks.

Family Engagement

Overall, the findings suggest that FAR makes a difference in engaging families and enhancing their experience with CPS. Based on findings from the family exit survey, FAR families reported feeling more engaged than did IR families. For example, FAR families were more likely to score their caseworkers high on demonstration of family-centered practice skills than were IR families. Relatedly, FAR families were more likely to rate satisfaction with their caseworker as high. Lastly and maybe most telling, FAR families reported being more likely to call CPS in the future than did IR families. These findings are aligned with the results from the case-specific report. Caseworkers reported that, compared to IR families, FAR families were more likely to show improvement in cooperation, receptivity to help, engagement, and a reduction in difficult behaviors. Furthermore, caseworkers observed that FAR families were more likely to receive higher levels of support from friends and relatives.

Caseworker Satisfaction

Based on the findings from the focus groups conducted during the two site visits, caseworkers and supervisors reported satisfaction with both the initial and ongoing implementation of differential response in their counties and were generally very positive about the Colorado DR model and the system change it engendered. Specifically, there was satisfaction regarding training, supervision, coaching, services, and the relationship between FAR and IR practice. Caseworkers found the practice-focused trainings and Signs of Safety training to be helpful, and reported that the half-day and full-day trainings conducted by state personnel were well-received. Caseworkers appreciated group supervision for the valuable insight provided by diverse perspectives and peer-validated decisions. Caseworkers who received coaching from supervisors, state staff, or other caseworkers found it to be a positive experience. These qualitative findings were supported by results from the caseworker general survey, which indicated that caseworkers and supervisors were somewhat satisfied with FAR in their counties. Caseworkers and supervisors also reported a slightly positive level of satisfaction concerning their child welfare jobs and reported that they are somewhat more likely to stay in the field because of the introduction of DR in Colorado.

Although the overall level of satisfaction was high, caseworkers did express some worries about training, supervision, coaching, and workload. The sequence and scope of training made it difficult to integrate and apply the knowledge gained, especially for IR caseworkers who were perceived to receive less practice specific training content. The biggest challenges for the group supervision model were the logistics, participants, format, consistency, and amount of time required. The main challenge during the second year of implementation was providing sufficient opportunities for formal coaching. After random assignment ended, there was a perceived increase in workload for FAR caseworkers, as more assessments were being assigned to FAR, and the optimal staff and caseload size had not yet been achieved. The overwhelming worry expressed by IR caseworkers was that of burnout, due to the high intensity of the cases being assigned exclusively to them. This was supported by findings from the caseworker general survey, in that caseworkers reported that FAR had slightly increased their caseload size, overall workload, and paperwork.

Community Buy-in

Overall, community stakeholders reported enhanced connections with DHS, which established new avenues of working together to serve children and families within a DR system. However, the biggest challenge of DR implementation in Colorado was for the counties and state to engage community stakeholders in the system reform. For example, there was a general lack of community awareness of the DR project and DR in general during the initial

implementation period. This led to the most demonstrable change observed during the project, as the concentrated efforts to generate community buy-in for DR were perceived to be effective. Another indicator of community buy-in was that families assigned to the FAR track were more likely than IR families to receive higher levels of support from the wider community as reported by caseworkers. That being said, the formal involvement of law enforcement, the courts, mandated reporters, and service providers in differential response is still evolving and faces challenges regarding roles, responsibilities, and expectations.

Costs

According to the comparative analysis findings, there was no statistically significant difference between the two tracks on initial costs. The follow-up costs for IR cases were significantly higher than the follow-up costs for FAR cases. There was no significant difference between the two tracks on overall costs.

Discussion of Findings

This section identifies limitations of the methods, data collection procedures, and data analysis techniques, provides implications for the practice and policy of DR in Colorado, and offers ideas for future research and evaluation of differential response.

Limitations

The limitations for the outcome, process, and cost evaluations are discussed to provide context for the key findings and recommendations and lay a foundation for future research and evaluation in this area.

Outcome Evaluation

This project had many strengths, including the implementation of a randomized controlled trial that allowed for a comparison of outcomes between two equivalent groups; an evaluation that examined the processes, outcomes, and costs of a DR practice model; the integration of multiple data collection methods including surveys, focus groups, interviews, and administrative data; and the utilization of appropriate data analyses techniques such as controlling for sampling bias, non-response bias, pre-existing group differences, and outliers through weighting, data transformation, and survival analyses. However, the outcome evaluation was challenged by several limitations common to child welfare research, such as a reluctance to randomly assign caseworkers to cases, ability to generalize findings to other counties or states, utilization of non-validated instruments, and measurement error including non-response bias. Furthermore, numerous programmatic, procedural, administrative,

supervisory, and practice reforms occurred in conjunction with the installation of two distinct tracks to respond to referrals of child abuse and neglect. As a result, it was somewhat difficult to isolate the relationship between the Colorado DR model and safety outcomes.

Process Evaluation

The process evaluation had several limitations regarding the selection of participants, the scope of the site visits, and the assessment of implementation fidelity. Focus group sampling was purposive and convenient, and not randomized, which could have yielded a sample of individuals more supportive of DR. However, the results show a clear continuum of opinion on the practice. Another limitation was that families were not interviewed for the process evaluation. The two rounds of site visits allowed participants to reflect on the initial stages of DR implementation in Colorado. However, a third site visit in the last year was planned but not implemented because sufficient data for the process evaluation were collected during the first two rounds. The implementation fidelity assessment for the FAR track also did not occur as planned due to resource and time limitations.

Cost Evaluation

There were several limitations to calculating start-up costs, estimating caseworker time, and collecting service cost data. The start-up effort for the CCDR may appear to be more costly and labor-intensive than the other demonstration projects because of the lack of a foundation for developing and implementing the DR model prior to the grant. Estimating caseworker time was the cost study's biggest challenge. The time spent by screeners, RED teams, supervisors, and administrators to manage FAR and IR cases was not included in the estimation of caseworker time, although no difference between the tracks was expected. One limitation to the collection of service costs is that data on contracted service costs paid for outside of child welfare funding were not collected.

Implications

The following practice and policy implications are based on the following key takeaways from the evaluation. First, FAR does no harm to the safety, well-being, and engagement of children and families. Second, a majority of child welfare practitioners, service providers, and community stakeholders were increasingly satisfied with and invested in the Colorado DR model and the DR system reform. Third, there were more caseworker contacts and services provided in a family assessment response at equally or potentially lesser costs.

Policy and Practice

Based on the totality of findings from the outcome, process, and cost evaluation, there was clear balance between the three main focus areas for DR in the five Colorado counties: (1) child safety, child well-being, and family engagement; (2) caseworker satisfaction and community buy-in; and (3) cost neutrality. The main policy implication is that there is evidence to support the expansion of FAR into new counties in Colorado that have sufficient staffing availability, service capacity, and community readiness. However, it is important to stress that if the eligibility criteria for family assessment response changes drastically, this implication may not apply and further evaluation would be necessary.

Five primary practice implications emerged from the study findings. The first is for each county to seek out consensus on the eligibility criteria for family assessment response. The development of the Agency Response Guide yielded more sound eligibility decisions, which was evidenced by the three percent re-track rate for the five counties. This indicates that almost all families were appropriately assigned to FAR based on their level of risk, and that families presenting with higher risk were determined to not be eligible for FAR and thus were assigned to receive a high risk assessment. However, some caseworkers and stakeholders continue to have reservations about certain types of assessments being assigned to FAR, urging ongoing review of the eligibility criteria. Stakeholders also call for more consistency and transparency in the assignment of assessments, including the re-tracking of assessments.

The second implication is to continue enhanced screening by refining the information gathered at this decision point. Most community stakeholders supported enhanced screening and want to see even more information collected at the time of referral. The downside of the new screening procedures was the resistance of some mandatory reporters regarding the increased time and information required to make a referral.

The third practice implication is for counties to pay special attention to managing caseloads and workloads for both IR and FAR caseworkers. For IR caseworkers, there should be a concerted effort to provide them with additional supports to alleviate some of the burden around such activities as court and facilitated family meetings. For FAR caseworkers, there is a need for new functionality in Trails to assist them in the timely documentation of FAR practice, including track change.

The fourth practice implication is for more integrated and seamless service delivery between county DHS and community providers. To keep up with the increasing demand for family-centered services, the state and counties must redouble efforts to increase service capacity by identifying external service providers and enhancing internal service offerings.

There also is a need for greater consistency in service plans, more follow-up on service referrals, and better communication between county DHS and providers.

The final practice implication is to clearly define what is different and what is the same about FAR and HRA casework, roles, and responsibilities. For example, regardless of assigned track, caseworkers try to engage with the family and always look for strengths and resources. Perhaps one of the emergent practice shifts for caseworkers in both tracks is the adoption of techniques promoted by the implementation of solution-focused assessment strategies and safety organized practice.

Expansion and Adoption

The Office of Children, Youth, and Families is in the process of rolling out Governor Hickenlooper's Child Welfare Plan "Keeping Kids Safe and Families Healthy 2.0." This plan includes improvements to the Child Welfare Training Academy; new prevention initiatives such as SafeCare, Community Response, and Nurse Family Partnership – Augmentation; mobile technologies for caseworkers; funding reform; a statewide child abuse and neglect hotline and public awareness campaign; and consistent screening rules and practices (including statewide training for enhanced screening and RED teams). These child welfare reforms align with Colorado's efforts to expand Differential Response and should serve to strengthen the supports for this program.

The primary consideration for new counties adopting DR is to have a clear understanding of the practice model's impact on workload and plan accordingly for training and staffing to minimize the adverse effects of turnover and burnout. The training should be front-loaded to align with implementation of the practice and to limit the training scope to discrete components rather than the overall picture of DR. It is imperative that supervisors receive extensive training in the model to better support caseworkers and serve families. Caseworkers repeatedly suggested that counties preparing to implement DR for the first time should place a heavy emphasis on coaching activities–specifically shadowing–while urging supervisors to focus more on mentoring. It should be noted that coaching is a significant component of the recently revised Colorado Child Welfare Training Academy curriculum.

New jurisdictions should use findings from the cost study to prepare their budgets for DR implementation while being prepared to make adjustments to labor and resource allocations as the practice evolves. Based on the difference between the tracks on follow-up costs, it can be inferred that FAR cases will be less costly than IR cases over time based on lower levels of re-involvement. This may ultimately create cost efficiency for family assessment response in achieving better outcomes at a lower cost. Lastly and perhaps most importantly,

new counties looking to successfully implement DR must earn the support and engagement of the community. To accomplish this, both DHS and partner agencies must commit to an open and honest dialogue about the DR model and its impact on child safety, service capacity, and stakeholder roles. At the state level, lessons learned include establishing a leadership structure that involves teams of individuals rather than a single project director, which allows for continuity of policy and practice.

These implications must be considered in light of concerns about how to allocate the increasingly limited resources in child welfare. Specifically, there are questions about whether it makes sense to target more services and resources to a low-risk population that is unlikely to be re-involved. Stated another way, some argue that scarce services and resources should be given primarily to higher-risk cases who are more likely to come back into the system. This study was not designed to address this concern, but did show that FAR families received more services and caseworker contacts than did IR families, but at a comparable cost. Furthermore, there was a decrease in post-FAR referral acceptance rates, which shows that DR in Colorado is not expanding the population of low- and moderate-risk families served by child welfare. Based on the overall safety, process, and cost results from the study, the question must also be asked about whether the Colorado DR model could benefit high-risk families as well.

Future Research and Evaluation

The three-year program evaluation also revealed avenues for future research of differential response. The hope is that researchers and evaluators will build on this work and continue to generate new evidence on this system reform for the field. Future research on DR must include more of the family voice. A more participatory evaluation design would better include families in all aspects of the evaluation. More attention should be paid to the instruments used to capture family and caseworker perspectives. Validated instruments that measure family engagement should be considered, and new instruments should be evaluated to ensure reliable and valid measurement. Future cost studies should consider administering a data collection template at the beginning of the project to more accurately and efficiently capture level of effort and financial start-up costs. Another area for researchers to build on is examining the relationship between race/ethnicity and safety outcomes.

Researchers also should consider the following suggestions to enhance the study of DR: (1) randomization at the caseworker level to control for differences in experience and skill level or inclusion of a cluster variable to account for differences in caseworker practices; (2) incorporation of implementation fidelity assessments to measure potential contamination of treatment practices; and (3) longitudinal data collection and analysis to explore the long-

term impact of family assessment response on child safety, family engagement, and follow-up costs. Finally, the results of this evaluation should be widely disseminated to child welfare practitioners and stakeholders to provide a common reference point for dialogue on the impact of differential response on child safety, well-being, and permanency outcomes. As such, the DR initiative may ultimately provide an opportunity to redefine the reputation of child welfare in the eyes of parents, practitioners, policymakers, and the public.

Program Evaluation of the Colorado Consortium on Differential Response

1. Background

In 2007, a series of high-profile child fatalities in Colorado prompted numerous examinations of the state's child welfare system. Then-Governor Bill Ritter formed the Governor's Child Welfare Action Committee to conduct a full systemic evaluation. Comprised of various local, state, and national experts on child welfare, this committee met over the course of several years and developed 35 recommendations for improvement. One of the recommendations was Differential Response (DR), which is an innovative system reform that allows child protective services (CPS) to address screened-in allegations of child maltreatment in different ways. Although there is some variation in how CPS systems implement DR, the typical approach includes responding to high-risk cases with a traditional investigation while responding to low- and moderate-risk cases with an assessment of the family's needs and strengths. Safety and risk assessments are completed for all cases.

There have been over a dozen program evaluations of DR in various states. Overall, these evaluations have demonstrated positive outcomes in terms of sustained child safety and improved engagement for families who receive assessments rather than investigations, in addition to enhanced community involvement and worker satisfaction.[1,2,3,4,5,6] However, these

[1] Center for Child and Family Policy, Duke University (2009). Multiple *response system (MRS): Evaluation report to the North Carolina Division of Social Services (NCDSS).* Durham, NC: Author. Retrieved from http://www.ncdhhs.gov/dss/mrs/docs/2009%20MRS%20Report.pdf

[2] Siegel, G. L., & Loman, L. A. (1997). *The Missouri family assessment and response demonstration: Final evaluation report.* St. Louis, MO: Institute of Applied Research. Retrieved from http://www.iarstl.org/papers/MO%20FAR%20Final%20Report-for%20website.pdf

[3] Loman, L. A., & Siegel, G. L. (2004a). *Differential response in Missouri after five years: Final report*. St. Louis, MO: Institute of Applied Research. Retrieved from http://www.iarstl.org/papers/MODiffResp2004a.pdf

[4] Loman, L. A., & Siegel, G. L. (2004b). *Minnesota alternative response evaluation: Final report*. St. Louis, MO: Institute of Applied Research. Retrieved from http://www.iarstl.org/papers/ARFinalEvaluationReport.pdf

[5] Siegel, G. L., & Loman, T. (2006). *Extended follow-up study of Minnesota's family assessment response: Final report.* St. Louis, MO: Institute of Applied Research. Retrieved from http://www.iarstl.org/papers/FinalMNFARReport.pdf

[6] Loman, L. A., Filonow, C. S., & Siegel, G. (2010). *Ohio alternative response evaluation: Final report*. St. Louis, MO: Institute of Applied Research. Retrieved from http://www.iarstl.org/papers/OhioAREvaluation.pdf

studies have been scrutinized for their methodological limitations and the inability to generalize the findings due to variations in how DR was defined and implemented.[7,8]

In October 2008, the American Humane Association (AHA), in partnership with Walter R. McDonald & Associates, Inc. (WRMA) and the Institute of Applied Research (IAR), received a five-year federal award to develop the National Quality Improvement Center on Differential Response in Child Protective Services (QIC-DR).[9] The agreement was awarded by the Children's Bureau of the Administration on Children, Youth and Families, U.S. Department of Health and Human Services. The key goals set forth under this agreement included:

1. Designing and conducting an evaluation to rigorously study implementation, outcomes and cost impact of differential response in research and demonstration (R&D) sites;
2. Learning if differential response is an effective approach in CPS; and
3. Building cutting-edge, innovative, and replicable knowledge about differential response, including guidance on best practices.

The major catalyst for the Colorado Consortium on Differential Response (CCDR) was the opportunity for the Colorado Department of Human Services (CDHS) to apply for a grant from the Children's Bureau through the QIC-DR. CDHS oversees the 64 Colorado county departments of social/human services in their administration of public assistance and child welfare programs, mental health services, juvenile corrections, state veteran nursing homes, services for the developmentally disabled, vocational rehabilitation programs, and senior services. In Colorado, child protective services (CPS) is state-supervised and county-administered with oversight by the Division of Child Welfare Services under the Office of Children, Youth, and Families in the CDHS.

CDHS solicited letters of interest from all 64 counties in Colorado and received five firm commitments. As part of the application to CDHS, counties were required to demonstrate their ability to implement DR by March 2010 and describe how they prepared their respective communities and staff for the shift in practice. CDHS also required each county to provide a

[7] Hughes, R. C., Rycus, J. S., Saunders-Adams, S. M., Hughes, L. K., & Hughes, N. K. (2013). Issues in differential response. *Research on Social Work Practice*, *23*, 493-520.

[8] Hughes, R. C., & Rycus, J. S. (2013). Discussion of issues in differential response. *Research on Social Work Practice*, *23*, 563-577.

[9] In 2012, the QIC-DR grant was relinquished by the American Humane Association. Since that time, the Kempe Center for the Prevention and Treatment of Child Abuse and Neglect at the University of Colorado Denver has managed the QIC-DR along with the support from the same partners.

mandatory letter of support from local law enforcement agencies endorsing the system reform to be implemented through DR and an agreement from county administration to provide a financial match. The state requested a commitment from local law enforcement because Colorado statute provides a standard for minor child abuse offenses that intersected with the population eligible for DR in some cases.

Administrators from the five committed counties saw DR as a logical extension of their current practice. Additionally, these five counties self-assessed that they were poised for a system reform of this magnitude. The counties already possessed a general familiarity with DR via articles, conferences, and talking with champions of the approach from other states. Additionally, information and consultation from individuals with experience in Minnesota influenced the initial consideration of DR for Colorado. In 2009, the QIC-DR selected Colorado, Illinois, and Ohio as the three Research & Development sites. The grantee for the Colorado site was CDHS, in conjunction with Arapahoe, Fremont, Garfield, Jefferson, and Larimer counties.

1.1. Making the Commitment to DR

To establish a strong commitment to DR among community stakeholders, Colorado appointed a Child Protection Manager to the CCDR, and identified county leaders for the initial consideration of the system reform. County Department of Human Services (DHS) directors and management teams, including administrators and supervisors, became the identifiable leaders during the commitment phase of the project. Key stakeholders included mental health providers, law enforcement personnel, probation departments, district attorney offices, judicial services, magistrates, school districts, child protection teams, and county commissioners. Temporary Assistance to Needy Families (TANF) personnel also were cited as key stakeholders because a risk factor for neglect is the inability to provide necessary food, clothing, and shelter.

Based on the experiences of other jurisdictions, the CCDR began under the assumption that community outreach and systematic response to community concerns would be an integral part of the work, particularly in the pre-implementation stage. The community outreach's primary purpose was to promote community buy-in and education. To that end, the Consortium held meetings and presentations in addition to frequent informal conversations with community stakeholders. Stakeholder meetings included treatment providers, school districts, law enforcement, mandated reporters, judges, and guardian ad litems (GAL). Trainings and presentations were provided to these same stakeholder groups, along with hospitals, Court Appointed Special Advocates (CASA), court services providers, public health agencies, the workforce center, Division of Youth Corrections (DYC), the District Attorney's (DA) Office, the SB94 (alternatives to detention for youth) Coordinator, the collaborative management team,

the Council of Governments, Boys and Girls Clubs, Head Start, school district leadership, faith-based groups, and interagency practice teams consisting of providers such as mental health, day treatment, domestic violence, and county-level child protection teams.

1.2. CPS System Prior to DR

Prior to the implementation of DR in the five participating counties, the CPS system consisted of one main intake response to screened-in referrals of maltreatment. Child welfare responses to abuse and neglect are considered Program Area 5 (PA5) in the Colorado system, whereas Program Area 4 (PA4) is reserved for children deemed beyond the control of their parents and/or involved in the juvenile delinquency system. A case is considered PA4 when the assessment of family and child needs and strengths does not result in a finding of abuse or neglect. However, there is much overlap between Program Areas 4 and 5, with assessments in PA5 being assessed as more appropriate for PA4 and vice versa.

1.3. Concurrent Child Welfare Reforms

The Division of Child Welfare (DCW) within CDHS uses collaboration to enhance practice and improve outcomes for children and families within the state. The DCW has initiated and facilitated numerous collaborative efforts, such as the Collaborative Management Program (HB1451). Within this statutory initiative, counties that choose to participate in the program are required to partner with one another to streamline all systems and achieve better access to services for children and families. The DCW program staff also facilitates the Collaborative Management Steering Committee and works in partnership with other collaborative child safety and permanency initiatives.

This strong push to reform the child welfare system intersected with other initiatives, including Family to Family (Annie E. Casey Foundation), Signs of Safety™, and Family Treatment Drug Courts. The DR initiative was generally viewed as complementary with these concurrent child welfare reforms, as they are all models that seek to increase engagement with families and collaboration across systems and with the community. Two initiatives in particular strongly aligned with differential response: the Colorado Practice Model (CPM) and the Colorado Disparities Resource Center (CDRC).

The Division of Child Welfare enlisted the assistance of the Mountains and Plains Child Welfare Implementation Center to develop, implement, and evaluate an overall practice model for child welfare in the state. The first phase of implementation strongly emphasized continuous quality improvement and the use of data in system reform. The first cohort for the CPM consisted of thirteen counties and one tribal community. Following the beginning of their

work and the instillation of quality practice teams, the five counties in the CCDR were asked to join in the leadership and steering of this reform. CPM in particular appeared to be an appropriate fit for differential response, and its practice components may be designated as an official "Promising Practice" in Colorado. This designation will be informed by the findings from the process, outcome, and cost evaluations for the DR project.

The CDRC was formed in 2009 through a partnership between AHA and DCW to address disparate and disproportionate outcomes for children of color in the child welfare system. In 2010, the CDRC approached CCDR to enlist participation in a series of learning communities on this issue. Since that time, the five CCDR counties have reviewed their data, and a few have taken steps beyond the differential response implementation to address outcomes. Participating counties have started to pay special attention to the service array for children in their communities, including the development of town hall meetings to address this issue. The learning communities concluded in April 2012.

1.4. Local Context

The five participating counties range in size from metropolitan to rural and are located in four of the five regions of Colorado. The five counties vary in geography, population, and diversity.

Arapahoe County is a county in the Denver metropolitan area with a child population of 142,851. The main cities are Aurora, Littleton, and Englewood. Arapahoe County has some very impoverished areas as well as the wealthy neighborhood of Cherry Hills Village. There are nine law enforcement jurisdictions. As displayed in Table 1.1, the child abuse and neglect referral acceptance rate in 2009 was 51 percent of 8,075 referrals received with seven percent of assessments opened into ongoing service cases. The primary family engagement effort by Arapahoe County DHS is the LINKS Program, which is a facilitated family engagement process designed to enhance the transition between intake and permanency (ongoing) caseworkers. The LINKS Program also promotes family-driven services and permanency planning.

Fremont County is the largest of the mid-sized counties in Colorado with a child population of around 8,500. Fremont County is located two hours southwest of Denver and lies at the foothills of the Rocky Mountains. The county seat is Canon City, and the Royal Gorge is a popular draw for visitors. Much of the local economy is supported by 15 prisons, including a Federal Supermax prison. There are three law enforcement jurisdictions. As displayed in Table 1.1, the child abuse and neglect referral acceptance rate in 2009 was 50 percent of 720 referrals received with 23 percent of assessments opened into ongoing service cases. Fremont

County DHS is involved with their community in a positive way, which was largely spurred by their work in the Collaborative Management Program and Family to Family.

Garfield County is a mid-sized county with a child population of approximately 12,000. Located three hours west of Denver across the Rocky Mountains, the main cities are Glenwood Springs and Rifle. Given its proximity to the resort community of Aspen and the draw of Glenwood Springs, Garfield County is diverse in income levels. There are seven law enforcement jurisdictions. As displayed in Table 1.1, the child abuse and neglect referral acceptance rate in 2009 was 69 percent of 350 referrals received with 12 percent of assessments opened into ongoing service cases. Garfield County DHS has been involved in the Collaborative Management Program and Family to Family. Their facilitated family meetings process includes Team Decision Making (TDM), family group conferencing, and family support planning meetings.

Jefferson County is a county in the Denver metropolitan area with a child population of 113,070. The main cities are Arvada, Golden, Lakewood, and Wheat Ridge. It also includes the mountain communities of Conifer and Evergreen. There are 12 municipal law enforcement jurisdictions. As displayed in Table 1.1, the child abuse and neglect referral acceptance rate in 2009 was 63 percent of 8,602 referrals received with 10 percent of assessments opened into ongoing service cases. Jefferson County DHS has been involved in various initiatives and projects over the years, including Systems of Care, Family to Family, and the Collaborative Management Program. Jefferson County has an active Parent Partners Program that assisted in the cognitive testing of the family exit survey used in this evaluation and provided valuable feedback prior to the implementation phase of the project. Additionally, Jefferson County completed a full-scale implementation of TDM prior to implementing DR.

Larimer County is a Front Range county with a child population of 64,000. Larimer County is located one hour north of Denver and is bordered on the north by Wyoming and on the west by the Rocky Mountains. The main cities are Fort Collins, Loveland, and Estes Park. There are eight law enforcement jurisdictions. As displayed in Table 1.1, the child abuse and neglect referral acceptance rate in 2009 was 45 percent of 5,589 referrals received with 27 percent of assessments opened into ongoing service cases. Larimer County DHS has a well-developed facilitated family meeting strategy that supports serving families at multiple points from assessment phase to ongoing involvement. The county also underwent extensive consultation with Minnesota practitioners during the pre-implementation phase of the CCDR.

Table 1.1

Child Population and 2009 Child Abuse and Neglect Referrals Statistics for CCDR Counties

County	Child Population	Number of Referrals	Referrals Accepted	Assessments Opened to Ongoing Services
Arapahoe	142,851	8,075	51%	7%
Fremont	8,500	720	50%	23%
Garfield	12,000	350	69%	12%
Jefferson	113,070	8,602	63%	10%
Larimer	64,000	5,589	45%	27%

1.5. Terminology

Given the variety of naming conventions used nationally in differential response, it is important to highlight and define these terms from the Colorado perspective.

- *Differential Response* – describes the overall system change necessary to implement a dual-track response system in child protective services.
- *Differential Response Model* – sets forth numerous *organizational processes* including 1) enhanced screening, 2) Review Evaluate and Direct (RED) teams, 3) group supervision, 4) facilitated family meetings, 5) front-loaded services, and 6) support planning. It also identifies *social work practices* such as 1) rigorous and balanced assessment, 2) strategies for including children, 3) Consultation and Information Sharing Framework, 4) evidence-based assessment tools, 5) risk and goal statements, 6) participation of extended networks, and 7) behaviorally-based safety and support plans that impact assessment and service delivery regardless of track assignment.
- *Investigative Response (IR)* – is a track of the dual-track response system in which county agencies respond to high-risk screened-in allegations of abuse or neglect. This response includes worker engagement strategies that assist in the assessment of safety, risk, family needs, and family strengths. Investigative response involves fact-finding to determine a preponderance of the evidence as to whether or not child maltreatment has occurred. This track is now referred to as High Risk Assessment (HRA).
- *Family Assessment Response (FAR)* – is a track of the dual-track response system in which county agencies respond to eligible low- and moderate-risk screened-in allegations of abuse or neglect. FAR is the non-investigative track of the dual track response system in that it does not require or allow the worker to determine a preponderance of the evidence as to whether or not child maltreatment occurred as alleged. This response includes engagement strategies that assist in the assessment of safety, risk, family needs, and family strengths.

2. Methodology

The Social Work Research Center (SWRC) in the School of Social Work at Colorado State University (Fort Collins, CO) and Westat (Rockville, MD) were selected by CDHS as the Colorado evaluation team. The required evaluation approach addressed three main focus areas: (1) an outcome evaluation that rigorously examined the impact of DR on key child welfare outcomes including child safety and family engagement; (2) a process evaluation that explored the effect of DR on caseworker satisfaction and community buy-in; and (3) a cost evaluation that documented start-up costs for DR implementation, and compared initial and follow-up costs for caseworker contact, service provision, and out-of-home placement. This chapter details the methodology for the outcome and process evaluations. The methodology for the cost evaluation is explained in Chapter 6.

2.1. Outcome Evaluation

This section describes the study design, research questions, eligibility criteria, random assignment procedures, random assignment monitoring, sampling, data collection, data analysis, and study logistics for the outcome evaluation.

2.1.1. Study Design

The evaluation team implemented an experimental design featuring a randomized controlled trial (RCT) to produce net impact estimates for the introduction of differential response in Colorado. Eligible families were randomly assigned to either family assessment response or investigation response. As part of the evaluation, the evaluation team measured the efficacy of the intervention using the following post-test only, control group design:

R	X	O_1
R		O_2

where R represents families randomly assigned to either the treatment (FAR) or control (IR) group; X represents the intervention of differential response; O_1 is the treatment group's outcomes and O_2 is the control group's outcomes.[10] Because outcomes were measured after the introduction of the intervention, this is called a posttest design.

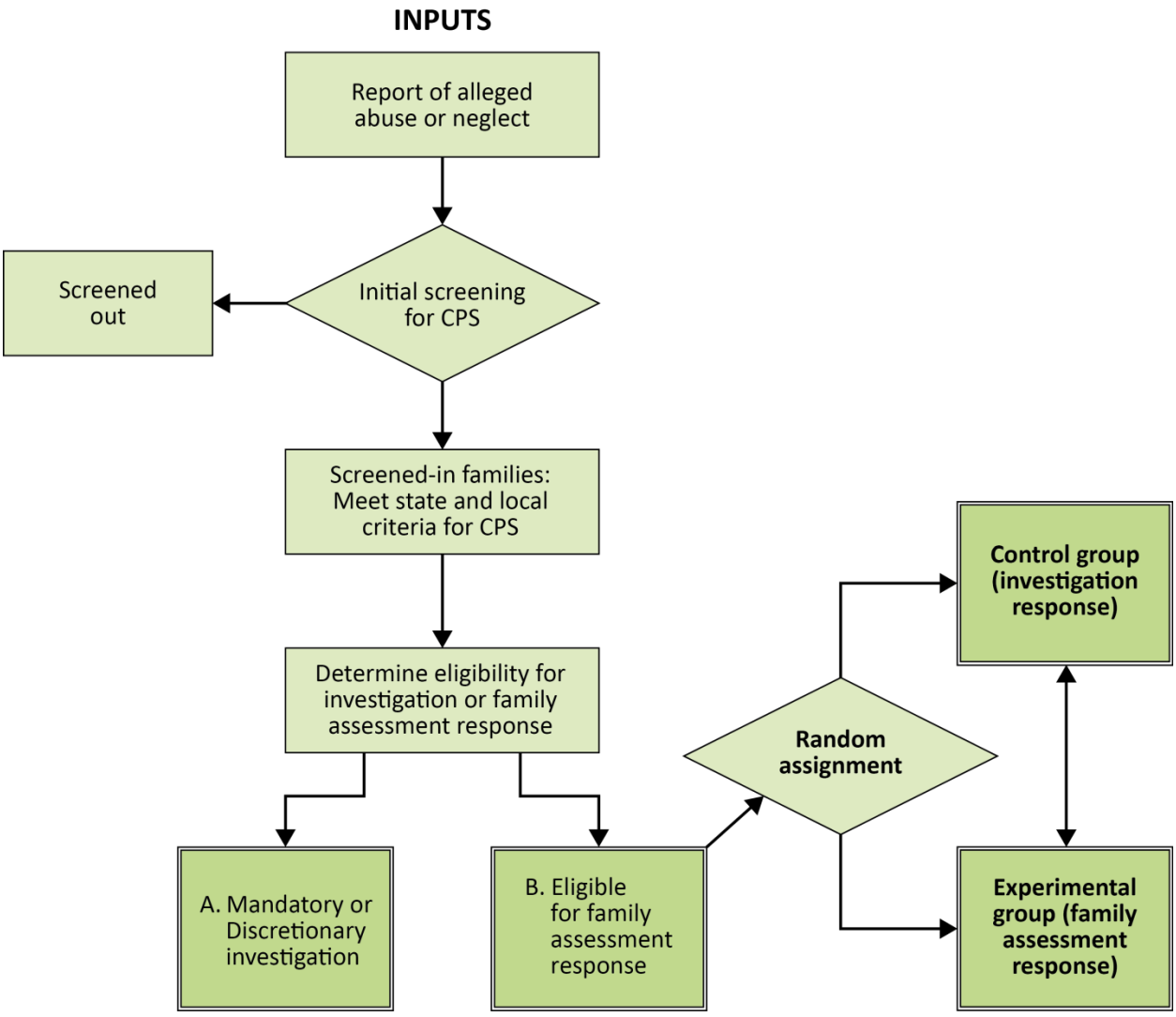
The control group received a traditional investigative response including a maltreatment finding and the possible provision of services (after opening a child welfare case). The treatment group received a comprehensive assessment of family needs and strengths. Families assigned to the FAR track did not receive a maltreatment determination. Caseworkers used a

[10] Royse, D., & Thyer, B. A. (1996). *Program evaluation: An Introduction (2nd edition).* Chicago: Nelson-Hall Publishers.

menu of family engagement strategies in the delivery of services and ongoing assessment of safety and risk for families in both tracks. These main tenants of the RCT did not change for either track during the course of the study. Figure 2.1 illustrates the process each site followed in identifying and randomly assigning eligible families for the study.

Figure 2.1

Randomized Controlled Trial Process



2.1.2. Research Questions

The outcome evaluation addressed two fundamental questions:

1. Are children whose families are assigned to a family assessment response as safe as or safer than children whose families are assigned to an investigation response?
2. Are there differences between the FAR and IR tracks in terms of family well-being, family engagement, and other factors that may affect child welfare outcomes?

2.1.3. Eligibility Criteria

To decide on the eligibility criteria for track assignment (see A and B in Figure 2.1), it was necessary to define low- and moderate-risk cases in comparison to high-risk cases.[11] The Screening and Referral workgroup developed the Agency Response Guide (see Appendix A) to assist in defining the mandatory or discretionary reasons for an investigation of high-risk cases. This was an important exercise from both a practice and outreach perspective as caseworkers and stakeholders needed to be on the same page regarding the eligibility determination.

In Colorado, each county DHS child welfare office receives the vast majority of reports of child abuse and neglect, and each report is subject to the Code of Colorado Regulations, Volume 7. A report was referred for one of the following decisions after an initial review by a supervisor: 1) immediate investigation in the event of serious bodily injury or high risk of abuse or neglect; 2) investigation within three calendar days or 5 business days of receipt of the report; or 3) screened-out for either community agency referral or no further action by the department. Information about each referral is entered into SACWIS (called Trails in Colorado) and an initial response assessment is completed.

All referrals that were screened out and for which there would be no further involvement by county CPS were not eligible for the study. For purposes of this project, any report referred for immediate investigation under Volume 7 was considered a mandatory investigation referral and was not eligible for the FAR track. Likewise, the following referral reasons also were not FAR-eligible:

1. Allegation of serious harm
2. Allegation of sexual abuse
3. Suspicious child fatality or homicide
4. Institutional referral

[11] The term "case" is used throughout the report for simplicity and consistency reasons. It can refer to a referral, an assessment, or a case depending on the nature of the variable, outcome, or analysis being described.

Additionally, a referral could be ineligible for FAR for the following discretionary reasons, which were selected by the team after a thorough review of presenting danger/harm, complicating/risk factors, strengths, cultural considerations, history, and child vulnerability.

1. Currently open investigation response
2. Frequent, similar, recent referrals
3. Violent activities in the household
4. Caregiver declined services in the past
5. Caregiver unwilling/unable to achieve safety
6. Past safety concerns not resolved
7. Previous serious child harm offenses
8. Credible reporting party alleges high safety concern
9. High child vulnerability
10. Substance abuse not manageable through FAR
11. Domestic violence not manageable through FAR
12. Court ordered investigation
13. FAR eligible, approved exemption – staffing
14. Not in FAR county jurisdiction
15. Randomizer down – project director notified
16. Insufficient information to assess for FAR eligibility

All remaining referrals were eligible for the FAR track and were subject to random assignment. The criteria for cases eligible for assignment to the FAR track included, but were not limited to, the following:

1. Mild to medium neglect
2. Educational neglect
3. Mild to moderate neglect from and Injurious environment due to domestic violence
4. Mild to moderate physical abuse

2.1.4. Random Assignment Procedures

A randomizer tool was used during the study period to determine the track, either FAR or IR, for screened-in referrals that were deemed to be FAR-eligible. The randomizer was a simple web-based application developed by the cross-site evaluator (WRMA) and accessed through a secure, encrypted website. Users entered a referral ID, county, and the first and last name of the primary caregiver into the program. The application then randomly assigned the case to the FAR or IR track based on the county's assignment ratio. A second-stage random assignment identified cases for survey data collection. The second stage sampling was done to

reduce the data collection burden on caseworkers. A sample ID was subsequently generated, and the user then entered the randomizer results into the track assign field in the Trails referral acceptance window. Caseworkers and screeners primarily used the random assignment results in Trails to check whether a family was already in the sample.

As displayed in Table 2.1, the proportion of cases assigned to the FAR and IR tracks was set to meet county requirements. In Larimer County, 20 percent of the FAR-eligible cases were randomly assigned to IR while 80 percent were assigned to FAR. The ratios were 50 percent FAR and 50 percent IR in the other four counties, although the proportion in Arapahoe County was adjusted to 40 percent IR and 60 percent FAR in January 2011 to balance workload across the two units. For the last three months of the study, the random assignment ratios were changed in favor of the FAR track so counties could adjust their staffing needs in anticipation of a higher rate of FAR cases after random assignment ended.

Table 2.1

Random Track Assignment and Random Survey Assignment Ratios

County	Start Date	Probability of Assignment to FAR	Probability of Survey Assignment - FAR	Probability of Survey Assignment - IR
Arapahoe	12/1/2010	50.0%	34.8%	34.8%
	1/7/2011	60.0%	29.4%	42.9%
	12/1/2011	70.0%	25.0%	57.1%
	1/3/2012	80.0%	22.2%	85.7%
	2/1/2012	90.0%	20.0%	100.0%
Fremont	12/1/2010	50.0%	100.0%	100.0%
	12/1/2011	75.0%	66.7%	100.0%
Garfield	12/1/2010	50.0%	100.0%	100.0%
	12/1/2011	75.0%	66.7%	100.0%
Jefferson	12/1/2010	50.0%	40.0%	40.0%
	12/1/2011	63.6%	30.8%	55.6%
	1/3/2012	75.0%	26.7%	80.0%
	2/1/2012	85.7%	23.1%	100.0%
Larimer	12/1/2010	80.0%	25.0%	100.0%
	12/1/2011	90.0%	22.2%	100.0%

2.1.5. Random Assignment Monitoring

The random assignment of cases was monitored to ensure that the results adhered to the programmed county assignment ratios. A randomizer exception report was continually generated to identify discrepancies between the randomizer results and the randomizer data

entered into Trails. Over a short period of time during the study, one county had several instances of a mismatch between the randomizer and Trails assignment. The evaluation team and Jefferson County senior data analyst discussed this concern with the county, and the issue was resolved. The evaluation team worked with the Trails design team to make changes as needed to the randomization procedure. Caseworkers accepted the use of the randomizer for the study, and according to the ad hoc reports, there were no problems with adherence to the protocols or assignment determinations. To address perceptions of workload changes related to the randomizer, adjustments were made early on in some counties by using blocked randomization. This prevented a long run of assignments to one track to avoid an imbalance in workload for a caseworker or team.

2.1.6. Sampling

This section describes the sampling for the pilot study, main study, and survey data collection.

2.1.6.1. Pilot study sample

A pilot study was conducted for two months prior to the start of the main study to allow for the utilization of the randomizer and the administration of the family exit survey and case-specific report. During the pilot period from October 5, 2010, to November 30, 2010, there were 1,523 accepted referrals in the five counties, of which 731 or 48 percent were identified as FAR eligible. Of the FAR eligible cases, 440 (60%) were assigned to the FAR track and 291 (40%) were assigned to the IR track.

2.1.6.2. Main study sample

As displayed in the CONSORT diagram in Appendix B, there were 10,799 accepted referrals in the five participating counties, of which 5,391 or 50 percent were identified as FAR eligible, during the main study period from December, 2010, through February, 2012. Of these referrals, 1,963 (36%) were randomly assigned to the IR track and 3,428 (64%) were randomly assigned to the FAR track. Approximately 7% of referrals were ultimately removed from the main study sample due to the following four reasons: (a) referral should have been screened-out, (b) referral should have had a required investigation, (c) referral was from a county not in the study, and (d) referral did not indicate anyone participating as a child (PAC). After the removal of 234 ineligible referrals, the final FAR sample size was 3,194 cases. After the removal of 161 ineligible referrals, the final IR sample size was 1,802 cases. The final overall sample size was 4,996 cases.

2.1.6.3. Survey sample

Since the percentage of cases assigned to each group was not the same in all counties, the random selection of cases for survey data collection attempted to gather equal numbers of FAR and IR cases for the survey sample within each county. As displayed in Table 2.1, a random sampling methodology was employed to select cases for data collection. For Arapahoe, 28 percent of FAR cases and 45 percent of IR cases were sampled for data collection. For Jefferson, 36 percent of FAR cases and 44 percent of IR cases were sampled for data collection. For Larimer, 24 percent of FAR cases and 100 percent of IR cases were sampled for data collection. All FAR and IR main study cases in Fremont and Garfield counties were sampled for data collection. The data collection ratios changed slightly in all five counties for the last three months of the study. This allowed for a higher percentage of IR cases to be sampled given the change to higher assignment ratios for the FAR track.

2.1.7. Quantitative Data Collection

The evaluation team used administrative outcome data from Trails, caseworker case-specific reports, caseworker general surveys, and family exit surveys to conduct the outcome evaluation. The instrument development process for the family exit survey is described in Appendix C. This section of the report details the data collection procedures for the family exit survey, case-specific report, and caseworker general survey (See Appendix D for more detail on these data sources).

2.1.7.1. Pilot study data collection

The evaluation team pilot-tested the caseworker case-specific report and family exit survey data collection procedures during a two-month period between October 1 and November 30, 2010. For cases assigned during the pilot period, 241 family exit surveys were mailed and 39 were completed and returned for a response rate of 16 percent. In addition, 298 case-specific reports were administered during the pilot period with 171 completed for a response rate of 57 percent. Subsequently, the main study data collection began in December 2010. Pilot data were not included as part of the main study data analysis. The pilot study ensured the stability and proper functioning of data collection procedures before moving ahead with main study data collection.

2.1.7.2. Family exit survey

The evaluation team flagged families sampled for survey data collection in Trails and generated reports upon case closure to identify the selected families ready to be surveyed. At the last meeting with the family, caseworkers were asked to encourage the family to complete the family exit survey. The evaluation team provided families with a prepaid $10 incentive in an

effort to promote the completion of the survey. On a weekly basis, the Jefferson County senior data analyst sent an Excel file to the evaluation team with the names and contact information for the families selected for data collection whose cases closed within the previous week.

The evaluation team made telephone calls to inform families that the survey and incentive were being mailed. They also provided an option to conduct a phone survey at that time or a time of the family's choosing. If accepted, the evaluation team read the survey to the family member and recorded the results for processing. The evaluation team mailed the family exit survey to the selected families that did not request a phone survey. A consent form accompanied the survey along with a business reply envelope so that the family could easily return the materials to the evaluation team, who receipted and processed them.

The evaluation team made reminder calls to families who did not return the survey within four weeks and offered the option to complete the survey with a trained telephone interviewer. Families received a second follow-up call if they did not return the survey within six weeks after the case closed. Families indicating they did not receive the initial survey received a second mailing with the incentive.[12]

The evaluation team mailed out the family exit surveys through December 2012. For cases selected for the main study, 1,089 surveys were mailed to FAR families and 257 completed surveys were returned (24% response rate), while 968 surveys were mailed to IR families and 206 were returned (21% response rate). This includes all completed surveys received after the first mailing, those received following the second mailing, and those completed by phone. Overall, 2,057 family exit surveys were administered with 463 completed (23% response rate). See Appendix E for the family exit survey.

2.1.7.3. Case-specific report

The evaluation team trained caseworkers in each county to complete the web-based case-specific reports. At case closure (based on the date the case closed in Trails), an email alert was sent to the last caseworker for the case. The email contained a link to the report, which was hosted on Limesurvey through Jefferson County DHS. Caseworkers received a follow-up email after two weeks and again after four weeks if the report was not completed.

The evaluation team collected the caseworker case-specific reports through December 2012. For cases selected to the main study, 1,089 FAR case-specific reports were administered

[12] In fall 2012, the evaluation team experimented with adding an additional mailing to non-respondents, with an offer of a second incentive if they mailed backed the survey. This option was added for all non-respondents throughout the remainder of the data collection period.

to caseworkers and 871 completed reports were received (80% response rate), while 968 IR case-specific reports were administered to caseworkers and 801 completed reports were received (83% response rate). Overall, 2,057 case-specific reports were administered with 1,672 completed (81% response rate). See Appendix F for the case-specific report.

2.1.7.4. Caseworker general survey

All active caseworkers and supervisors as of October, 2011, who had at least one FAR or IR case during the study timeframe, were eligible to complete the caseworker general survey. At the time the survey was fielded, the Jefferson County senior data analyst sent an email to the eligible caseworkers with a link to the survey, which also was hosted on Limesurvey through Jefferson County DHS. A follow-up email was sent to caseworkers after four weeks if they had not yet completed the survey.

For eligible caseworkers, 143 surveys were administered with 89 completed (62% response rate). For eligible supervisors, 39 surveys were administered with 30 completed (77% response rate). Overall, 182 general surveys were administered with 119 completed (65% response rate). See Appendix G for the caseworker general survey.

2.1.8. Quantitative Data Analysis

For the quantitative data analysis, the following procedures were employed for quality assessment, case assignment errors, handling missing data, preparing the quantitative data, and estimation and weighting. The analytical approaches for intent-to-treat and quantitative data also are described.

2.1.8.1. Quality assessment

The evaluation team worked with the Jefferson County senior data analyst to conduct quality assessment (QA) on data collected from Trails. In addition, all administrative and survey data submitted to WRMA went through rigorous QA checks to ensure compliance with the record extract layouts and mapping forms. The following QA strategies were employed:

1. Examine frequencies and distributions for multiple data fields;
2. Conduct manual quality checks of the data (e.g., manual look-ups to supplement information learned from aggregated indicators);
3. Maintain regular communication with caseworkers and supervisors about how and when certain data were entered and updated, and how particular fields were interpreted and used for case management and planning.

2.1.8.2. Case assignment errors

The evaluation team received weekly reports from the Jefferson County senior data analyst with the IDs of assigned cases, the pathway assignment of each case, and whether the case was sampled for follow-up data collection. The evaluation team compared the two files to make sure that the data entered into Trails matched that from the randomizer. The evaluation team immediately notified DHS of any errors so corrections could be made expediently. The evaluation team also scrutinized all violations that occurred (e.g., case assignment incorrectly recorded in Trails) to determine whether additional training of workers in data entry was required or if the random assignment process was possibly being undermined by workers' desires to make exceptions. The evaluation team found no evidence of this during the study.

2.1.8.3. Estimation and weighting

The probability of assignment to FAR differed among counties and over time.[13] The survey weights adjusted for different assignment probabilities so that the weighted distribution of FAR cases among counties and over time was the same as both the distribution of all cases across counties and over time and the distribution of IR cases across counties and over time. As a result, a weighted mean (or percentile) based on either FAR or IR cases, estimates the mean that would have been obtained if all cases had been assigned to FAR or IR, respectively. The weighted difference between FAR and IR cases is the difference that would be observed if FAR had been used for all cases as opposed to using IR for all cases.

Different weight variables were used for analysis of the administrative data, the case-specific reports, and the family exit survey. For analysis of the administrative data, the analysis weight for a FAR case is the inverse of the probability of assigning the case to FAR. Similarly, the weight for an IR case is the inverse of the probability of assigning the case to IR. The weights for analysis of the family exit survey and the case reports have two components: one based on the probability that a case is selected for survey data collection, and one based on the probability that a completed survey is received (for those selected for data collection) to adjust for non-response. Analysis of weighted survey data requires procedures that use the weights and calculate the appropriate standard errors and *p*-values. Appropriate analysis procedures were used for all of the results presented in this report. Appendix H has details on weight calculation.

2.1.8.4. Missing data

For the cases selected for data collection, data may be missing for all questions or variables (called survey non-response), or data may be missing for only some questions (called

[13] For example, 28% of all FAR cases were from Larimer County and 10% of all IR cases were from Larimer County; however only 21% of cases assigned to the study were from Larimer County.

item non-response). Missing data are not problematic for analysis if it occurs randomly. Systematically missing data, however, may result in biased estimates. In general, the magnitude of the bias, if present, increases as the proportion of missing values increases. The possible bias due to non-response is minimized by the non-response adjustment to the survey weights. For item non-response, the prediction formula can be used to estimate the missing response if the variable with the missing responses can be reasonably predicted from other variables for which the data are not missing. This process for estimating the missing responses is called imputation. The study team reviewed the survey data and determined that since missing data was minimal, imputation was not necessary.

2.1.8.5. Quantitative data preparation

For the family exit survey, the evaluation team developed scale scores using items that appeared to be related. Reliability of scales consisting of multiple items was evaluated by calculating a Cronbach's alpha, or internal consistency reliability, for each scale. Factor analysis confirmed each scale was measuring a distinct construct. The Cronbach's alpha for each scale and the final data analysis matrix for the family exit survey are included in Appendix I.

Many case-specific survey items were recoded prior to analysis. Most recoding simplified outcomes to represent two levels: a relatively worse outcome and a relatively better one.[14] There were several reasons that such recoding facilitated analysis. First, most of the case-specific survey items had only four levels. An item should have at least five levels in order to entertain the possibility that the data can be treated as continuous or approximately normally distributed.[15,16] Even if four categories were enough, the response data from the case-specific survey items are often non-normal; responses for many items were either flat or markedly skewed. If the four-level items had not been dichotomized, the evaluation team would have used either cumulative logit regression[17] or non-parametric analysis. Cumulative logit modeling would be a rather complex method requiring substantial explanation for readers. Non-parametric analysis would not allow for the control of demographic variables which are significantly different between FAR and IR. The use of dichotomous coding allowed the

[14] For example, a four-level item recording the level of improvement in family functioning was recoded to indicate a 'no/little' or 'moderate/much' level of improvement. The original responses of 'none' and 'little' correspond to 'no/little' improvement and responses of 'moderate' and 'much' correspond to 'moderate/much' improvement.

[15] Morgan, G. A., Leech, N. L., Gloeckner, G. W. & Barrett, K. C. (2007). *SPSS for introductory statistics* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

[16] Trochim, W. M. K., & Donnelly, J. P. (2008). *The research methods knowledge base* (3rd ed.). Mason, OH: Cengage Learning.

[17] Agresti, A. (2007). *An introduction to categorical data analysis* (2nd ed.). Hoboken, NJ: John Wiley & Sons.

evaluation team to balance considering distributional assumptions and controlling for covariates, while still providing a clear and useful analysis for the reader. The final data analysis matrix for the case-specific report is included in Appendix J.

2.1.8.6. Intent-to-treat

To maintain the integrity of the experimental design, an "intent-to-treat" (ITT) approach was employed, whereby families originally assigned to the treatment condition remained a treatment case and families originally assigned to the control condition remained a control case for all data analyses.[18] Families assigned to the FAR track that crossed over to the IR track because of a track change remained in the FAR group for analytic purposes. The ITT approach also accounts for any families randomized to the IR track but erroneously assigned to the FAR track (or vice versa), which was considered a randomizer data entry error. This occurred for three families that were erroneously assigned to the FAR track and 23 families that were erroneously assigned to the IR track during the study period.

Although ITT analysis diminishes analytic power, it avoids spuriously basing conclusions on characteristics related to attrition and crossover. The ITT approach allows for inferences that observed differences between treatment and control families after the program are, in fact, due to the program, with a small and known probability of error. Data completeness is particularly important for ITT analyses because the "problem" cases (e.g., drop-outs, crossovers) are those most likely to be associated with missing data. With other approaches, such as eliminating drop-out families from the analysis, there is a risk of comparing the most motivated treatment families with all the control families, thus incorrectly attributing observed differences to program effects.

2.1.8.7. Analytical procedures

Differences in data type determined which analytical procedures were used. Categorical data represents discrete characteristics of cases, such as the race of the youngest child or the Yes/No response to a question. Descriptive statistics used for categorical data include frequencies and percentages. A chi-square statistic was used for comparing the counts between the FAR and IR cases. Logistic regression was used for more complicated models predicting categorical outcomes. Counts such as the number of face-to-face meetings could be recoded into several ordered categories (e.g., "1", "2 to 5", or "6 or more"). Ordinal logistic regression was used to model ordered categorical outcome variables.

[18] Hollis, S., & Campbell, F. (1999). What is meant by intention to treat analysis? Survey of published randomised controlled trials. *BMJ,* 319, 670-674.

Numerical data includes numerical values that can have many distinct values (and possibly an infinite number of distinct values), such as counts (e.g., number of referrals or days from assignment to the first referral after the initial referral). Descriptive statistics used for numerical data include frequencies, means, medians, modes, and ranges. Differences between the FAR and IR cases were assessed using means and t-tests. Linear regression was used for more complicated models predicting numerical outcomes from multiple predictors. In some cases, the predictor or outcome variables were coded into discrete categories or transformed (such as with a log transformation) to reduce the influence of a few extreme values on the model assessment. The linear regression model predicts the mean of the outcome data. For time-to-event data (e.g., time to the first referral after the initial referral), the primary interest is in the distribution of time rather than the mean. Cox proportional hazard models were used for the analysis of time-to-event data. This model can handle cases for which the time is censored (i.e., unknown, but known to be greater than some value).

For the administrative data, stepwise regression[19] was used to identify statistically significant predictors (both main effect and interactions) from the following groupings:

- *Demographic characteristics.* As a result of random assignment, the demographic characteristics of the FAR and IR families were expected to be similar such that any remaining small differences are unlikely to affect the results. Occasionally, the demographic differences are larger than expected. Inclusion of demographic factors in the model removes the effect of any remaining differences on the assessment of FAR versus IR.
- *Variables that affect the outcome.* Some variables, even those balanced between the FAR and IR cases, may predict the outcome. For example, family satisfaction may be related to race. Including race in the model provides additional information about the effect of race on the outcome even if it has little effect on the assessment of FAR versus IR.
- *Interactions of FAR/IR and other predictors.* Interactions were considered as candidate variables only if the factors in the interaction were already in the model as

[19] In this context, "stepwise" regression refers to alternating forward inclusion and backwards elimination steps until no further changes are made. Inclusion or elimination is based on the p-value. Although SAS does not currently provide a stepwise variable selection procedure for weighted survey data, the stepwise selection for the administrative data can be closely approximated using the SAS LOGISTIC procedure specifying the survey weights for the administrative data. As a result, the LOGISTIC procedure was used to select predictors for categorical outcomes using the administrative data. The final model from the LOGISTIC procedure was fit using the SAS SURVEYLOGISTIC procedure to calculate the final p-values. Similarly, the SAS PHREG procedure was used to select predictors for the Cox Proportional Hazards model predicting time-to-event data (survival analysis) and the SURVEYPHREG procedure was used to calculate the final *p*-values.

main effects. Up to two-way interactions were included in the model. Significant interactions of the FAR/IR indicator and other predictors (e.g., race) indicate that the FAR/IR difference depends on the other predictors. Such interactions provide a more detailed assessment of the FAR/IR differences. Interactions were only entered into the administrative data models because of the insufficient number of observations for the family exit survey and case-specific report and the complexity of using the stepwise unweighted algorithm for the survey data.

2.1.9. Study Logistics

The evaluation team obtained approval from the CSU Institutional Review Board (IRB) in November 2010. The CSU IRB served as the IRB of record for the evaluation and conducted the annual reviews after each year of the study. The evaluation team completed all necessary data sharing agreements for the cross-site and local evaluations. The project also formed multiple workgroups very early in the process to determine the practices, policies, and philosophical underpinnings of DR as it would be practiced in Colorado. Appendix K details the logistics of the evaluation including IRB approval, data sharing and evaluation agreements, and workgroup support for the design and implementation of the study.

2.2. Process Evaluation

This section describes the research design, research questions, sampling, data collection, and data analysis for the process evaluation.

2.2.1. Research Design

The evaluation team used a qualitative research approach with a grounded theory design featuring focus groups, structured interviews, and document review.

2.2.2. Research Questions

The process evaluation addressed one fundamental question:

1. What are the perceptions of caseworkers, supervisors, and community stakeholders in regard to caseworker satisfaction and community buy-in related to the implementation of DR in the five counties?

2.2.3. Sampling

The process evaluation employed a non-random sampling design that was purposeful and convenient. Each participating county in the CCDR designated a practice lead to assume responsibility for focus group recruitment with direction from the lead evaluator. The county practice leads distributed recruitment emails to the identified constituent groups, including caseworkers, supervisors, and community stakeholders. The stakeholders group included members of child protection teams (CPT), RED teams, school district staff, law enforcement personnel, service providers, and judicial representatives.

2.2.4. Qualitative Data Collection

The evaluation team used focus groups and interviews to conduct the process evaluation. The QIC-DR developed the focus group and interview protocols with input from the local evaluation teams for the three R&D sites. The evaluation team conducted the focus groups and interviews. Representatives from WRMA and American Humane Association/Kempe Center assisted in taking notes during the focus groups and interviews. A digital audio recorder was used to record the focus groups and interviews.

The evaluation team conducted two rounds of site visits to the five participating counties to collect qualitative data through focus groups with caseworkers, supervisors, administrators, and community stakeholders. The first site visit, which focused on the start-up of the program and early implementation issues, occurred during April and May of 2011, which was six months into implementation. The evaluation team conducted 31 focus groups during the Year 1 site visits: five each in Fremont and Arapahoe counties, six in Larimer County, seven in Garfield County, and eight in Jefferson County. There were a total of 10 caseworker focus groups, eight stakeholder focus groups, six supervisor focus groups, five administrator focus groups, and two screener interviews. Each focus group had between four and 10 participants with a median of six. The focus groups lasted between 60 and 90 minutes and were conducted over a two-day period at each site.

The second round of site visits, which focused on the continued implementation of DR and the implications for sustainability and replication of the child welfare reform effort occurred during July and August of 2012, which was 21 months into implementation, 14 months after the first round of site visits was completed, and four months after the random assignment of cases ended. The evaluation team conducted 16 focus groups during the Year 2 site visits. Three focus groups were held in each county in addition to one interview with state personnel, totaling five caseworker, five supervisor, and five stakeholder focus groups and one project director interview. Each focus group had between four and ten participants with a

median of five. The focus groups lasted between 60 and 90 minutes and were conducted during one day at each site. The evaluation team also reviewed documents for the site-visit reports and informally interviewed the project director (PD) to gain more insight into the pre-, initial, and ongoing implementation policies and practices.

2.2.5. Qualitative Data Analysis

A professional transcriber hired by the evaluation team transcribed the audio files from the focus groups and structured interviews. The evaluation team employed a constant comparative analysis approach to analyze the qualitative data through open, axial, and selective coding, which yielded narratives for each group that participated in the site visits. The evaluation team then combined findings across groups for the evaluation report.

3. Intervention

This chapter defines the differential response intervention examined by the program evaluation, including the core elements, dual track response system, organizational processes, social work practices, and practice principles of the Colorado DR model. Intervention components such as case flow, assessment, services, assessment closure, and re-referrals also are detailed.

3.1. Colorado DR Model

For several months prior to hiring the project director, groups of practitioners in Colorado gathered in workgroups to discuss the values and principles of DR. Using tools gathered from Olmsted County, Minnesota, the group established a foundational document to provide vision for the work. As the workgroups added both practice and Colorado-specific rule and statute components over subsequent months, the Colorado DR Model took shape. The model was designed to facilitate consistent implementation, serve as an outline for practice coaching, and define process measures for fidelity. Little has changed across the different iterations of the model, even as the workgroups enhanced and clarified the original model to serve as an example for counties interested in implementing DR and a resource for new caseworkers during their training experience.

The Colorado DR Model represents an organizational shift in child welfare agencies in Colorado that impacts all parts of the organization. It includes a series of infrastructure changes within the organizations as well as a deepened and enhanced social work practice set for all staff. Infrastructure refers to the daily processes within the organization (such as supervision and family meetings) that need to be aligned and congruent with new practices in order for any real uptake of organizational change to occur. These may include changes to policy, procedure, and practice.

3.1.1. Eight Core Elements

As displayed in Table 3.1, the Colorado DR model is an integration of the eight core elements defined by the QIC-DR, practice principles that guide decision points, promising practice components, and procedures as defined in Colorado law. Some statutory adaptation was necessary, including HB1226, which allowed for two discrete tracks of intervention with no determination of maltreatment in the FAR track.

Table 3.1

Eight Core Elements of Colorado DR Model

Core Element (QIC-DR)	Colorado Implementation
Two or more discrete tracks of intervention.	A new, non-investigatory track was developed: Family Assessment Response.
Multiple responses for reports of maltreatment that are screened in and accepted for response.	Track assignment decision is made at the time of screening.
No substantiation of alleged maltreatment for families served in a non-investigation track and services offered without a formal determination of child maltreatment. This means that perpetrators and victims are not identified for the alleged reports of maltreatment that receive a non-investigation response.	Though the roles of alleged person responsible for abuse or neglect (PRAN) and alleged victim will continue to be entered in Colorado Trails prior to track assignment for the purpose of National Child Abuse and Neglect Data System (NCANDS) reporting, there will be no finding made or determined PRAN entered for those cases served in the FAR track.
Differential use of central registry depending on track, meaning the name of the alleged perpetrator is not entered into the central registry for those individuals who are served through a non-investigation track.	Colorado does not utilize a central registry, but rather records background information in Trails. For cases served in the FAR track, that information cannot be used for the purpose of determining employment eligibility, as investigation will not be conducted.
Response pathways are formally established via legislation, policy or protocols.	HB1226 established the FAR track as an option in all five counties in the project.
The CPS agency determines pathway assignment based on an array of factors (e.g., alleged maltreatment type; presence of imminent danger; risk level; number of prior reports; age of child).	Colorado has developed an agency response guide that provides structured analysis of an array of factors.
Initial pathway assignment can change if the agency obtains new information that alters risk level or safety concerns.	A track change from FAR to IR occurs if any of the following is present within the assessment period: A. The family requests investigation response. B. New information is received that might warrant a change in response. C. There is insufficient engagement or ability to conduct an assessment of child safety.
Services are voluntary – families may accept or refuse services so long as there are no safety concerns.	If no safety factors are present in the safety assessment, the family may accept or refuse services from the agency without consequence.

3.1.2. Dual-track Response System

Colorado implemented a dual-track response system with an investigation response (now called high risk assessment) track and a family assessment response track. The IR track involved a county agency response to higher-risk screened-in allegations of abuse or neglect. Investigative responses involved fact-finding to determine a preponderance of the evidence as to whether or not child maltreatment occurred as alleged. As the non-investigative track of the dual-track response system, FAR did not require or allow the worker to determine a preponderance of the evidence as to whether or not child maltreatment occurred as alleged. Only cases of lower and moderate risk screened-in allegations of abuse or neglect were eligible for the FAR track.

For both the FAR and IR tracks, the services provided include engagement strategies that assist in the assessment of safety, risk, family needs, and family strengths. However, there were some distinct differences in how caseworkers approached the families in each track. One of the main differences occurred at the initial contact with families following a screened-in referral of child maltreatment. Since they were not required by state law (Code of Colorado Regulations, Volume 7) to initially see the child outside the presence of the alleged PRAN, FAR caseworkers were less likely to use the element of surprise to see children within response times set forth by the RED teams. They also were allowed to initially interview the family as a whole. This approach was perceived as less stressful on a family than having to interview each child alone. However, if it appeared that a child was fearful of talking in front of their family, the option remained to talk to the child separately (as is normal in the IR approach). FAR caseworkers typically called families to alert them that a referral had been received and assigned. Caseworkers then asked families how best to complete an initial visit while ensuring that response times for child contact were met.

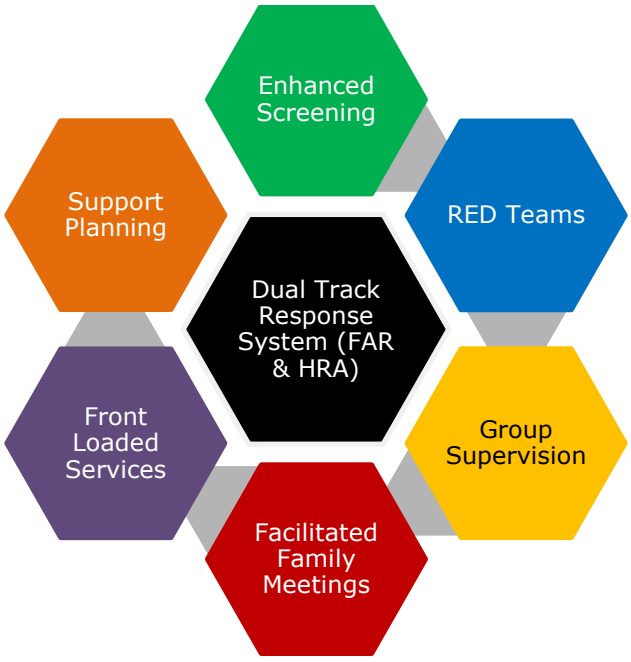
Unlike an IR case, in which collaterals (friends, neighbors, relatives, and other interested parties who may be supports for the family) were typically contacted separately, a FAR case allowed for collaterals to participate in the initial contact or any other scheduled meetings. During the first meeting, FAR workers gave families a brochure explaining the FAR approach. Additionally, workers used a script to describe the project's evaluation elements to families. Initial meetings were typically longer in FAR cases to allow caseworkers sufficient time in obtaining the family history. An additional change was the shift in the workday schedule. Rather than having appointments during the day or seeing children at school, FAR appointments also were scheduled in the late afternoon or evening.

3.1.3. Organizational Processes

As displayed in Figure 3.1, the organizational processes of the Colorado DR Model are: enhanced screening, RED teams, group supervision, facilitated family meetings, group supervision, front-loaded services, and support planning.

Figure 3.1

Colorado DR Model – Organizational Processes



Enhanced Screening – Given the need to make two substantial decisions at the point of referral, a common format for approaching reporting parties clarifies information about the allegation of abuse or neglect as well as obtaining information about family supports and strengths from the reporter's perspective.

RED Teams – Prior to the formation of the CCDR, the customary method for decision-making at the point of referral was a review by one supervisor. One county had experimented with group decision-making at this point, and the other counties became interested in pursuing this strategy. To that end, the CCDR adopted the RED team strategy, which is an innovative decision-making approach to DR track assignments.[20]

[20] Sawyer, R., & Lohrbach, S. (2005). Differential response in child protection: Selecting a pathway. *Protecting Children, 20* (2/3), 44-53.

Group Supervision – All five counties implemented group supervision in some form. The rationale for group supervision is that implementation of a true philosophical shift might best take place in groups of workers and supervisors meeting to discuss decision points and intervention plans. This also flows well with the integration of solution-focused skills.

Facilitated Family Meetings – Prior to the start of the project, all five counties had structures in place for facilitated family meetings which included Family Group Decision Making (FGDM), Team Decision Making (TDM), Family Unity Meetings (FUM), Family Group Conferencing (FGC), Listening to the Needs of Kids (LINKS), and other county-specialized processes. These facilitated family meetings are used to promote family engagement for cases at all levels of risk. Additionally, facilitated family meetings are meant to assist with safety and support planning, which includes formal and informal safety networks for families.

Front-Loaded Services – This refers to the ability of workers to access effective and specialized formal Core Services for families earlier then in the past. Typically, a family is unable to access major agency services until an investigation has been completed. In the Colorado DR model, service referrals are made as soon as family needs are identified and the family agrees to the service. This prevents unnecessary wait times and allows families to begin a change process as soon as possible.

Support Planning – Most child welfare jurisdictions make two kinds of plans with the families they work with: A short-term "safety plan" that typically describes how a child can safety remain in the community over a small time horizon and a "service plan" that typically describes the role formal service providers will play with the family. The Colorado DR model includes a third kind of kind of planning – a support plan. A support plan allows for the creation of agreements between the family, their informal support network and the agency. The time horizon for these plans is typically longer and the plans themselves include detailed, behavioral action-steps the family, network and agency agree to take in order to ensure safety, permanency and well-being are demonstrated over time. Support plans could begin as soon as a case is opened with the agency and could be a critical part of after-care planning.

3.1.4. Social Work Practices

As displayed in Figure 3.2, the social work practices of the Colorado DR Model are: rigorous and balanced assessment, strategies for including children, Consultation and Information Sharing Network, evidence-based assessment tools, risk and goal statements, participation of extended networks, and behaviorally-based safety and support plans.

Figure 3.2

Colorado DR Model – Social Work Practices

Rigorous and Balanced Assessment – In the Colorado DR model, significant attention is paid to interviewing strategies with children, parents, and support networks that surface not only the history of any harm that occurred to the child, but also protective actions parents and networks took to keep their children safe. Both of these approaches are enhanced through the use of solution-focused interviewing (SFI). Originating at the Milwaukee Brief Treatment Center,[21,22,23] SFI is a questioning approach or interviewing practice based on a simple idea with profound ramifications—that the areas people pay attention to will grow. It highlights strategies for child welfare professionals to inquire about acts of protection in as rigorous a way as with acts of risk, and provides a series of micro-practices (e.g., exception questions, relationship questions) to assist with this approach.

Strategies for Including Children – While children are the focus of any child welfare intervention, and most professionals agree that children's perspectives are vital for child welfare work, consistently maintaining this focus can be a daunting task even for seasoned professionals. The temptation to make the work with children a superficial part of child welfare

[21] de Shazer, S. (1985). *Keys to Solution in Brief Therapy*. New York: Norton.

[22] Berg, I. K. (1994). *Family Based Services: A Solution-Focused Approach*. New York: Norton.

[23] Berg, I. K., & Kelly, S. (2000). *Building Solutions in Child Protective Services*. New York: Norton.

casework is great. To help children meaningfully contribute to both assessment and case planning in developmentally-appropriate ways, the Colorado DR model incorporates a series of practices and conversation aids such as the Three Houses[24] and The Safety House.[25]

The Consultation and Information Sharing Framework – This framework assists in analyzing and organizing all the information known about a child and family at any given time into key domains. [26] This process is typically used by child welfare workers and supervisors in case consultations and multi-disciplinary teams, but is designed to be inclusive of the family perspective. It makes use of common language to help sort and prioritize case information, allowing increased clarity about the hopes, concerns, and purpose of any particular child welfare intervention.

Evidence-based Assessment Tools – Regular key decisions (i.e., opening and closing cases, bringing a child into care, what should go in a case plan) need to be considered in almost every child welfare case. Research into child welfare decision-making[27] indicates key decisions are frequently made in an inconsistent fashion using inconsistent criteria. Evidence-based decision support tools use the best of child welfare research and aggregate data to help caseworkers "check" their intuition at these key decision points, ensuring these immensely important decisions are consistent and congruent with both research and organizational policy.

Risk Statements – Once a good analysis by caseworkers and supervisors is completed with the Framework, it becomes possible for the agency to articulate detailed, short, and behaviorally-based risk statements[23] employing clear, non-judgmental language to relay:

- What the agency is concerned may happen to the child if nothing else changes; and
- When the agency is most worried this may happen.

Such statements provide a clear rationale for child welfare involvement and are a foundation for making clear goals about the work. These deceptively simple statements take some time to construct; but once they are made, they can be shared with family members, community partners, court officials, and anyone else working with the children and family.

---

[24]Weld, N. (2008). The three houses tool: Building safety and positive change. In M. Calder (Ed.), *Contemporary risk assessment in safeguarding children.* Lyme Regis, England: Russell House Publishing.

[25] Parker, S. (2009). *The safety house: A tool for including children in safety planning.* Perth, Australia: Aspirations Consultancy.

[26] Lohrbach, S. (2000). *Child protection practice framework – Consultation and information sharing.* Unpublished manuscript.

[27] Children's Research Center. (2008). *Structured decision making: An evidence-based approach to human services*. Retrieved from http://nccdglobal.org/sites/default/files/publication_pdf/2008_sdm_book.pdf

Goal Statements – Goals in child welfare are often service-driven rather than safety-driven and behaviorally-based. Everyone working with a family in an open child welfare case should understand and be able to articulate the necessary steps for adequate protection demonstration and case closure. These goals should:

- Address the risk statement;
- Be created collaboratively with the family members whenever possible;
- Be written in clear, everyday language; and
- Describe the presence of new, observable behaviors or actions (particularly behaviors with the children) rather than simply the absence of old, problematic behavior.

Participation of Enhanced Networks – The axiom that "it takes a village to raise a child" is never truer than in child welfare work when caregivers are found to be a danger to their children. Drawing on the wisdom of family teaming approaches, the Colorado DR Model offers strategies for identifying and enlarging the network of people around the child and communicating the risk statements to them. The DR model also allows workers to enlist the network's help in developing and implementing plans to keep children safe and address the reasons for referral to the system.

Behaviorally-based Safety and Support plans – Safety and support planning must be more than a simple "laundry list" of services to which a family has agreed. The Colorado DR Model treats services and safety as separate but equal concerns. In this approach, services are a means to an end—that end being actual safety for the child. All planning is collaboratively made and includes detailed actions agreed to by parents and extended family members in order to show everyone involved that children will be safe.

3.1.5. Practice Principles

These practice principles were adapted from a model utilized in Olmsted County, Minnesota.[28] They were further defined during pre-implementation by the project's intake/services workgroup.

Focus on Safety – Achieved through family engagement, child safety is the primary focus. It is important to understand the danger/harm from the points of view of the caseworker and family members. The safety concerns of the agency are clearly articulated. Caseworkers develop safety-focused partnerships with families and conduct a comprehensive assessment of the family's strengths, resources, and supports to help address the identified safety and risk

[28]Sawyer, R. (2010, April). *Introduction to differential response in child protective services*. Community Presentation, Fort Collins, CO.

concerns. Caseworkers assess safety by using scaling questions with the family. Effort is made to identify exceptions that could be interpreted as strengths and built upon in the plan. Natural support systems committed to supporting the family over time are included to help bring about necessary change to ensure children's well-being. Safety plans give explicit and careful focus to the goals of the family and agency that ensure child safety.

Constructive Engagement (Partnership with Families) – Constructive engagement occurs when the caseworker practices out of a belief that every family has strengths, resources, their own way of solving problems, and their own goals. Family members are seen as future cooperative partners. To facilitate this, agency staff use solution-focused skill sets. Serving the unique needs of various families also demands creative and at times non-traditional service delivery strategies.

Collaborative Engagement (Collaboration with Communities) – Community agencies and informal family supports are valued as partners in promotion and maintenance of child welfare. Caseworkers recognize the value of collaborating with other natural supports and community agencies to help the family, engaging members of these support systems in the same way they engage the family.

Family & Community Inclusion – Caseworkers are open-minded about family and community members identified as partners in building safety. Families are encouraged to look at extended family and community supports as helpers in problem-solving from the very beginning of the case. Whenever possible, family meetings are called to create and expand networks of family support. These meetings facilitate constructive and collaborative engagement in an integrated fashion. Efforts are made to engage fathers and the paternal side of the family in addition to mothers and maternal supports. Caseworkers are familiar with community resources, particularly those that focused on meeting basic family needs. Before case closure, caseworkers develop support plans with families to assist in maintaining networks. Services are sustainable post-departmental involvement if needed.

Assessment of Risk and Protective Capacity – Structured decision-making with the Risk Assessment tool balances with the intentional inclusion and integration of a family's perspective. This represents a bridging of professional and family knowledge. A consistent framework (included in Trails) is used throughout agency involvement for assessment, service planning, supervision, and family meetings, providing the opportunity for delineation of needs, strengths, and next steps.[29] It also provides a common point of reference for group

[29]Turnell, A., & Edwards, S. (1999). *Signs of Safety: A solution and safety oriented approach to child protection casework.* New York: Norton.

consultation/supervision.[30] Workers develop a treatment plan with the family that is concise, behaviorally specific, individualized, and measurable using the family's language. The focus remains on safety instead of services, compliance, and completion as the main criteria for closing the case (i.e., is there enough safety to close?).

Transparency – Transparency in child welfare represents clear communication of roles, responsibilities, and agency authority. There are multiple levels of authority implicit in the child welfare system. Efforts are made at all levels to utilize authority in a responsible manner with respect for individual needs, differences, culture, and diversity. Transparency is demonstrated by explaining to families how decisions were made or sharing discussions that occur when they were not present. There is a general practice of "no decision about a family without the family." Caseworkers recognize the power of language, foregoing "professional jargon" in favor of the family's language, empathic and active listening, and checking for understanding by the family. When possible, agency authority is used to advocate for the needs of families. Additionally, there is transparency of practice within the agency, and agency staff is engaged in group decision-making, including RED teams.[20] Transparency of practice between workers is a main focus for group supervision and consultation.[30]

3.2. DR Fidelity Assessment Matrix

The fidelity assessment matrix was designed to operationalize the practices and practice principles of the Colorado DR model to allow for an assessment of implementation fidelity. It was developed through conversations with project workgroups, consultation with various venues practicing DR, guidance from state program staff, and direction from the evaluation team. The implementation of fidelity assessment did not occur as planned because of resource and time limitations, but the matrix could serve as a valuable tool for the ongoing evaluation of DR in Colorado if the practice expands statewide. The matrix (see Appendix L) displays the interplay between practices, practice principles, fidelity indicators, and data sources.

3.3. Intervention Components

This section describes intervention components of differential response in Colorado such as case flow, assessment, services, assessment closure, and re-referrals.

---

[30]Lohrbach, S. (2007). Supervision in child welfare practice. *Social Work Now: The Practice Journal of Child, Youth, and Family*, *40*, 19-24.

3.3.1. Case Flow

The evaluation team worked with representatives from each county to develop case flow charts to visually depict the DR practice model. As displayed in Figure 3.3, the case flow provides an example of the initial screening and eligibility determination for the DR pilot. A referral went through three decision points before being deemed as FAR eligible.

Initial Screening – All referrals went through an enhanced screening process using a comprehensive referral Screening Guide to gather as much explicit information as possible about the allegation. This information assisted in the determination of whether the referral needed an immediate response or could be passed to a RED team. The screener also asked questions about family support and protective factors to provide a comprehensive view of the family at the time of the referral. If the referral warranted an immediate response, the screener or the screener's supervisor assigned the referral to IR. If not, the screener sent the referral to the RED team, which consists of 2-5 caseworkers, one or more supervisors, and other staff.

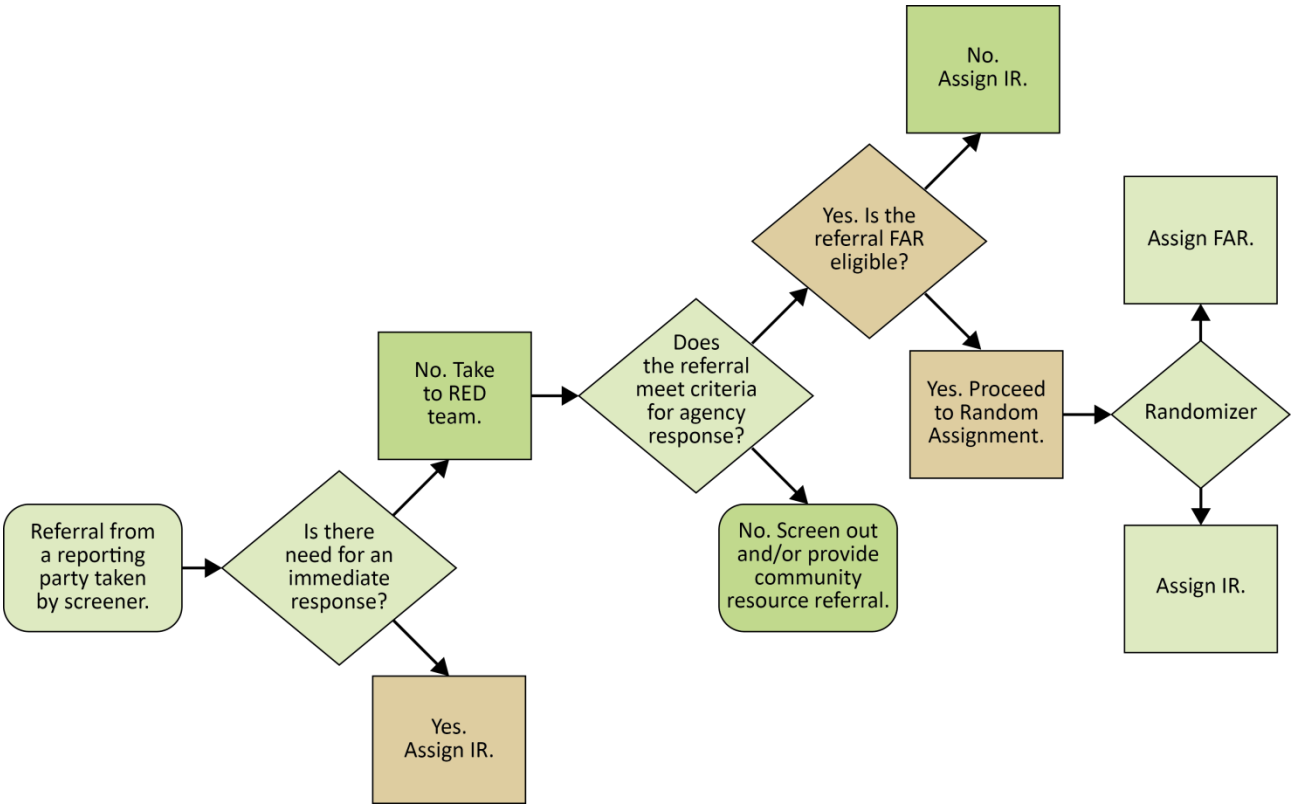
Criteria for Agency Response – Using the Agency Response Guide (see Appendix A), the RED team examined the screening information to determine whether the referral met the criteria for agency response. If the referral did not require an agency response, the RED team screened it out at this stage and may have provided community resource referrals to the family. If the referral met the agency response criteria, the RED team determined its response time and track assignment eligibility.

Track Assignment – The RED team also used the Agency Response Guide to determine whether a referral was eligible for FAR based on a balanced assessment of safety and protective capacities. They followed a process of visually outlining each referral to compare the family strengths, history, danger/harm, and complicating and risk factors. If the referral met the criteria for a required or discretionary investigation as found in the Agency Response Guide, the RED team assigned it to the IR track. If it was eligible for FAR, it proceeded to randomization.

Re-track Assignment – The caseworker had the discretion to re-track a referral randomly assigned to the FAR track to the IR track if additional and/or more serious safety concerns were uncovered or revealed during the assessment. This occurred in 3% of the FAR assessments.

Figure 3.3

Generic Case Flow Chart for Colorado



3.3.2. Assessment

Colorado employed the Colorado Assessment Continuum (CAC) in both the IR and FAR tracks, which involves conducting safety and risk assessments at or before 30 days from the date of assignment. The CAC was implemented prior to differential response to guide decisions for working with families. The CAC assessments, or tools, are the same in both tracks and were administered at the same points in involvement with the family (prior to 30 days from the date of assignment). The Colorado Safety and Risk Assessments were designed to be completed by the caseworker based on conversations with the family and collateral sources. The risk assessment includes factors such as the number of children, age of children, and prior referrals.

If safety factors applied, the caseworker proceeded to an assessment of protective capacity and an in-home safety analysis. If the caseworker determined that a protective parent mitigated child safety in lieu of a safety factor, there was no need to proceed. If needed and possible, safety planning attempted to keep the children safely in their home. If safety was not attainable through this means, the child was removed from the home and placed in out-of-

home care. Caseworkers completed an additional safety assessment anytime there was a significant change in family circumstance or if changing situations posed a new or renewed threat to child safety. Additionally, caseworkers completed a new safety assessment prior to supervisory approval for case closure on all FAR and IR cases opened over 60 days.[31]

3.3.3 Services

Post-assessment services differed between FAR and IR cases. Service delivery in IR cases was designated by a service plan or a service authorization in Trails. For FAR cases, caseworkers administered services through a new document called the Family Assessment Response Service Plan (FARSP). Developed using input from Parent Partners in Jefferson County, this tool was designed to be open-ended and family-friendly. Under Volume 7, services were required to formally begin by 60 days from the date of referral unless the FAR was closed in the assessment phase.

3.3.4. Assessment Closure

When safety remained a concern, the caseworker could not end involvement with the family. However, in the absence of safety concerns warranting a safety plan or removal, families receiving FAR had the choice of staying involved with the agency by participating in services. When there was no need for services or if the family declined further involvement, the FAR closed at the assessment phase. FAR assessments closed when the caseworker determined in partnership with the family that the agency and family goals were met and/or the family chose not to complete services. However, in FAR assessments where families needed or volunteered for services, the agency engaged in post-assessment service provision.

In 3% of FAR assessments, there was a need to take steps for child safety by removing the child from home and placing them in out-of-home care. In 4% of the FAR assessments, steps needed for child safety required family participation in services through court involvement. Based on data from the cross-site evaluation, there was a statistically-significant difference ($p < .001$) between the length of time FAR and IR assessments remained open. Specifically, the mean length of a FAR assessment was 43 days as compared with 30 days for an IR assessment. During the second year of implementation, increasing attention was paid to the timely closure of assessments, both FAR and IR.

[31]At the time of the evaluation, the CAC tools also included the North Carolina Family Assessment Scale (NCFAS), which caseworkers completed at 60 days from the date of assignment or at case opening, whichever came first, for both FAR and IR cases. The NCFAS was mandatory at case closure. Until it was phased out of the CAC in late 2012, the NCFAS was the main assessment tool used to determine family service needs and measure family status and progress on five domains.

3.3.5. Re-referrals

A re-referral is when a new referral meeting agency response criteria comes in on a family after their initial involvement period is closed. Once it was determined that a screened-in re-referral had been received on a previously randomized family, it was passed on to the RED team. If a re-referral was received on a family randomized into the FAR track, it remained in the FAR track unless the RED team determined that the new report required a HRA response (and thus was not FAR-eligible). Similarly, if a re-referral was received on a family that was randomized into the IR track, it stayed in the IR track. However, a reminder in Trails for this requirement was not functional for a short period of time, and a very small number of originally-assigned IR cases ended up being assigned to FAR after a screened-in re-referral.

4. Client, Caseworker, and Case Characteristics

This chapter presents a descriptive analysis of the client, caseworker, and case characteristics for the outcome evaluation. The client characteristics are based on weighted administrative data collected from Trails on all families randomly assigned to the study (see the CONSORT table in Appendix B for a description of this sample). The caseworker characteristics are based on survey data collected from the caseworker general survey. The case characteristics are based on weighted administrative data and family exit survey and case-specific report data that reflect services, caseworker contacts, and "during initial involvement" events (i.e., D&N filing, case opening, OOH placement) for families in both tracks.

4.1. Client Characteristics

The client characteristics are described for referral findings, family demographics, and presenting issues.

4.1.1. Referral Findings

Among families assigned to IR, 56% of the initial referrals were unfounded, 22% were inconclusive, 21% were founded, 1% were served in the FAR track (due to data entry error in Trails) and had no finding, and in less than 1% of IR cases, no abuse or neglect investigation occurred (typically due to inability to locate the family after assignment). Among the 103 FAR families who experienced a track change and were served in the IR track, 40% of the referrals were unfounded, 35% were founded, 22% were inconclusive, and in 3% of FAR cases, no abuse or neglect investigation occurred.

4.1.2. Family Demographics

The evaluation team examined family demographics and presenting issues using administrative data to confirm the equivalence of the treatment and control groups.[32] As displayed in Table 4.1, no statistically significant difference was found between FAR and IR families for ethnicity of the youngest child. To capture ethnicity of the families, the team used the ethnicity of the youngest child in the home.[33] For almost half of the families in the sample, the youngest child in the home was Caucasian (46%), followed by Hispanic (24%), and African-American (11%). For about 19% of the study families, the ethnicity of the youngest child was not entered in Trails.

[32] Given the number of comparisons (i.e., five for demographics and 14 for presenting issues), the team used a Bonferroni adjustment to set a more conservative *p*-value (.05/19 statistical tests = .003) to assess statistical significance.

[33] Ethnicity of the youngest child was determined by the following hierarchy: Hispanic - African American – Other – Caucasian. The "other" ethnicity category consisted of American Indian, Asian, Alaska Native, and Native Hawaiian/Other Pacific Islander.

Table 4.1

Family Demographics of Study Sample

Demographic	Response Categories	FAR Weighted *N* (%)	IR Weighted *N* (%)	Total Unweighted[34] *N* (%)
Ethnicity of Youngest Child	Caucasian	1,447 (45%)	799 (44%)	2,276 (46%)
	Hispanic	758 (24%)	423 (23%)	1,196 (24%)
	African-American	316 (10%)	225 (12%)	534 (11%)
	Other	37 (1%)	18 (1%)	59 (1%)
	Unknown	636 (20%)	337 (19%)	931 (19%)
	Total	3,194 (100%)	1,802 (100%)	4,996 (100%)

Note: Percentages may not add up to 100% due to rounding.

As displayed in Table 4.2, there was no statistically significant difference between FAR and IR families on caregiver age. The age of the caregivers ranged from 15 to 82 with an overall average of 33 years of age. There were three statistically significant differences in the family demographics between FAR and IR families including the number of children in the home, the number of caregivers in the home, and the age of the youngest child in the home. The number of children in the home ranged from 1-9, with IR families (M = 2.0) having more children in the home, on average, than did FAR families (M = 1.8), t = 3.66, p = .0003. The age of the youngest child in the home ranged from 0-18 with FAR families (M = 5.9 years) having an "older" youngest child, on average, than did IR families (M = 5.4 years), t = 3.69, p = .0002. The number of caregivers in the home ranged from 1-2 with IR families (M = 1.7) having more caregivers, on average, than did FAR families (M = 1.6), t = 4.13, p = .0001. The significant differences between the groups on these family demographics were likely driven by the large sample size, and do not appear to be practically significant or related to track assignment.

[34] The FAR and IR samples were not directly comparable due to the differing assignment probabilities among counties and over time. However, the analysis weights adjust for the differing assignment probabilities. As a result the weighted FAR and IR estimates can be compared. Because the weights are designed to weight the FAR and IR cases up to all cases eligible for the study, the weighted FAR (or IR) estimates approximate an expected response if all cases had been assigned to FAR (or IR). As a result, for characteristics known for all cases from the administrative data, the weighted FAR and IR distributions can be compared to the unweighted distribution for all cases. For data that would not be affected by the assignment, the weighted FAR and IR values can be compared to assess the performance of the random assignment and weighting procedures. For characteristics that might be affected by the FAR/IR assignment, the difference between the weighted FAR and IR results is a measure of the effect of the FAR versus IR assignment.

Table 4.2

Family Demographics of Study Sample

Demographic	FAR Weighted *M* (*SD*)	IR Weighted *M (SD)*	Total Unweighted[34] *M (SD)*
Caregiver Age (*N* = 4,819)	33.35 (9.45)	33.47 (9.66)	33.39 (9.48)
Youngest Child Age (*N* = 4,996)*	5.92 (4.87)	5.36 (4.70)	5.75 (4.82)
Number of Children in the Home (*N* = 4,996)*	1.84 (1.04)	1.97 (1.11)	1.88 (1.07)
Number of Caregivers in the Home (*N* = 4,996)*	1.60 (.49)	1.67 (.47)	1.62 (.48)

* $p < .003$

4.1.3. Presenting Issues

As displayed in Table 4.3, there were no statistically significant differences in presenting issues between the two groups. This includes overall risk level (as measured by the Colorado Risk Assessment tool), specific risk factors (i.e., domestic violence, mental health, and substance abuse), prior involvement (i.e., referrals and assessments), and allegation type (i.e., abuse and neglect).[35] Almost all study families (90%) were assessed to be at low or medium risk for future maltreatment. A little over one-third of the study caregivers were assessed as having a domestic violence risk factor (36%), while less than a quarter were assessed as having a mental health risk factor(22%) or substance abuse risk factor (21%).

As displayed in Table 4.3, within 24 months prior to study enrollment, about half of the families (49%) had a referral while more than a quarter (29%) had an assessment (i.e., a screened-in referral). About one-third of the study families (33%) had an allegation of abuse, while almost three-quarters (73%) had an allegation of neglect. Among the allegations of neglect, 69% were for an injurious environment, 9% were for a lack of supervision, 4% were for educational neglect, 3% were for medical neglect, 3% were for a failure to protect, 2% were for deprivation of necessities, and 10% were for other types of neglect.

[35] All of these presenting issues are considered in relation to risk of future child maltreatment.

Table 4.3

Presenting Issues of Study Sample

Presenting Issue	Response Categories	FAR Weighted *N* (%)	IR Weighted *N* (%)	Total Unweighted[34] *N* (%)
Risk Level	Low	418 (14%)	259 (15%)	679 (14%)
	Medium	2,313 (76%)	1,330 (75%)	3,643 (76%)
	High	305 (10%)	173 (10%)	486 (10%)
	Total	3,036 (100%)	1,762 (100%)	4,808 (100%)
Risk – Domestic Violence	Yes	1,143 (36%)	672 (37%)	1,816 (36%)
	No	2,051 (64%)	1,130 (63%)	3,180 (64%)
	Total	3,194 (100%)	1,802 (100%)	4,996 (100%)
Risk – Mental Health	Yes	730 (23%)	373 (21%)	1,113 (22%)
	No	2,464 (77%)	1,429 (79%)	383 (78%)
	Total	3,194 (100%)	1,802 (100%)	4,996 (100%)
Risk – Substance Abuse	Yes	667 (21%)	374 (21%)	1,041 (21%)
	No	2,527 (79%)	1,428 (79%)	3,955 (79%)
	Total	3,194 (100%)	1,802 (100%)	4,996 (100%)
Had Prior Referral in Past 24 months	Yes	1,592 (50%)	856 (48%)	2,456 (49%)
	No	1,602 (50%)	946 (52%)	2,540 (51%)
	Total	3,194 (100%)	1,802 (100%)	4,996 (100%)
Had Prior Assessment in Past 24 Months	Yes	928 (29%)	521 (29%)	1,457 (29%)
	No	2,266 (71%)	1,281 (71%)	3,539 (71%)
	Total	3,194 (100%)	1,802 (100%)	4,996 (100%)
Allegation – Abuse	Yes	1,042 (33%)	571 (32%)	1,626 (33%)
	No	2,152 (67%)	1,231 (68%)	3,370 (67%)
	Total	3,194 (100%)	1,802 (100%)	4,996 (100%)
Allegation – Neglect	Yes	2,346 (73%)	1,343 (75%)	3,668 (73%)
	No	848 (27%)	459 (25%)	1,328 (27%)
	Total	3,194 (100%)	1,802 (100%)	4,996 (100%)

As displayed in Table 4.4, the abuse, neglect, and total risk scores for study families (as measured by the Colorado Risk Assessment tool) were low (indicating lower risk). The abuse risk score ranged from -2 to 17 with an average of 1.7, the neglect risk score ranged from -3 to 17 with an average of 2.0, and the total risk score ranged from -5 to 30 with an average of 3.7. The number of safety concerns (as measured by the Colorado Safety Assessment tool) ranged from 0 to 8 with an average of 0.17 with 88% of study families having no safety concerns (i.e., no threat of imminent danger to the children). The number of prior referrals within 24 months ranged from 0-16 with an average of 1.2, and the number of prior assessments within 24

months ranged from 0-7 with an average of 0.5, both of which indicate that families had limited involvement with the child welfare system prior to enrollment in the study.

Table 4.4

Presenting Issues of Study Sample

Presenting Issue	FAR Weighted *M (SD)*	IR Weighted *M (SD)*	Total Unweighted[34] *M (SD)*
Abuse Risk Score (*N* = 4,824)	1.77 (3.02)	1.62 (2.98)	1.71 (3.01)
Neglect Risk Score (*N* = 4,860)	2.00 (3.04)	1.97 (3.01)	1.98 (3.03)
Risk Total Score (*N* = 4,820)	3.77 (5.36)	3.59 (5.32)	3.70 (5.36)
Number of Safety Concerns (*N* = 4,996)	0.15 (0.58)	0.20 (0.52)	0.17 (0.57)
Number of Prior Referrals within 24 months (*N* = 4,996)	1.14 (1.69)	1.17 (1.77)	1.15 (1.71)
Number of Prior Assessments within 24 Months (*N* = 4,996)	0.44 (0.85)	0.46 (0.88)	0.45 (0.85)

4.2. Caseworker Characteristics

This section describes the characteristics of the caseworkers and supervisors who responded to the caseworker general survey. This was not meant to be representative sample and the findings should not be generalized to all caseworkers and supervisors in Colorado.

4.2.1. Demographics

As displayed in Table 4.5, 40% of respondents to the caseworker general survey were from Jefferson, 24% were from Arapahoe, 24% were from Larimer, 8% were from Garfield, and 5% were from Fremont. For primary responsibility, 75% of respondents identified as caseworkers and 25% identified as supervisors. For gender, 91% of respondents identified as female, while 9% of respondents identified as male.

For race/ethnicity, 91 % of respondents identified as white/Caucasian, 10% identified as Hispanic/Latino/Spanish origin, 4% identified as Asian, 3% identified as Black/African American, 3% identified as American Indian, and 3% of respondents identified as 'other.' For educational level, 48.7% of respondents had a four-year college degree, while 51.3% had completed additional graduate coursework or had a graduate degree.

Table 4.5

Demographics of Caseworker General Survey Respondents (N = 119)

Characteristic	Frequency	Percentage
County		
Arapahoe	29	24.4
Fremont	6	5.0
Garfield	9	7.6
Jefferson	47	39.5
Larimer	28	23.5
Primary Responsibility		
Caseworker	89	74.8
Supervisor	30	25.2
Gender		
Male	11	9.2
Female	108	90.8
Race/Ethnicity		
Hispanic/Latino/Spanish Origin	12	10.1
Black/African-American	4	3.4
Native Hawaiian/Other Pacific Islander	0	0.0
Asian	5	4.2
Alaska Native	0	0.0
White/Caucasian	108	90.8
American Indian	3	2.5
Other	4	3.4
Educational Level		
Four-year College Degree	58	48.7
Some graduate school or graduate degree	61	51.3

Note. Respondents could indicate more than one racial or ethnic category. As a result, the total number of responses is 136; however, percentages were calculated as a proportion of the 119 respondents.

As displayed in Table 4.6, the mean age of respondents was 36 years old, with a median age of 34 years, a minimum age of 23 years, and a maximum age of 64 years. The mean level of child welfare experience was 6.8 years, with a median experience level of 5 years. Respondents ranged from having less than one year of experience to having almost 40 years of experience.

Table 4.6

Age and Years of Experience of Caseworker General Survey Respondents (N = 119)

Characteristic	Mean	Median	Min	Max
Age of Respondent	36	34	23	64
Years of Child Welfare Experience	6.8	5	0	39

4.2.2. Professional Experiences

The following tables summarize the professional experiences of the 119 caseworkers and supervisors who completed the caseworker general survey. Table 4.7 displays areas of responsibility for caseworkers and supervisors at the time they responded to the survey. Respondents could indicate more than one work area, as each area of responsibility could be marked either yes or no on the survey. For general areas of responsibility, 71% of respondents had intake responsibilities, 31% had ongoing responsibilities, and 18% had screening responsibilities. For specific work areas, 56% of respondents conducted assessments for the IR track, 45% conducted assessments for the FAR track, 29% handled out-of-home placement and custody cases, 19% provided family preservation/prevention services, 8% carried adoption cases, and 22% worked in other areas.

Table 4.7

Work Areas of Caseworkers and Supervisors

Outcome	Frequency	Percent
Intake		
Yes	85	71.4
No	34	28.6
Ongoing Casework/Intact Family		
Yes	37	31.1
No	82	68.9
Screening		
Yes	21	17.6
No	98	82.4
Assessment-Investigative Response Track		
Yes	66	55.5
No	53	44.5
Assessment-Alternative Response Track		
Yes	54	45.4
No	65	54.6
Out-of-home Placement/Custody Cases		
Yes	35	29.4
No	84	70.6
Family Preservation/Prevention Services		
Yes	23	19.3
No	96	80.7
Adoption		
Yes	10	8.4
No	109	91.6
Other		
Yes	27	21.7
No	92	77.3

For the impact of FAR on caseloads, workloads, and paperwork, caseworkers completed a scale that ranged from large decrease (1) to a large increase (5) with 3 being no change. As displayed in table 4.8, caseworkers reported that FAR had slightly increased their caseload size, overall workload, and paperwork at the time of survey administration.[36] For caseload, 36% of respondents experienced a small or large increase after FAR, while 51% experienced no change in caseload. For workload, 41% of respondents experienced a small or large increase after FAR, while 48% experienced no change in workload. For paperwork, 46% of respondents experienced a small or large increase after FAR, while 43% experienced no change in paperwork.

Table 4.8

Impact of FAR on Caseloads, Workload, and Paperwork for Caseworkers (N = 89)

Question	Large Decrease	Small Decrease	No Change	Small Increase	Large Increase
Has FAR in any way caused an increase or decrease in your caseload size?	3.4%	10.1%	50.6%	24.7%	11.2%
Has FAR in any way caused an increase or decrease in your overall workload?	2.2%	9.0%	48.3%	28.1%	12.4%
Has FAR in any way caused an increase or decrease in your paperwork?	6.7%	4.5%	42.7%	23.6%	22.5%

4.3. Case Characteristics

The case characteristics are described for services, caseworker contact, and initial involvement events.

4.3.1. Services

The tracks were compared on service provision, availability of services, timeliness of services, and service need.

4.3.1.1. Service provision

The case-specific report was used to answer the question "are there differences in the provision of service referrals and receipt of actual services between families assigned to FAR

[36] These are descriptive results so cannot be generalized beyond the sample of 119 caseworkers and supervisors who responded to the survey.

and IR?" [37] The frequency of providing service referrals and actual services was compared between families assigned to FAR and families assigned to IR using a weighted multiple logistic regression model fit to the data.[38] The difference between DR tracks was statistically significant for provision of referrals, $\chi^2(1) = 36.37$, $p < .0001$, OR = 2.14, CI[1.67, 2.74][39], and for service provision, $\chi^2(1) = 22.62$, $p < .0001$, OR = 1.76, CI[1.39, 2.22].[40] These results indicate that there was a statistically significant difference between the DR tracks in frequency of both referrals provided by child protective services and in frequency of service receipt by families. On average, families assigned to the FAR track had 2.1 times the odds of being provided with a service referral than did families assigned to the IR track. Families assigned to FAR had 1.8 times the odds, on average, of receiving services than did families assigned to the IR track.

The family exit survey was used to answer the question "do families assigned to FAR receive different services and in different quantities than do families assigned to IR?" Three different approaches were used to answer this question.[41] First, differences by track based on receipt of at least one service were examined. Second, differences by track based on the total number of services received were examined. Third, differences by track based on the type of services received were examined.

Any services received: Using logistic regression, the receipt of at least one service variable[42] was included as a dependent variable in the model. The model revealed no statistically significant differences between the tracks on receipt of at least one service as reported by

---

[37] The following survey items were used: 'Was information about or referral to services given to the family' (yes, no, uncertain) and 'were any services (traditional or non-traditional) or supports provided to this family (caregivers or children)' (yes, no, uncertain). Note that responses of 'uncertain' were considered missing. The actual number of families with non-missing responses used for the referral outcomes analysis was 1,651 (99% response rate; 99% for FAR and 98% for IR). For the service provision outcome analysis it was 1,353 (81% response rate; 80% for FAR and 82% for IR).

[38] Both the 'referral provided' and the 'service provided' dependent variables for each model could take a value of 'yes' or 'no'. The logistic regression models were run separately for referrals and services. For both regressions, the independent variables included track, county, number of children in the home, number of caregivers and age of youngest child. County was statistically significant for referrals, $\chi^2(4) = 15.30$ and $p = .0041$, but was not significant for service provision. None of the three demographic variables were statistically significant for either outcome, so they were dropped from the models.

[39] See Table M.1 in Appendix M for details on the logistic regression model.

[40] See Table M.2 in Appendix M for details on the logistic regression model.

[41] For each of the logistic regression models, five covariates were included in the analysis: respondent gender, income, number of children in the home, number of caregivers in the home, and age of the youngest child. Variables were dropped from the model if they were not significant at $p < .05$. Track and county were also included as predictors in the models.

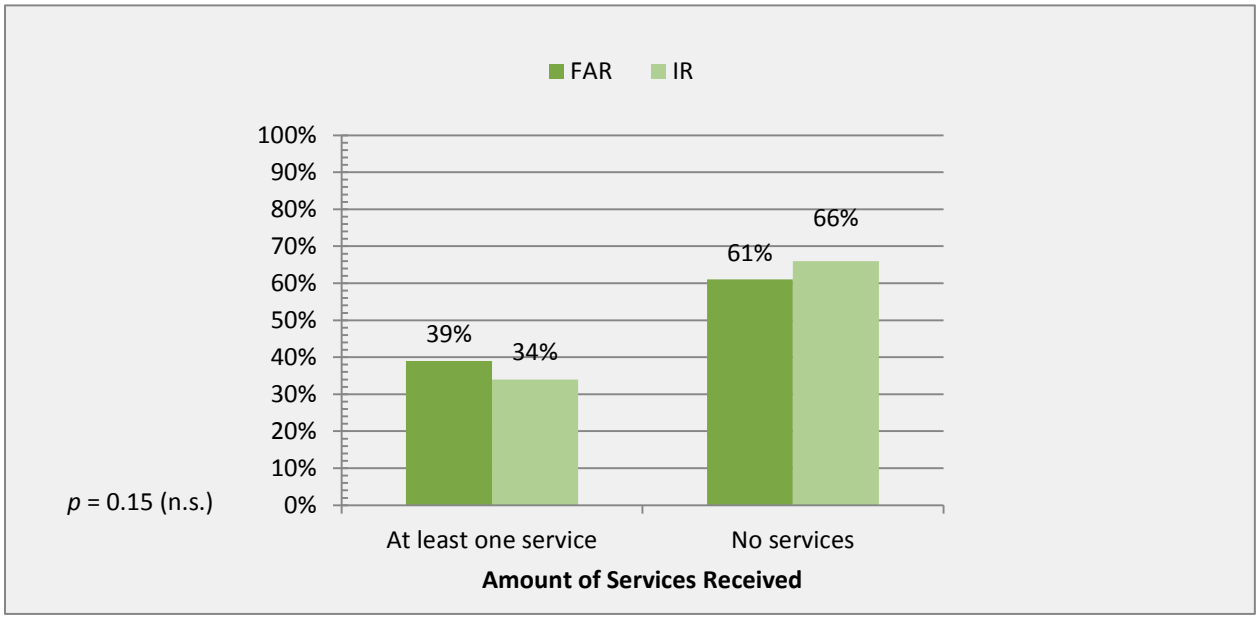
[42] The number of services received variable was recoded to describe the receipt of at least one service.

families, $\chi^2(1) = 2.05$, $p = .152$, OR = 1.36, CI[0.89, 2.08].[43] In total, 37% (n = 1,870) of all respondents reported receiving at least one service. As shown in Figure 4.1, 39% (n = 1,253) of FAR respondents and 34% (n = 617) of IR respondents reported receiving at least one service. It should be noted that 63% of all respondents (61% for FAR; 66% for IR) reported receiving no services. There were no statistically significant differences between the groups on receipt of at least one service or the receipt of no services as reported by families.[44]

Number of services received: To examine differences in track by the total number of services received, the evaluation team created a scaled item from a count of the total services indicated on the survey. The logistic regression model revealed no statistically significant differences between the tracks on the number of services received as reported by families, $\chi^2(1) = 1.54$, $p = .215$.[45] Both FAR and IR respondents received, on average, approximately one service per family (FAR = .89; IR = .76), with FAR respondents reporting a range of 1-15 services compared to 1-8 services among IR respondents.

Figure 4.1

Services Received by Track



[43] See Table N.1 in Appendix N for details on the logistic regression model.

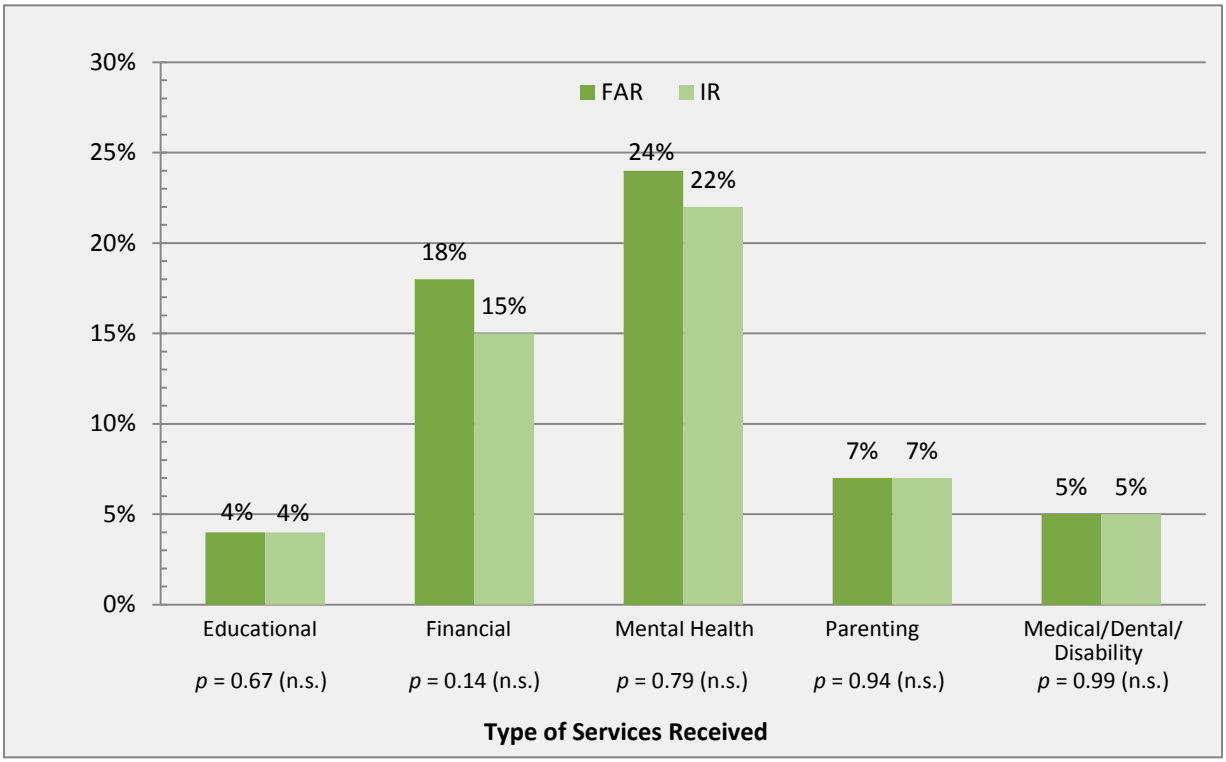
[44] See Table N.2 in Appendix N for details on the logistic regression model.

[45] See Table N.3 in Appendix N for details on the linear regression model.

***Type of services received**:* Family exit survey respondents were asked to identify the types of services they received as part of their experience with CPS by checking them off from a list of services displayed in the survey. The evaluation team condensed this list of 25 items into five major service categories: educational services, financial services, mental health services, parenting services, and medical/dental/disability services.[46] The regression model for each service category revealed no statistically significant differences between the tracks on the receipt of services as reported by families, $\chi^2(1) < 2.19$, $p > .139$.[47] Figure 4.2 shows the weighted percentage of FAR and IR families who received each type of service.

Figure 4.2

Type of Services Received by Track



[46] The number of services received in each category was recoded as receiving "no services" or "one or more services" for the logistic regression analysis.

[47] See Tables N.4-N.8 in Appendix N for details on the logistic regression models for each type of service. All models included track and county, neither of which were significant for any model. For the educational services model, income was statistically significant ($\chi^2(1) = 11.06$, $p = .011$), with use of services decreasing with increasing income. For the financial services model, income was statistically significant ($\chi^2(1) = 31.43$, $p < 0.0001$), with use of services decreasing as income increased. For the mental health services model, the number of children in the home was statistically significant ($p = 0 .037$), with use of services decreasing with more children in the home.

4.3.1.2. Availability of services

The availability of services was discussed by caseworkers, supervisors, and community stakeholders during the focus groups. Although funds were available for FAR and IR families equally, caseworkers reported that FAR caseworkers had access to more financial resources than did IR workers for the following reasons: (1) management was more open to providing financial assistance in FAR assessments, and (2) FAR assessments received help from Medicaid, food stamps, TANF, and housing through a "contact person" working directly with FAR caseworkers. An additional funding difference identified by caseworkers was related to an increased attempt to utilize community-based services before tapping into Core Services. Community stakeholders reported that FAR caseworkers were more creative with resources and had more "leeway" with regard to what services to provide. Some supervisors reported that FAR caseworkers became more educated around community resources. Furthermore, FAR caseworkers were perceived by supervisors to have more time to help connect families directly to community resources.

Table 4.9 summarizes caseworker and supervisor responses from the caseworker general survey to two items about perceptions of service differences between FAR and IR in their own county.[48] The scale was 1 to 5 with 1 being 'much more likely with FAR,' 5 being 'much more likely with IR,' and 3 being 'no difference' between FAR and IR. For receiving services, the mean was 2.4, which indicates that caseworkers and supervisors perceived that families assigned to FAR were somewhat more likely to receive the services they need than were families assigned to IR. For referrals, the mean was 2.4, which indicates that caseworkers and supervisors perceived that families assigned to FAR were somewhat more likely to be referred to other community resources than were families assigned to IR.

Table 4.9

Perceptions of Service Differences between FAR and IR (N = 119)

Question	Mean	Median	Max	Min
Families receive services they need	2.4	3.0	1	4
Families referred to other resources or agencies in community	2.4	2.0	1	4

[48] These are descriptive results so cannot be generalized beyond the sample of 119 caseworkers and supervisors who responded to the survey.

4.3.1.3. Timeliness of services

As displayed in Table 4.10, caseworkers and supervisors who responded to the caseworker general survey perceived that families assigned to the FAR track were somewhat more likely to receive services quickly.[49] Based on findings from the site-visit report, a majority of caseworkers replied that FAR caseworkers were able to access services more quickly in instances of county-provided services, court services, life skills, in-home services, and financial resources. Specifically, it was reported that FAR families can access Core Services from the point of assessment, which theoretically means that the services could be accessed more quickly.

Table 4.10

Perceptions of Service Differences between FAR and IR (N = 119)

Question	Mean	Median	Max	Min
Families receive services quickly	2.4	3.0	1	4

However, results from the caseworker general survey and the focus groups were not supported by findings from the case-specific report, which was used to answer the question "are there differences in the timing of services received between families assigned to FAR and IR?[50] The timeliness of service provision was compared between families assigned to FAR who received services and families assigned to IR who received services using a weighted multiple logistic regression model fit to the data.[51] The difference between the FAR and IR tracks was statistically significant for timeliness of service provision, $\chi^2(1) = 5.66$, $p = 0.02$, OR = 0.49, CI[0.28, 0.88].[52] This result indicates that, on average, families assigned to the FAR track had

[49] The scale was 1 to 5 with 1 being 'much more likely with FAR,' 5 being 'much more likely with IR,' and 3 being 'no difference' between FAR and IR. For receiving services quickly, the mean was 2.4. These are descriptive results so cannot be generalized beyond the sample of 119 caseworkers and supervisors who responded to the survey.

[50] The following item was used to answer this research question: "If yes [i.e., if services received], how soon after the initial report date did the family receive services?' (within one week, within two weeks, within three weeks, within four or more weeks, family was not offered services, uncertain). The responses were combined to create a variable indicating whether service was provided within three weeks or after three weeks. Three weeks was chosen as a reasonable timeframe during which to organize service provision. Note that 'uncertain' responses were considered missing. Responses which indicated the family did not receive services were also considered missing because lack of service receipt is inconsistent with the premise of the item (i.e., that it should only be answered if services were received). Because families which did not receive services were excluded and because some responses were missing, the actual sample size for the analysis was 539 families.

[51] The timeliness of providing services for families who did receive them was compared between families assigned to FAR and families assigned to IR. Weighted multiple logistic regression models were fit to the data. The analysis included *only* cases where the caseworker indicated that services were actually provided to the family. The 'how soon' dependent variable could take a value of 'less than three weeks' or 'more than three weeks'. The independent variables included DR track, county, number of children in the home, number of caregivers and age of youngest child. However, neither county, nor any of the demographic variables were statistically significant. The three demographic variables were dropped from the models.

[52] See Table M.3 in Appendix M for details on the logistic regression model.

half the odds of receiving services within three weeks than did families assigned to the IR track.[53] As shown in table 4.11, almost 84% of FAR families (who received services) received services within three weeks while approximately 16% waited more than three weeks to receive services. More than 90% of IR families (who received services) received services within three weeks while fewer than 10% waited more than three weeks to receive services. However, these results compare two very small and similar rates of occurrence (i.e., only 16% of FAR and 10% of IR did not receive services within 3 weeks), thus there are likely no practical differences in timeliness of service provision between the two tracks.

Table 4.11

Weighted Frequency of Service Timing for Families that Received Services (N = 1,719)

Track	Within 1 week	Within 2 weeks	Within 3 weeks	More than 3 weeks	Total
FAR	510 (40.6%)	332 (26.4%)	211 (16.8%)	204 (16.2%)	1,257 (73.1%)
IR	227 (49.1%)	139 (30.1%)	55 (11.9%)	41 (8.9%)	462 (26.9%)
Total	737 (42.9%)	471 (27.4%)	266 (15.4%)	245 (14.3%)	1,719 (100.0%)

4.3.1.4. Service needs

The case-specific report was used to answer the question "are there differences in matching of service to needs and in effectiveness of services between families assigned to FAR and IR?"[54] The matching of services to needs and the effectiveness of services was compared between families assigned to FAR and families assigned to IR using a weighted multiple logistic regression model fit to the data.[55] The difference between DR tracks was not statistically

[53] This result was poorly estimated, as indicated by a wide confidence interval. When using logistic regression, this can happen when one of the two binary outcomes occurs very infrequently. In this case, the 'more than 3 weeks' response was not frequently used.

[54] The following items were used to answer these research questions: 'overall, how well were the services that were actually provided matched to the service needs of the family' (very well matched, somewhat matched, not very well matched, not at all matched) and 'overall, how effective were the services provided to the family in solving their problems or in producing needed changes' (very effective, somewhat effective, not very effective, not at all effective). The service-match and service-effectiveness items were recoded to reflect 'not/not very' or 'somewhat/very' matched or effective. Responses of 'not at all' or 'not very' corresponded to 'not/not very' match/effectiveness and responses of 'somewhat' or 'very' corresponded to somewhat/very matched or effective. These variables were analyzed only for families who did receive services. The actual sample size for both analyses (service match and service effectiveness) was 607 families which both received services and had non-missing responses.

[55] The difference in service match and service effectiveness for families who received services was compared between families assigned to FAR and families assigned to IR. Weighted multiple logistic regression models were fit to the data. The dataset included *only* cases where the caseworker indicated that services were actually provided to the family. ' statistically significant for service match, $\chi^2(4) = 2790.91$ and $p < .0001$, and for service effectiveness, $\chi^2(4) = 18.54$ and $p = 0.001$. However, none of the three demographic variables were statistically significant, so they were dropped from the models.

significant for level of service match, $\chi^2(1) = 0.14$, $p = 0.712$, OR = 0.83, CI[0.30, 2.28][56] or for level of service effectiveness, $\chi^2(1) = 0.48$, $p = 0.489$, OR = 0.82, CI[0.47, 1.43].[57] This result indicates that, on average, FAR families received a similar level of service match and service effectiveness as did IR families as reported by caseworkers.

Information about service needs was also captured through the family exit survey. The survey included an open-ended follow-up question that asked respondents to describe the services they needed but did not receive. Of the 144 families that completed this question, 70% reported needing help that was not received. The most common type of help that was not received was financial assistance (aid and referrals), which was identified by 42% of respondents to this question. Nineteen percent of respondents to this question reported needing but not receiving mental health services or referrals (e.g., domestic violence services, individual and family counseling), ten percent mentioned medical, dental, or disability services, eight percent indicated job or education-related assistance, and seven percent reported needing but not receiving advocacy help to address concerns with their case or caseworker. Less than five percent of respondents to this question reported needing but not receiving help in other areas, such as legal services, parenting education, and respite services.

4.3.2. Caseworker Contacts

The case-specific report was used to answer the question "are there differences in the number of face-to-face and telephone contacts between caseworkers and families assigned to FAR and IR?"[58] The number of face-to-face (F2F) and phone contacts with a caseworker was compared between families assigned to FAR and families assigned to IR using a weighted multiple logistic regression model fit to the data.[59] The track variable was statistically significant, $F(1, 1671) = 5.88$, $p = .02$, $\beta_{FAR} = 3.90$ and CI[3.45, 4.35], $\beta_{IR} = 3.46$, CI[2.98, 3.93].[60] The estimated average difference in face-to-face contacts for families assigned to FAR versus IR was statistically significant and equal to 0.44, CI [0.08, 0.80], $t(1671) = 2.43$, $p = .02$. This means

[56] See Table M.4 in Appendix M for details on the logistic regression model.

[57] See Table M.5 in Appendix M for details on the logistic regression model.

[58] The two items used for these analyses were: 'number of contacts with family: face-to-face meetings…' and 'number of contacts with family: telephone contacts….' There was no recoding of variable values for these analyses. The actual sample for the caseworker contact analyses was 1,672 families, as no responses were missing.

[59] Independent variables in the model included DR track, county, number of children in the home, number of caregivers and age of youngest child. County was statistically significant, $F(4, 1671) = 4.46$ and $p = .004$. Of the demographic variables tested, none were statistically significant. Therefore, all three were dropped from the model.

[60] See Table M.6 in Appendix M for details on the logistic regression model.

that, on average, families assigned to the FAR track received almost half of a face-to-face contact *more* than did families assigned to the IR track. As displayed in Table 4.12, families in the FAR tracked averaged 3.5 F2F contacts as compared to 3.1 F2F contacts for families in the IR track.

The number of telephone contacts with a caseworker was also compared between families assigned to FAR and families assigned to IR using a weighted multiple logistic regression model fit to the data.[61] The track variable was statistically significant, $F(1, 1671) = 40.02$, $p = <.0001$, $\beta_{FAR} = 10.02$ and CI[8.63, 11.42], $\beta_{IR} = 7.41$, CI[6.13, 8.68].[62] The estimated average difference in telephone contacts for families assigned to FAR versus IR was 2.62, CI [1.80, 3.42], $t(1671) = 6.33$, $p < .0001$. This means that, on average, families assigned to the FAR track received about two and one-half telephone contacts *more* than did families assigned to the IR track. As displayed in Table 4.12, families in the FAR tracked averaged 7.3 phone contacts as compared to 4.7 phone contacts for families in the IR track.

Table 4.12

Estimated Average Number of Caseworker Contacts for FAR and IR Families

Track	Average Number of Face-to-face Contacts	Average Number of Telephone contacts
FAR	3.54	7.33
IR	3.09	4.71
Difference	0.44*	2.62*
* Statistically significant difference; $p < .05$.		

The results from the case-specific report were not supported by the findings from the family exit survey, which was used to answer the question, "do families assigned to FAR experience more face-to-face meetings with the caseworker than do families assigned to IR?"[63] Findings from the ordered logistic regression model[64] indicated no statistically significant

[61] Independent variables included DR track, county, number of children in the home, number of caregivers and age of youngest child. County was statistically significant, $F(4, 1671) = 6.71$ and $p = <.0001$. Of the demographic variables tested, none were statistically significant. Therefore, all three were dropped from the model.

[62] See Table M.7 in Appendix M for details on the logistic regression model.

[63] The survey question that inquired about the number of face-to-face meetings with a caseworker was used to answer this question. Due to low frequencies, the evaluation team recoded the survey response options by condensing the last two categories (6-10 and more than 10 meetings) into one (6 or more meetings).

[64] Item non-responses were removed from the analysis due to using ordered logistic regression. Maintaining the non-responses in analysis would make ordering the response categories and interpreting the analysis difficult.

relationship between the tracks on number of face-to-face meetings ($\chi^2(1) = 0.22$, $p = .639$, OR = 0.91, CI[0.62, 1.34]) based on the caregiver perspective.[65]

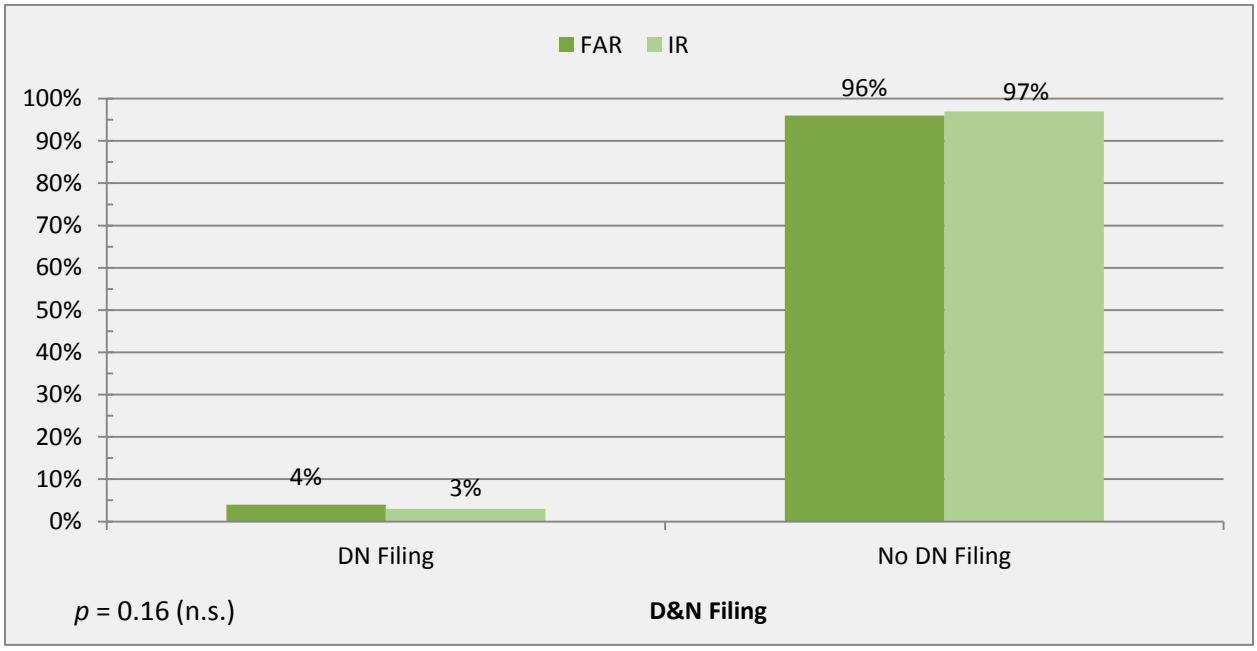
4.3.3. Initial Involvement Events

In describing the results for each initial involvement event, the differences by track are presented first followed by the relationships between the event and family and case characteristics.

D&N Filing during initial involvement: When dependency and neglect (D&N) filings during initial involvement were examined using the administrative data, the stepwise regression analysis did not find a statistically significant difference between FAR and IR families (p = 0.158).[66] As displayed in Figure 4.3, the percentage of FAR families with a D&N filing was 4%, while the percentage of IR families with a D&N filing was 3%.

Figure 4.3

D&N Filing during Initial Involvement by Track



Traditional child welfare case opened during initial involvement: The stepwise regression model results indicate that there was a statistically significant difference between the tracks on

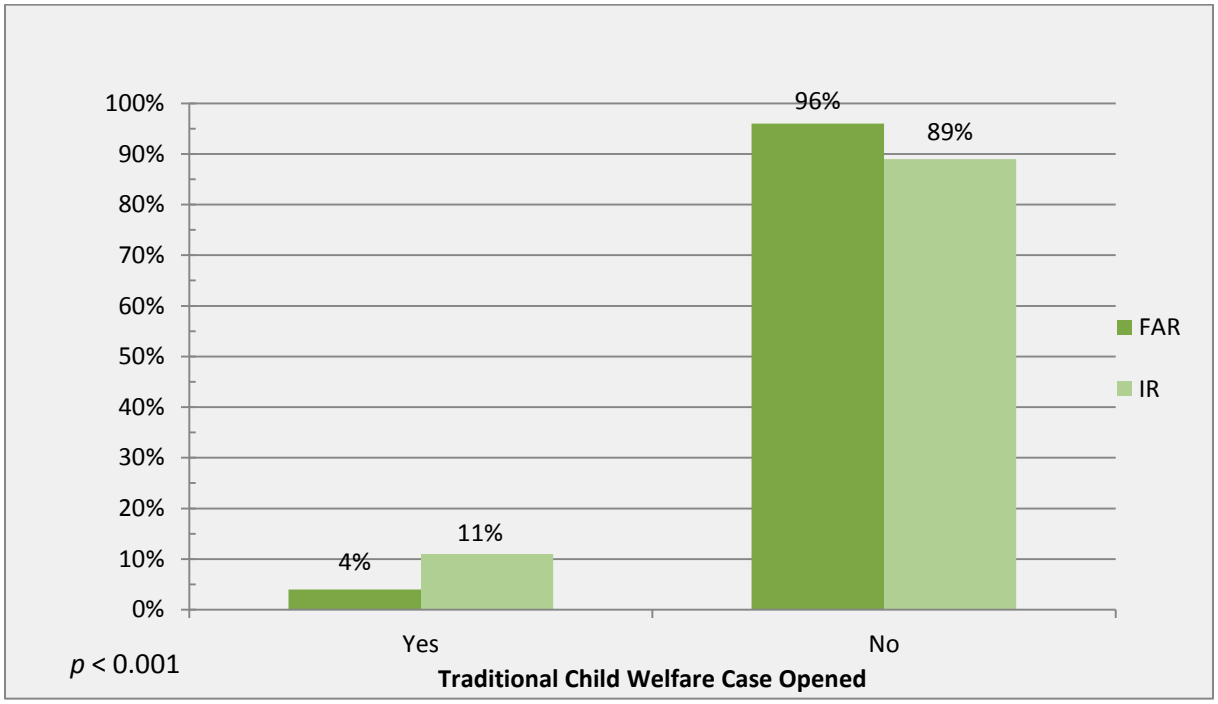
[65] See Table N.9 in Appendix N for details on the logistic regression model. The logistic regression model used track and county as predictors. There were statistically significant differences between counties ($\chi^2(4) = 15.11$, $p = .0045$).

[66] See Table O.1 in Appendix O for the details on the logistic regression model.

traditional child welfare case opened during initial involvement ($p < .001$).[67] FAR families were less likely to have a traditional child welfare case open during initial involvement than were IR families. As displayed in Figure 4.4, fewer FAR families (4%) had traditional CW cases opened during initial involvement than did IR families (11%). However, this is an expected finding because traditional child welfare cases had to be opened so that families assigned to the IR track could receive Core Services. This was not a requirement for families assigned to the FAR track. There were interaction effects between track assignment and number of prior assessments and track assignment and counties for the traditional CW case opened during initial involvement event. Specifically, as the number of prior assessments increased the likelihood of a case opening during initial involvement for FAR families also increased. Fremont County had a higher percentage of cases with a case opened during initial involvement (for both FAR and IR cases), while Larimer County had a higher percentage of case openings for FAR cases.

Figure 4.4

Traditional Child Welfare Case Opening during Initial Involvement by Track



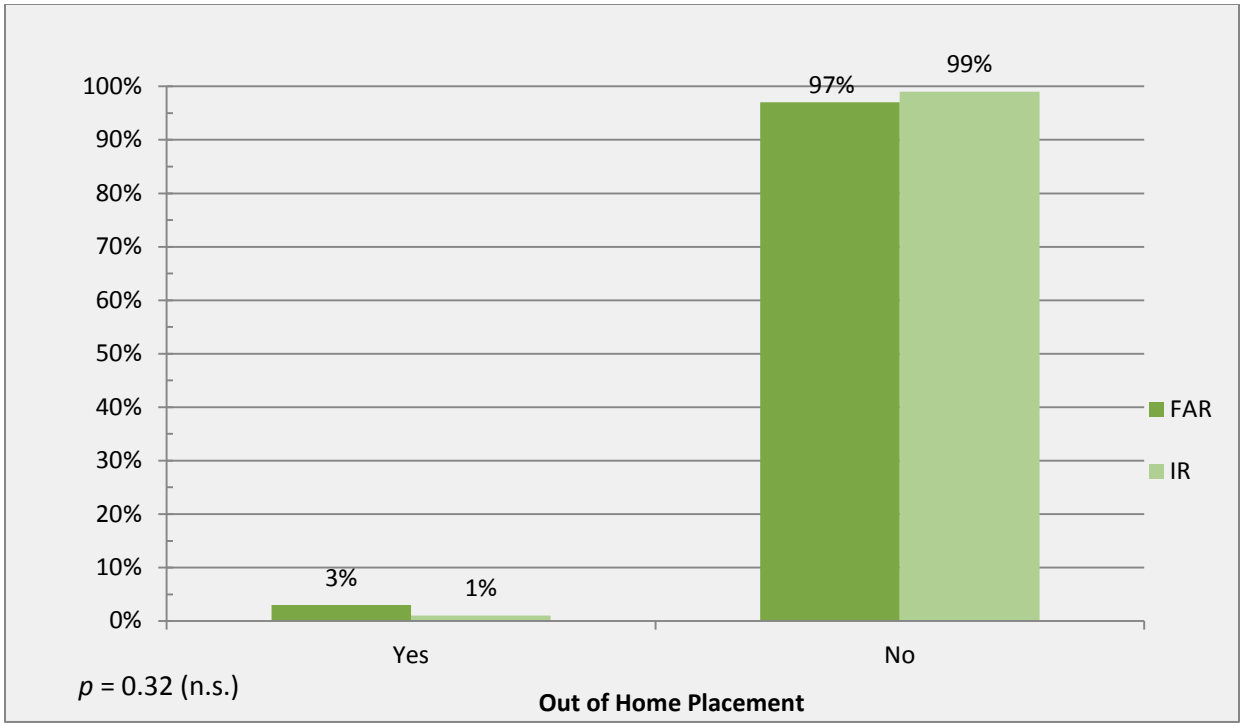
Out-of-home placement during initial involvement: When out-of-home placement during initial involvement was examined, the stepwise regression analysis did not find a statistically

[67] See Table O.2 in Appendix O for the details on the logistic regression model.

significant difference between FAR and IR families (p = 0.317).[68] As displayed in Figure 4.3, the percentage of FAR families with an OOH placement was 3%, while the percentage of IR families with an OOH placement was 1%. There also was an interaction effect between track assignment and number of prior referrals for the OOH placement during initial involvement event. Specifically, as the number of prior referrals increased the likelihood of having an OOH placement during initial involvement for FAR families also increased. In addition, there was an interaction effect between track assignment and domestic violence risk factors for the OOH placement during initial involvement event. Specifically, having a domestic violence risk factor was related to not having an OOH placement during initial involvement for FAR families.

Figure 4.5

Out-of-Home Placement during Initial Involvement by Track



[68] See Table O.3 in Appendix O for the details on the logistic regression model.

5. Results

The results from the outcome evaluation are presented by the main outcome categories: child safety, family well-being, family engagement, caseworker satisfaction, and community buy-in. Table 5.1 presents a summary of the study outcome measures and data sources for each category.

Table 5.1

Summary of Study Outcomes, Measures, and Data Sources

Outcome Category	Outcome Measures	Data Sources
Child Safety	Referral within 365 days of initial referral	Administrative Data
	Assessment within 365 days of initial referral	Administrative Data
	HRA within 365 days of initial referral	Administrative Data
	Founded HRA within 365 days of initial referral	Administrative Data
	CW Case Opened after initial involvement	Administrative Data
	Out-of-home Placement after initial involvement	Administrative Data
	Improvements in Child Safety	Family Exit Survey
Family Well-Being	Improvement in Family Well-being	Family Exit Survey
	Improvement in Family Functioning	Case-Specific Report
	Meeting Family Needs	Family Exit Survey
Family Engagement	Family-centered Practice	Family Exit Survey
	Satisfaction with Caseworker	Family Exit Survey
	Feelings toward CPS	Family Exit Survey
	Family Involvement	Caseworker Survey
	Engagement Characteristics	Case-Specific Report
Caseworker Satisfaction	Overall Satisfaction	Caseworker General Survey
	Training Satisfaction	Focus Group
	Supervision Satisfaction	Focus Group
	Coaching Satisfaction	Focus Group
	Workload Satisfaction	Focus Group
	Service Satisfaction	Caseworker Survey
	Satisfaction with FAR/IR Relationship	Focus Group
Community Buy-in	Awareness and Understanding of DR	Focus Group
	Community Outreach and Training	Focus Group
	Formal Involvement of Stakeholders	Focus Group

5.1. Child Safety Outcomes

To answer the research questions about safety outcomes using administrative data, the evaluation team used stepwise regression[69] to: (1) identify relationships between family and case characteristics and key safety outcomes, and (2) explore differences in safety outcomes across FAR and IR families when adjusting for the effects of these family and case characteristics. Since the probability of assignment to FAR or IR differed among counties and changed over time within counties, the evaluation team weighted the administrative data (see Appendix H for details on the weighting procedures).

The evaluation team worked with the Leadership Team and Data/Evaluation Workgroup to identify a comprehensive list of safety outcomes. Table 5.2 provides a list of the safety outcomes analyzed using stepwise regression. The period examined for most of the outcomes was within 365 days of initial referral; however, several outcomes measured after initial involvement also were analyzed. Safety outcomes within three days of the original referral were assumed to be related to the initial referral. Safety outcomes later than three days after the initial referral were classified as re-involvement outcomes for the analyses. In addition, survival analyses were conducted to examine certain safety outcomes over a longer period of time. In describing the results for each outcome, the differences by track are presented first followed by the relationships between the outcome and family and case characteristics. The tables display the results by track (whether statistically significant or not) and by statistically significant predictors. The complete regression model tables are included in Appendix O.

[69] County and track were included in all models, even if they were not significant. Candidate predictors included family and case characteristics and two-way interactions of factors that were already included in the model as main effects. The models selected in the stepwise procedure were then fit using the SAS SURVEYLOGISTIC procedure to get the final parameter estimates, standard errors, and *p*-values.

Table 5.2

Safety Outcome Descriptions

Name of Outcome	Outcome Description
Referral Within 365 Days of Initial Referral	A new report of child abuse or neglect after initial report regardless of whether report was screened-in or screened-out
Assessment Within 365 Days of Initial Referral	An accepted/screened-in report of child abuse or neglect after the initial report
High-risk Assessment (HRA) Within 365 Days of Initial Referral[70]	An accepted/screened-in report of child abuse or neglect after the initial report that does not meet FAR eligibility criteria (low to moderate risk) and was determined to require a high risk assessment response
Founded HRA Within 365 Days of Initial Referral	An accepted/screened-in report of child abuse or neglect after the initial report that meets the HRA criteria and was determined to be founded as a result of an investigation
Traditional Child Welfare (CW) Case Opened after Initial Involvement	An ongoing child welfare services case opened due to court action or out-of-home placement as a result of a subsequent assessment
Out-of-Home (OOH) Placement after Initial Involvement	A removal from home and placement in OOH care as a result of a subsequent assessment

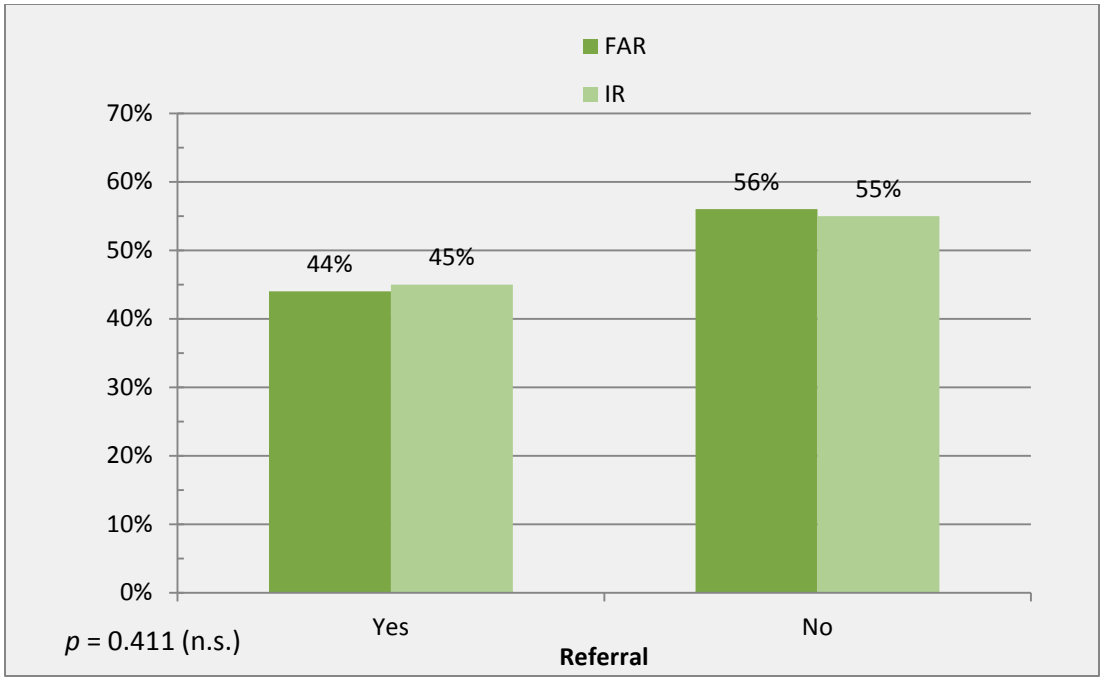
5.1.1. Referral within 365 Days of Initial Referral

The stepwise regression model results indicate that there was no significant difference between the tracks on referral within 365 days of initial referral, $\chi^2(1) = 0.68$, $p = .411$, OR = 0.94, CI[0.82, 1.09]. As displayed in Figure 5.1, the percentage of FAR families that were re-reported was 44%, while the percentage of IR families that were re-reported was 45%.

[70] This outcome allowed for a fairer comparison between the FAR and IR tracks than did subsequent investigation response, because subsequent referrals that were FAR-eligible had to be assigned to investigation response if the family was originally assigned IR, while subsequent FAR-eligible referrals for families originally assigned FAR had to be assigned family assessment response.

Figure 5.1

Referral within 365 Days of Initial Referral by Track



The findings indicate that there were associations between this outcome and a number of family and case characteristics. As displayed in Table 5.3, families in Fremont County and families with more children had a higher probability of a referral within 365 days of initial referral. Families with a youngest child of "other/unknown" ethnicity had a lower probability of this outcome.

Furthermore, there were statistically significant associations between this outcome and a series of risk factors. Families with a higher total risk score, a prior referral within 24 months, a higher number of prior referrals within 24 month, a higher number of total prior referrals, and a domestic violence risk had a higher probability of a referral within 365 days of initial referral.

The model also revealed statistically significant interactions between several family and case characteristics and this outcome including caregiver age and mental health risk factors, risk total score and number of prior referrals, and number of children in the home and prior referral within 24 months of initial referral.[71]

[71] See Table O.4 in Appendix O for details on the logistic regression model.

Table 5.3

Track and Significant Predictors of Referral within 365 Days of Initial Referral

Characteristic	Response Categories	Referral Weighted *N* (%) (*N* = 2,227)	No Referral Weighted *N* (%) (*N* = 2,769)	Total Weighted *N* (%) (*N* = 4,996)
Track	FAR	1,407 (44%)	1,787 (56%)	3,194 (100%)
	IR	820 (45%)	982 (55%)	1,802 (100%)
County	Arapahoe County	716 (42%)	1,007 (58%)	1,723 (100%)
	Fremont County	85 (61%)	53 (39%)	138 (100%)
	Garfield County	123 (41%)	175 (59%)	298 (100%)
	Jefferson County	767 (43%)	1,016 (57%)	1,783 (100%)
	Larimer County	536 (51%)	518 (49%)	1,054 (100%)
Ethnicity of Youngest Child	Caucasian	1,129 (50%)	1,117 (50%)	2,246 (100%)
	Hispanic	593 (50%)	587 (50%)	1,180 (100%)
	African-American	249 (46%)	293 (54%)	542 (100%)
	Other or Unknown	256 (25%)	772 (75%)	1,028 (100%)
Prior Referral in Past 24 months	Yes	1,412 (58%)	1,037 (42%)	2,449 (100%)
	No	815 (32%)	1,732 (68%)	2,547 (100%)
Domestic Violence Risk	Yes	910 (50%)	905 (50%)	1,815 (100%)
	No	1,317 (41%)	1,864 (59%)	3,181(100%)
		Weighted M	*Weighted M*	*Total Weighted M*
Risk Total Score (N = 4,820)		5.21	2.49	3.70
Number of Prior Referrals within 24 months (N = 4,996)		1.66	.74	1.15
Number of Prior Referrals (N = 4,996)		4.04	1.95	2.89
Number of Children in the Home (N = 4,996)		2.01	1.79	1.89

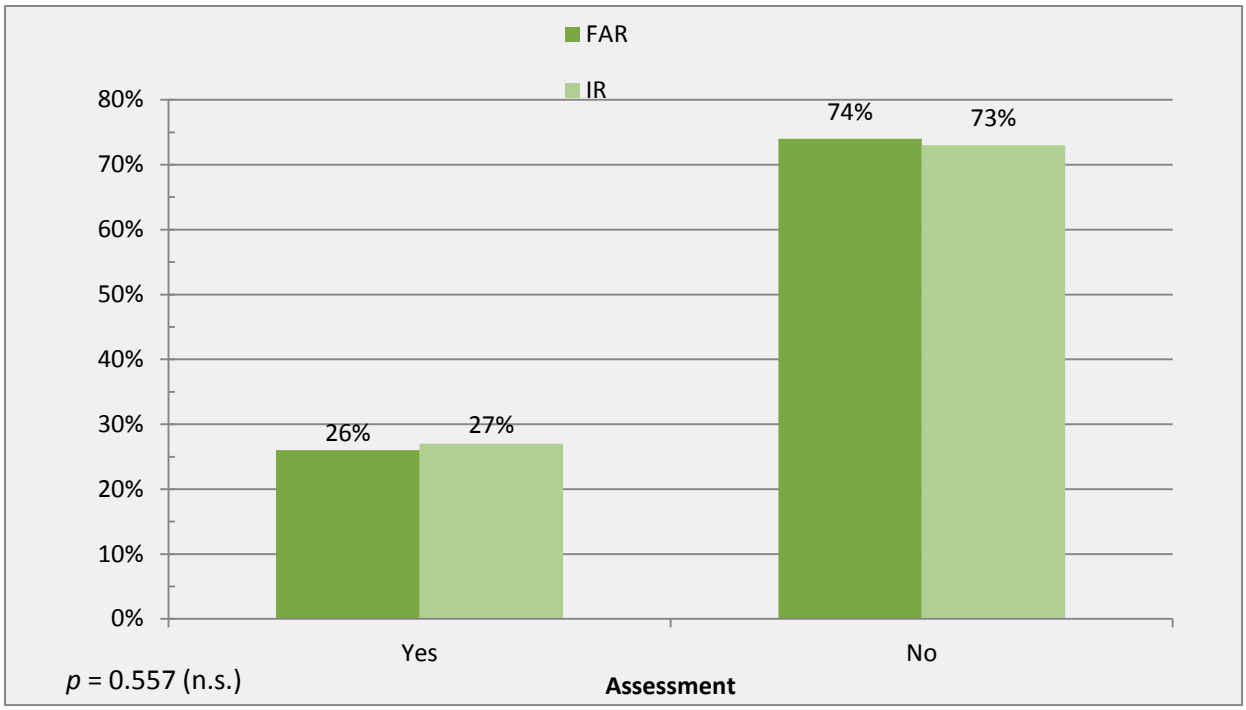
5.1.2. Assessment within 365 days of Initial Referral

The stepwise regression model results indicate that there was no significant difference between the tracks on assessment within 365 days of initial referral, $\chi^2(1) = 0.35$, $p = .557$, OR = 0.95, CI[0.82, 1.12]. As displayed in Figure 5.2, the percentage of FAR families that were re-assessed was 26%, while the percentage of IR families that were re-assessed was 27%.

The findings indicate that there were associations between this outcome and a number of family and case characteristics. As shown in Table 5.4, families in Fremont County and families with a "younger" youngest child had a higher probability of an assessment within 365 days of initial referral. Families with a youngest child of "other/unknown" ethnicity and families with an "older" caregiver had a lower probability of this outcome.

Figure 5.2

Assessment within 365 Days of Initial Referral by Track



Furthermore, there were statistically significant associations between this outcome and a series of risk factors. Families with a higher risk total score, mental health risk factor, prior child welfare case, prior referral within 24 months, and a higher number of prior referrals had a higher probability of an assessment within 365 days of initial referral.

The model also revealed statistically significant interactions between several family and case characteristics and this outcome including county and ethnicity of youngest child, risk total score and number of prior referrals, and mental health risk factors, and prior referral within 24 months of initial referral.[72]

[72] See Table O.5 in Appendix O for details on the logistic regression model.

Table 5.4

Track and Significant Predictors of Assessment within 365 Days of Initial Referral

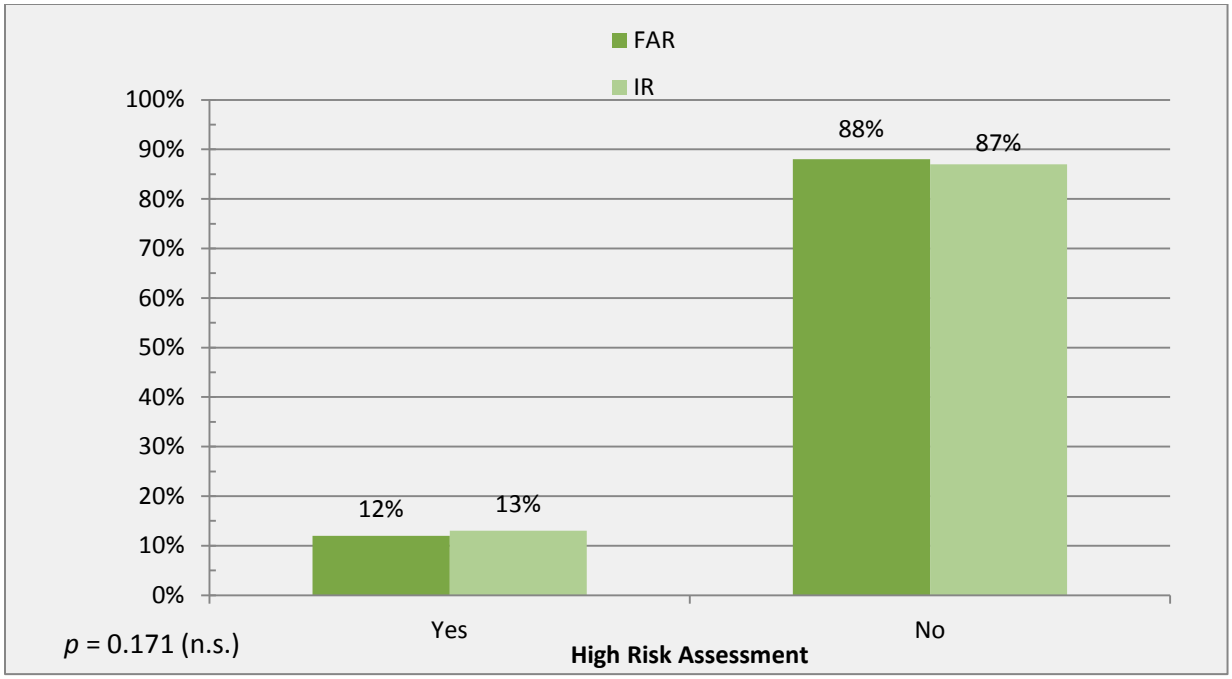
Characteristics	Response Categories	Assessment Weighted *N* (%) (*N* = 1,327)	No Assessment Weighted *N* (%) (*N* = 3,669)	Total Weighted *N* (%) (*N* = 4,996)
Track	FAR	837 (26%)	2,357 (74%)	3,194 (100%)
	IR	490 (27%)	1,312 (73%)	1,802 (100%)
Mental Health Risk	Yes	385 (35%)	718 (65%)	1,103 (100%)
	No	942 (24%)	2,951 (76%)	3,893 (100%)
County	Arapahoe County	419 (24%)	1,304 (76%)	1,723 (100%)
	Fremont County	56 (40%)	82 (60%)	138 (100%)
	Garfield County	74 (25 %)	224 (75%)	298 (100%)
	Jefferson County	503 (28%)	1,281 (72%)	1,783 (100%)
	Larimer County	276 (26%)	779 (74%)	1,054 (100%)
Ethnicity of Youngest Child	Caucasian	684 (30%)	1,562 (70%)	2,246 (100%)
	Hispanic	382 (32%)	799 (68%)	1,181 (100%)
	African-American	167 (31%)	374 (69%)	541 (100%)
	Other or Unknown	94 (9%)	935 (91%)	1,028 (100%)
Prior Referral in Past 24 months	Yes	888 (36%)	1,561 (64%)	2,449 (100%)
	No	439 (17%)	2,108 (83%)	2,547 (100%)
Prior Child Welfare Case	Yes	331 (36%)	580 (64%)	911 (100%)
	No	996 (24%)	3,089 (76%)	4,085 (100%)
		Weighted M	*Weighted M*	*Total Weighted M*
Age of Youngest Child (N = 4,996)		5.16	5.92	5.72
Caregiver Age (N = 4,819)		32.58	33.70	33.39
Risk Total Score (N = 4,820)		5.59	3.02	3.70
Number of Prior Referrals (N = 4,996)		4.26	2.38	2.88

5.1.3. High Risk Assessment within 365 Days of Initial Referral

The stepwise regression model results indicate that there was no significant difference between the tracks on high-risk assessment within 365 days of initial referral, $\chi^2(1) = 1.88$, $p = .171$, OR = 0.87, CI[0.71, 1.06]. As displayed in Figure 5.3, the percentage of FAR families that had a HRA was 12%, while the percentage of IR families that had a HRA was 13%.

Figure 5.3

High Risk Assessment within 365 Days of Initial Referral by Track



The findings indicate that there were associations between this outcome and a number of family and case characteristics. As shown in Table 5.5, families in Fremont County and families with more children in the home had a higher probability of a HRA within 365 days of initial referral. Families with a youngest child of "other/unknown" ethnicity had a lower probability of this outcome.

Furthermore, there were statistically significant associations between this outcome and a series of risk factors. Families with a higher risk total score, a mental health risk factor, and a prior referral within 24 months had a higher probability of a HRA within 365 days of initial referral.

The model also revealed a statistically significant interaction between family and case characteristics and this outcome including county and ethnicity of youngest child.[73]

[73] See Table O.6 in Appendix O for details on the logistic regression model.

Table 5.5

Track and Significant Predictors of HRA within 365 Days of Initial Referral

Characteristic	Response Categories	HRA Weighted *N* (%) (*N* = 633)	No HRA Weighted *N* (%) (*N* = 4,363)	Total Weighted *N* (%) *N* = (4,996)
Track	FAR	390 (12%)	2,804 (88%)	3,194 (100%)
	IR	243 (13%)	1,559 (87%)	1,802 (100%)
Mental Health Risk	Yes	187 (17%)	916 (83%)	1,103 (100%)
	No	446 (11%)	3,447 (89%)	3,893 (100%)
County	Arapahoe County	173 (10%)	1,550 (90%)	1,723 (100%)
	Fremont County	37 (27%)	101 (73%)	138 (100%)
	Garfield County	39 (13%)	258 (87%)	297 (100%)
	Jefferson County	273 (15%)	1,510 (85%)	1,783 (100%)
	Larimer County	111 (11%)	943 (89%)	1,054 (100%)
Ethnicity of Youngest Child	Caucasian	327 (15%)	1,919 (85%)	2,246 (100%)
	Hispanic	183 (16%)	997 (84%)	1,181 (100%)
	African-American	78 (14%)	463 (86%)	541 (100%)
	Other or Unknown	45 (4%)	983 (96%)	1,028 (100%)
Had Prior Referral in Past 24 months	Yes	443 (18%)	2,006 (82%)	2,449 (100%)
	No	190 (7%)	2,357 (93%)	2,547 (100%)
		Weighted M	*Weighted M*	*Total Weighted M*
Age of Youngest Child (N = 4,996)		4.91	5.84	5.72
Risk Total Score (N = 4,820)		6.15	3.35	3.70
Number of Children in the Home (N = 4,996)		2.16	1.85	1.89

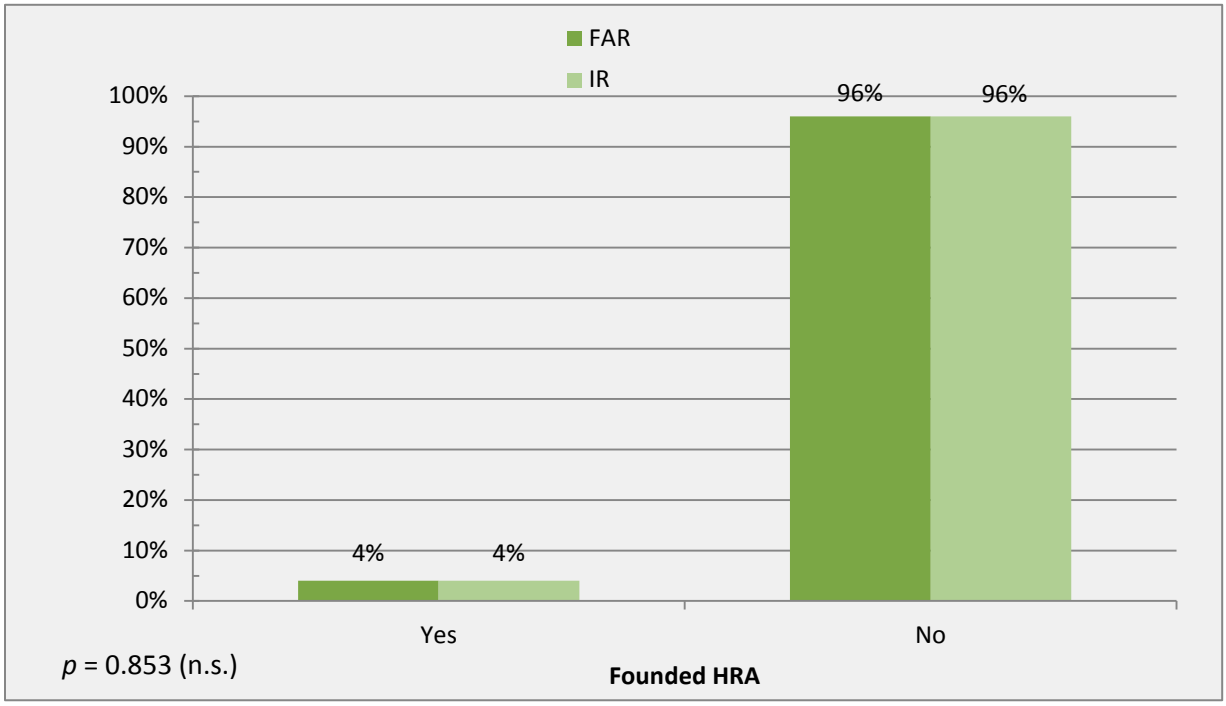
5.1.4. Founded HRA within 365 Days of Initial Referral

The stepwise regression model results indicate that there was no significant difference between the tracks on founded high-risk assessment within 365 days of initial referral, $\chi^2(1) = 0.03$, $p = .853$, OR = 1.03, CI[0.74, 1.44]. As displayed in Figure 5.4, the percentage of FAR families that had a founded HRA was 4%, while the percentage of IR families that had a founded HRA also was 4%.

The findings indicate that there were associations between this outcome and a number of family and case characteristics. As shown in Table 5.6, families in Fremont County and families with a "younger" youngest child had a higher probability of a founded HRA within 365 days of initial referral. Families with a youngest child of "other/unknown" ethnicity had a lower probability of this outcome.

Figure 5.4

Founded HRA within 365 Days of Initial Referral by Track



Furthermore, there were statistically significant associations between this outcome and a series of risk factors. Families with a higher total risk score, having a prior referral within 24 months, and having a neglect allegation had a higher probability of a founded HRA within 365 days of initial referral.

The model also revealed a statistically significant interaction between family and case characteristics and this outcome including age of youngest child and neglect allegation.[74]

[74] See Table O.7 in Appendix O for details on the logistic regression model.

Table 5.6

Track and Significant Predictors of Founded HRA within 365 Days of Initial Referral

Characteristic	Response Categories	Founded HRA Weighted *N* (%) (*N* = 220)	No Founded HRA Weighted *N* (%) (*N* = 4,776)	Total Weighted *N* (%) (*N* = 4, 996)
Track	FAR	142 (4%)	3,052 (96%)	3,194 (100%)
	IR	79 (4%)	1,723 (96%)	1,802 (100%)
County	Arapahoe County	72 (4%)	1,651 (96%)	1,723 (100%)
	Fremont County	13 (9%)	125 (91%)	138 (100%)
	Garfield County	13 (4%)	284 (96%)	298 (100%)
	Jefferson County	92 (5%)	1,692(95%)	1,783 (100%)
	Larimer County	30 (3%)	1,024 (97%)	1,054 (100%)
Ethnicity of Youngest Child	Caucasian	119 (5%)	2,127 (95%)	2,246 (100%)
	Hispanic	69 (6%)	1,111 (94%)	1,181 (100%)
	African-American	29 (5%)	512 (95%)	541 (100%)
	Other or Unknown	3 (<1%)	1,025 (100%)	1,028 (100%)
Neglect Allegation	Yes	188 (5%)	3,501 (95%)	2,449 (100%)
	No	32 (2%)	1,291 (98%)	2,547 (100%)
Had Prior Referral in Past 24 months	Yes	153 (6%)	2,296 (94%)	2,449 (100%)
	No	67 (3%)	2,480 (97%)	2,547 (100%)
		Weighted M	*Weighted M*	*Total Weighted M*
Age of Youngest Child (N = 4,996)		4.15	5.79	5.72
Risk Total Score (N =4,280)		6.56	3.58	3.70

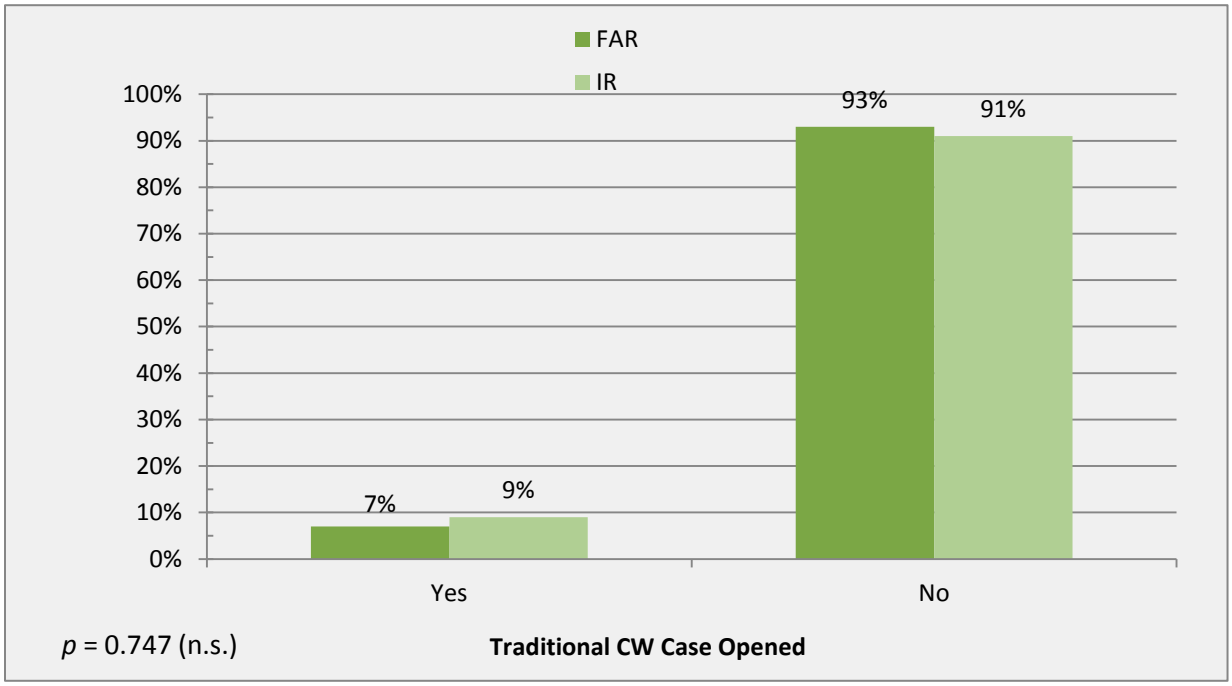
5.1.5. Traditional Child Welfare Case Opened after Initial Involvement

The stepwise regression model results indicate that there was no significant difference between the tracks on traditional child welfare case opened after initial involvement, $p = .747$. As displayed in Figure 5.5, the percentage of FAR families that had a traditional CW case opened after initial involvement was 7%, while the percentage of IR families that had a traditional CW case opened after initial involvement was 9%.

The findings indicate that there were associations between this outcome and a number of family and case characteristics. As shown in Table 5.7, families in Fremont County and families with a "younger" youngest child had a higher probability of a traditional CW case opened after initial involvement. Families with a youngest child of "other/unknown" ethnicity had a lower probability of this outcome.

Figure 5.5

Traditional CW Case Opened after Initial Involvement by Track



Furthermore, there were statistically significant associations between this outcome and a series of risk factors. Families with a higher risk total score, mental health risk, prior referral within 24 months, and higher number of prior referrals had a higher probability of a traditional CW case opened after initial involvement.

The model also revealed statistically significant interactions between several family and case characteristics and this outcome including age of youngest child and neglect allegation, county and risk total score, risk total score and mental health risk, risk total score and number of prior referrals, and track and county.[75]

[75] See Table O.8 in Appendix O for details on the logistic regression model.

Table 5.7

Track and Significant Predictors of Traditional CW Case Opened after Initial Involvement

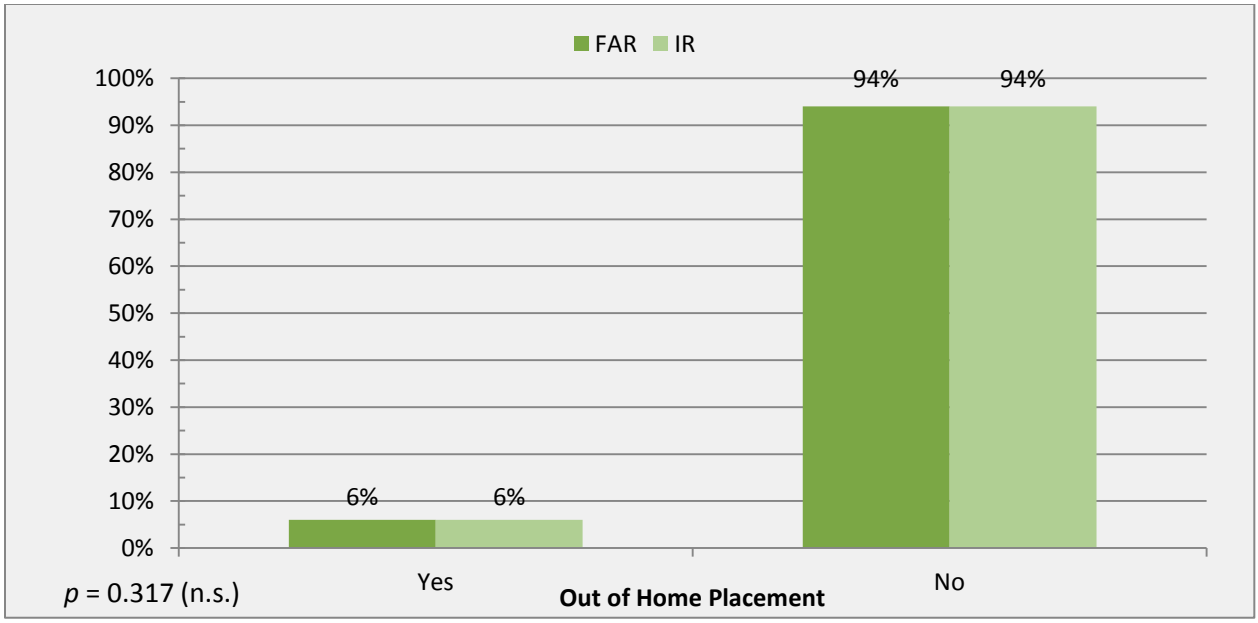
Characteristic	Response Categories	Subsequent Traditional CW Case Weighted *N* (%) (*N* = 394)	No Subsequent Traditional CW Case Weighted *N* (%) (*N* = 4,602)	Total Weighted *N* (%) (*N* = 4,996)
Track	FAR	234 (7%)	2,960 (93%)	3,194 (100%)
	IR	160 (9%)	1,642 (91%)	1,802 (100%)
County	Arapahoe County	106 (6%)	1,617 (94%)	1,723 (100%)
	Fremont County	23 (16%)	115 (84%)	138 (100%)
	Garfield County	23 (8%)	274 (92%)	297 (100%)
	Jefferson County	130 (7%)	1,653 (93%)	1,783 (100%)
	Larimer County	112 (11%)	942 (89%)	1,054 (100%)
Ethnicity of Youngest Child	Caucasian	222 (10%)	2,024 (90%)	2,246 (100%)
	Hispanic	120 (10%)	1,061 (90%)	1,181 (100%)
	African-American	46 (8%)	496 (92%)	541 (100%)
	Other or Unknown	7 (1%)	1,021 (99%)	1,028 (100%)
Mental Health Risk	Yes	122 (11%)	981 (89%)	1,103 (100%)
	No	272 (7%)	3,621 (93%)	3,893 (100%)
Prior Referral in Past 24 months	Yes	283 (12%)	2,166 (88%)	2,449 (100%)
	No	111 (4%)	2,436 (96%)	2,547 (100%)
		Weighted M	*Weighted M*	*Total Weighted M*
Age of Youngest Child (N = 4,996)		5.46	5.74	5.72
Risk Total Score (N =4,820)		6.28	3.48	3.70
Number of Prior Referrals (N =4,996)		5.18	2.68	2.88

5.1.6 Out-of-home Placement after Initial Involvement

The stepwise regression model results indicate that there was no significant difference between the tracks on OOH placement after initial involvement, $\chi^2(1) = 0.09$, $p = .759$, OR = 0.96, CI[0.73, 1.26]. As displayed in Figure 5.6, the percentage of FAR families that had an OOH placement after initial involvement was 6%, while the percentage of IR families that had an OOH placement after initial involvement also was 6%.

Figure 5.6

Out-of-home Placement after Initial Involvement by Track



The findings indicate that there were associations between this outcome and a number of family and case characteristics. As shown in Table 5.8, Families with a youngest child of "other/unknown" ethnicity had a lower probability of OOH placement after initial involvement.

Furthermore, there were statistically significant associations between this outcome and a series of risk factors. Families with a mental health risk, higher number of prior assessments within 24 months of initial referral, and a higher number of prior referrals had a higher probability of an OOH placement after initial involvement.

The model also revealed statistically significant interactions between several family and case characteristics and this outcome including risk neglect score and ethnicity of youngest child, county and mental health risk, and risk neglect score and mental health risk.[76]

[76] See Table O.9 in Appendix O for details on the logistic regression model.

Table 5.8

Track and Significant Predictors of Out-of-Home Placement after Initial Involvement

Characteristic	Response Categories	OOH Placement Weighted N (%) (*N* = 295)	No OOH Placement Weighted *N* (%) (*N* = 4,701)	Total Weighted *N* (%) (*N* = 4,996)
Track	FAR	188 (6%)	3,006 (94%)	3,194 (100%)
	IR	108 (6%)	1,694 (94%)	1,802 (100%)
Ethnicity of Youngest Child	Caucasian	158 (7%)	2,087 (93%)	2,246 (100%)
	Hispanic	86 (7%)	1,095 (93%)	1,181 (100%)
	African-American	40 (7%)	501 (93%)	541 (100%)
	Other or Unknown	11 (1%)	1,017 (99%)	1,028 (100%)
Mental Health Risk	Yes	108 (10%)	995 (90%)	1,103 (100%)
	No	187 (5%)	3,706 (95%)	3,893 (100%)
		Weighted M	*Weighted M*	*Total Weighted M*
Number of Prior Referrals (N =4,996)		5.83	2.70	2.88
Number of Prior Assessments within 24 months (N =4,996)		1.01	0.42	0.45

5.1.7. Survival Analysis

Survival analysis is useful for analyzing time-to-event data, such as the time to a first referral after the initial referral in the study. Survival analysis adjusts for different periods of follow-up including possible future referrals that are not observed because the study period is over. An analysis on the number of referrals in a one-year period after the initial referral provides similar information to a survival analysis of the time to the first referral after the initial referral. These two analyses differ in that the survival analysis only looks at the first referral (and not all referrals) and the survival analysis uses data over a longer period of time. Both analyses can model the effects of various predictors on the outcome.[77]

The administrative data provide four dates for survival analysis:

1. Date of the first referral;
2. Date of the first assessment (not all referrals result in an assessment);
3. Date of the first high-risk assessment (not all assessments are HRA); and
4. Date of the first founded high-risk assessment (not all HRA are founded).

[77] There are two types of survival models. The accelerated lifetime model requires an assumption regarding the distribution of the time-to-event data but can be used on relatively small datasets. The Cox proportional hazards model, used for this analysis, requires fewer assumptions and constructs a non-parametric estimate for the distribution of the time-to-event data. Using SAS, the Cox proportional hazards model can be fit using either the PHREG or SURVEYPHREG procedure. For weighted survey data, the SURVEYPHREG procedure provides a better assessment of significance than the weighted PHREG procedure; however, the SURVEYPHREG procedure does not have the capability of doing stepwise selection of predictors.

For each of these dates, the time from initial referral date in the study was calculated and used for the survival analysis. Dates beyond February 28, 2013 were censored, that is, any events after that date were not recorded in the administrative data file used for analysis.

The four outcome measures are related because an assessment or high risk assessment is not conducted unless there is a referral and a founded high risk assessment can only occur if there is a high risk assessment. As a result, factors that predict referral are also likely to predict other outcomes and visa-versa. Therefore, the same final model was fit to each outcome. The predictors in the final model were those that were significant for predicting at least one of the four outcomes. Main effects that were significant at the 5% level and interactions that were significant at the 1% level were kept in the final model.

The Cox proportional hazards model predicts the hazard rate, which is the probability of an event (such as a future referral) per unit time. The model parameters can be expressed as hazard ratios. For continuous variables the hazard ratio is the proportional change in the probability of the event associated with a one unit increase in the predictor. Using referral as an example, a hazard ratio greater than 1.0 indicates an increase in the probability of a first referral per unit time. For a selected level of a categorical variable, the hazard ratio is the ratio of the probability of an event for that category to the probability for the reference category (the category for which the hazard ratio is 1.0). Several steps were used to identify the final survival model fit to the data.[78]

[78] 1. The distributions of the number of prior referrals and assessments were highly skewed. To minimize the effect of a few large counts on the predictions, these variables were transformed using Ln(Count+1).

2. For each outcome, the SAS PHREG procedure was used to select predictors of the outcome variables using the default stepwise selection algorithms. The analysis used the survey weights for the administrative data. Two-way interactions were considered if the associated main effects were already selected for the model.

3. The selected effects for the four outcomes were pooled to construct a draft final model. Main effects that were significant in any model at the 5% level and interactions that were significant in any model at the 1% level were included in the draft final model, based on the overall p-value. Given the relatively large number of possible interactions, the more stringent criteria for selection of interactions was used to reduce the number of interactions selected by chance.

4. The draft final model was fit to all outcomes using the SAS SURVEYPHREG procedure. The results were reviewed and the final model was adjusted. The following changes were made:

- The interaction of county by youngest child's ethnicity was dropped because of numerical problems associated with small counts in some cells.
- The interaction of county by transformed number of prior referrals was not significant at the 1% level in any model based on the SURVEYREG procedure.

5. The resulting final model was fit to all four outcome variables using the SURVEYPHREG procedure. The multi-step process for selecting the final model was used, in part, because the SURVEYREG procedure provides a better assessment of significance than the weighted PHREG procedure; however, the SURVEYREG procedure does not have the capability of doing stepwise selection of predictors.

Table 5.9 shows survival analysis results for predicting the time to the first referral, assessment, high risk assessment, or founded high risk assessment.[79] Since the same model was fit to all four outcome measures, the parameters can be directly compared. The sign and relative magnitude of the various predictors are generally consistent across the outcome measures. The predicted hazard rates for each of the significant predictor variables are the same for both FAR and IR tracks, as there were no significant interactions between tracks and the predictors. The following summarizes the relationships in Table 5.9:

- Compared to an IR assignment, a FAR assignment has a lower probability for referrals, assessments, high risk assessments, and founded high risk assessments. However, the FAR/IR difference was only statistically significant for predicting the probability of a high-risk assessment (p = .01). Thus, FAR assignments were 18% less likely to have a HRA, over time, than were IR assignments (i.e., less by a factor of 0.82).

- Having an "older" youngest child was significantly associated with a lower probability for each of the four outcomes. Thus, families with an "older" youngest child were 2% less likely to have a referral, 4% less likely to have an assessment, 6% less likely to have a HRA, and 6% less likely to have a founded HRA than were families with a "younger" youngest child.

- Having a prior referral within 24 months of initial referral was significantly associated with a higher probability for each of the four outcomes. Thus, families with a prior referral within 24 months were 39% more likely to have a referral, 41% more likely to have an assessment, 52% more likely to have a HRA, and 44% more likely to have a founded HRA than were families without a prior referral within 24 months of initial referral.

- Having a lower or medium risk level was significantly associated with a lower probability for each of the four outcomes. Thus, families with a low risk level were 36% less likely to have a referral, 36% less likely to have an assessment, 41% less likely to have a HRA, and 70% less likely to have a founded HRA than were families with a high risk level. Families with a medium risk level were 20% less likely to have a

[79] The results shown are the parameter estimates and associated p-values. For categorical predictors, the table also shows the overall p-value for assessing the joint contribution of all levels of the categorical values. The parameter estimates are for predicting the hazard rate for a first referral, assessment, HRA, and founded HRA after initial referral. Negative values correspond to lower hazards or probabilities per unit time (i.e., preferred values in the context of child welfare where lower probabilities of referral, assessment, etc. are preferred).

referral, 19% less likely to have an assessment, 33% less likely to have a HRA, and 47% less likely to have a founded HRA than were families with a high risk level.

- Having more children in the home was significantly associated with a higher probability for referrals. Thus, each additional child in the home is associated with a 6% increase in the probability of a referral.

- There were significant county differences for each of the four outcomes. The patterns among counties were somewhat consistent, with Fremont County having the highest probability for referral, assessment, HRA, and founded HRA. The interaction of county and track was not significant in any of the models.

- Caregiver age was a significant predictor for the probability of a referral and assessment, but not for HRA and founded HRA. Families with older caregivers had a lower probability of a referral and assessment than did families with younger caregivers. For every one year increase in caregiver age, there was a 1% decrease in the likelihood of the family having a referral or assessment.

- The interaction between the youngest child's ethnicity and number of prior referrals (transformed) makes interpretation of the main effects for child's ethnicity and number of referrals more complex. However, for referrals, assessments, and high risk assessments the interaction is not significant; as a result the following are based on the parameters for the main effects predicting these outcomes:

    - There are significant differences between the ethnicity of youngest child groups. The patterns among the ethnicity groups are consistent, with the "other/unknown" group having a lower probability for the re-involvement outcomes than the other ethnicity groups.

    - Number of prior referrals (transformed) was a significant predictor for the probability of referral and assessment, but not for founded HRA. Families with fewer prior referrals had a lower probability for referral and assessment than did families with more prior referrals. Although similar patterns were found for HRA, the interaction was significant, indicating that the effect of prior referrals differs somewhat among ethnicity groups.

See Figures P.1 through P.8 in Appendix P for the survival curves corresponding to the various factors in the Cox proportional hazards models. Each figure has four plots for the four outcome events, referral, assessment, high risk assessment, and founded high risk assessment.

Each plot shows the proportion of children with no re-involvement versus days since referral. In this context, higher values are preferred. Predictors that are continuous in the survival model were recoded into roughly equal sized categories for presentation. The survival curves are not adjusted for other predictors in the model. As a result, differences portrayed in the plots may be somewhat different than indicated by the parameter values in Table 5.9.

Table 5.9

Survival Analysis Results Predicting Time to First Referral, Assessment, HRA, and Founded HRA

	Time to first:	Referral		Assessment		HRA		Founded HRA	
Parameter	**Level**	**Hazard Ratio**	**p-value**	**Hazard Ratio**	**p-value**	**Hazard Ratio**	**p-value**	**Hazard Ratio**	**p-value**
Youngest Child Age		0.985	0.0135	0.958	<.0001	0.946	<.0001	0.942	0.0008
Prior Referral, 24 Months		1.394	<.0001	1.411	<.0001	1.520	<.0001	1.440	0.0284
Risk Level	Overall		<.0001		0.0027		0.0006		<.0001
	Low	0.639	<.0001	0.643	0.0008	0.592	0.0038	0.296	0.0002
	Medium	0.802	0.0014	0.810	0.0089	0.670	0.0002	0.528	<.0001
	High	1.000	.	1.000	.	1.000	.	1.000	.
Youngest Child Ethnicity	Overall		<.0001		<.0001		<.0001		<.0001
	Caucasian	1.862	<.0001	3.394	<.0001	4.275	<.0001	20.005	<.0001
	Hispanic	1.626	<.0001	3.180	<.0001	4.842	<.0001	22.726	<.0001
	African American	1.827	<.0001	3.915	<.0001	4.541	<.0001	20.086	<.0001
	Other/Unknown	1.000	.	1.000	.	1.000	.	1.000	.
Transformed Number of Prior Referrals		1.541	<.0001	1.707	0.0006	2.389	<.0001	1.094	0.9172
Number of Children in the Home		1.058	0.0073	1.009	0.7476	1.054	0.1516	1.060	0.3184
County	Overall		0.0136		<.0001		<.0001		0.0148
	Arapahoe	0.866	0.0225	1.130	0.1131	0.993	0.9541	1.517	0.0320
	Fremont	1.161	0.1779	1.790	<.0001	2.594	<.0001	2.426	0.0053
	Garfield	0.826	0.0454	1.104	0.4044	1.448	0.0209	1.563	0.1037
	Jefferson	0.894	0.0679	1.308	0.0004	1.535	0.0001	1.805	0.0021
	Larimer	1.000	.	1.000	.	1.000	.	1.000	.
Caregiver Age		0.993	0.0223	0.989	0.0035	1.005	0.2858	0.989	0.2046
Transformed Number of Prior Referrals* Youngest Child Ethnicity	Overall		0.1670		0.1060		0.0059		0.8172
	Caucasian	0.835	0.0869	0.786	0.1217	0.568	0.0042	1.161	0.8620
	Hispanic	0.877	0.2333	0.787	0.1348	0.489	0.0005	1.014	0.9869
	African American	0.760	0.0377	0.647	0.0168	0.513	0.0049	1.007	0.9935
	Other/Unknown	1.000	.	1.000	.	1.000	.	1.000	.
Track	FAR	0.961	0.3816	0.975	0.6427	0.820	0.0100	0.932	0.5829

5.1.8. Survey Safety Measures

The family exit survey was used to answer the question, "do families assigned to FAR perceive more improvements in child safety than do families assigned to IR?" A one-item yes/no question about perceptions of improvements in child safety as a result of experiences with CPS was used as the dependent variable for the logistic regression model.[80] The logistic regression model findings indicate no statistically significant difference between tracks on this child safety measure, $\chi^2(1) = 0.21$, $p = .644$, OR = 1.10, CI[0.73, 1.66].[81] However, there was a statistically significant relationship between the child safety measure and parent education, $\chi^2(3) = 18.11$, p = .0004.[82] Specifically, as parent education level increases self-reported child safety decreased. Respondents with "some college or a 2-year degree" had 1.3 times the odds, on average, of reporting that their child was safer than did respondents with a "four year degree, some graduate school, or a graduate degree." Respondents with "less than high school diploma" had 8.0 times the odds, on average, of reporting that their child was safer than did respondents with a "four year degree, some graduate school, or graduate degree."

5.2. Family Well-being Outcomes

Family well-being was measured by family and caseworker responses captured in the family exit survey and case-specific report. Specifically, questions were asked about improvements in family well-being and family functioning, and meeting of family needs.

5.2.1. Improvement in Family Well-being

The family exit survey was used to answer the research question, "do families assigned to FAR perceive more improvements in family well-being than do families assigned to IR?[83] The perception of improvements in family well-being was compared between families assigned to FAR and families assigned to IR using weighted logistic regression models fit to the data. The logistic regression model findings suggested no statistically significant differences between the

[80] To fit the model to a binary logistic regression, we removed the item non-responses.

[81] See Table N.10 in Appendix N for details on the logistic regression model. The logistic regression model included track, county, and income as predictors. Only differences among income groups were statistically significant, ($\chi^2(3) = 10.68$, $p = .014$), with increasing income associated with decreasing perception of improvement in child safety.

[82] See Table N.11 in Appendix N for details on the logistic regression model. The logistic regression model included DR track, county, income, education, and race as predictors. Only differences among education groups were statistically significant.

[83] A family well-being scale was constructed by recoding response options for the questions related to the respondent being better off overall, better as a parent, and being better able to provide for his/her family as a result of their experience with CPS. Response items were recoded as "positive" or "less positive" and a mean score was calculated for each respondent.

tracks on perceptions of improvements in family well-being, $\chi^2(1) = 1.00$, $p = .316$, OR = 1.20, CI[0.84, 1.73].[84] However, there was a statistically significant relationship between parents' education level and self-reported well-being scale scores.[85] Specifically, as education level increases, the likelihood of reporting higher family well-being scores as a result of involvement with CPS decreases. Parents with "less than a high school diploma" had 3.7 times the odds, on average, of reporting higher family well-being scores than did caregivers with a "four year degree or higher."

5.2.2. Improvement in Family Functioning

The case-specific report was used to answer the question, "are there differences in the addressing of and improvement in family functioning between families assigned to FAR and IR?"[86] Table 5.10 displays the percentage of cases which had an indicated need, a need addressed and which experienced moderate or much improvement. The frequency of meeting family needs, as reported by the caseworker, was compared between families assigned to FAR and families assigned to IR using weighted multiple logistic regression models fit to the data.[87]

---

[84] See Table N.12 in Appendix N for details on the logistic regression model. The ordinal logistic regression model included factors for DR track, county, and income. Only differences by income level were significant ($\chi^2(3) = 8.48$, $p = .037$), with increasing income associated with decreasing perception of well-being.

[85] See Table N.13 in Appendix N for details on the logistic regression model. The ordinal logistic regression model included DR track, county, income, education, and race as predictors. Only differences by education level were significant, ($\chi^2(3) = 17.47$, $p = 0.0006$).

[86] Eight family needs were described in the survey: Material needs, Substance abuse, Physical health, Mental health, Parenting skills/discipline, Domestic violence, Education, and Social supports. For each need, the caseworker was asked to 'check all family needs present at case opening.' Then, the caseworker was asked (for each need) whether the condition was addressed while the case was open (yes or no) and what the level of improvement for the need was (none, little, moderate, much). The level of improvement item was recoded to reflect two levels of improvement. Responses of 'none' or 'little' correspond to 'no/little' improvement and responses of 'moderate' or 'much' correspond to 'moderate/much' improvement. The actual sample size for this analysis was 1,672 families.

[87] 'Need addressed' was the dependent variable and it could take a value of 'yes' or 'no'. The logistic regression models were run separately for each of the eight needs. The datasets for each of the models included *only* cases where the need to be analyzed had been indicated *present for the family at case opening*. For each of the eight regressions, the independent variables included DR track, county, number of children in the home, number of caregivers and age of youngest child. Age of youngest child was statistically significant for the physical health needs model, $\chi^2(1) = 7.11$, $p = .008$, OR = .90, CI [.83, .97] and number of caregivers was statistically significant for the social support needs model, $\chi^2(1) = 4.39$, $p = .04$, OR = .30, CI [.10, .93]. The odds ratio for age of youngest child in the physical health needs model indicates that, on average, the odds of meeting a physical health need are about 10% lower for each year increase in the age of the youngest child. The odds ratio for number of caregivers in the social support needs model indicates that, on average, the odds of meeting a social support need are about 70% lower if the family had two caregivers than if it had only one. However, this result was very poorly estimated, probably due to the fact that none of FAR families had both one caregiver *and* unmet social support needs. When there are no observations for a particular combination of variable values, then odds ratio estimates can be poor.

Table 5.10

Weighted Frequency of Cases with Needs, with Needs Addressed, and with Improvement

Need	Percent of Cases with Indicated Need^	Percent of Cases with Indicated Need where 'Need Addressed'*	Percent of Cases with Need Addressed where Improvement Equals 'Moderate/Much'†
Material	996 (19.9%)	839 (84.2%)	413 (49.3%)
FAR	661 (20.7%)	601 (90.9%)	312 (51.9%)
IR	335 (18.6%)	238 (71.0%)	102 (42.9%)
Substance Abuse	1,287 (25.8%)	1,151 (89.4%)	587 (51.0%)
FAR	838 (26.2%)	758 (90.5%)	396 (52.2%)
IR	449 (24.9%)	393 (87.5%)	192 (48.9%)
Physical Health	539 (10.8%)	415 (77.0%)	160 (38.6%)
FAR	349 (10.9%)	270 (77.4%)	112 (41.5%)
IR	190 (10.5%)	144 (75.8%)	48 (33.3%)
Mental Health	1,146 (22.9%)	1,047 (91.3%)	428 (40.9%)
FAR	767 (24.0%)	722 (94.1%)	295 (40.9%)
IR	379 (21.0%)	324 (85.5%)	133 (41.0%)
Parenting Skills	1,501 (30.1%)	1,431 (95.3%)	692 (48.3%)
FAR	1,002 (31.4%)	961 (95.9%)	462 (48.1%)
IR	499 (27.7%)	470 (94.2%)	230 (48.9%)
Domestic Violence	1,034 (20.7%)	955 (92.3%)	550 (57.6%)
FAR	657 (20.6%)	617 (93.9%)	346 (56.1%)
IR	377 (20.9%)	338 (89.7%)	205 (60.7%)
Educational	578 (11.6%)	552 (95.4%)	283 (51.4%)
FAR	398 (12.5%)	382 (96.0%)	195 (51.0%)
IR	180 (10.0%)	170 (94.4%)	88 (51.8%)
Social Support	563 (11.3%)	504 (89.6%)	193 (38.3%)
FAR	382 (12.0%)	348 (91.1%)	143 (41.1%)
IR	181 (10.0%)	156 (86.2%)	50 (32.1%)

^*N* = 4,996 (total sample), 3,194 (FAR), 1,802 (IR).
**N* = Number of Cases with Indicated Need for each Need Area (for total sample, FAR, and IR)
†*N* = Number of Cases with Indicated Need where 'Need Addressed' (for total sample, FAR, and IR)

The difference between tracks in meeting family needs (given that a particular need was present) was statistically significant for two of the eight needs measured by the caseworker case-specific survey: material needs, $\chi^2(1) = 15.84$, $p < .0001$, OR = 3.77, CI[1.96, 7.24][88] and mental health needs, $\chi^2(1) = 7.55$, $p = .006$, OR = 2.90, CI[1.36, 6.19].[89] On average, families assigned to the FAR track had four times the odds of having material needs addressed than did families assigned to the IR track. Families assigned to FAR were also almost had three times the odds of having mental health needs addressed than did families assigned to the IR track. The difference between tracks in meeting family needs (given that a particular need was present) was not statistically significant for six of the eight needs measured by the caseworker case-specific survey: substance abuse needs, physical health needs, parenting skills needs, domestic violence needs, education needs, and social supports needs.[90]

The improvement in family needs ('no/little' or 'moderate/much'), given that the need was addressed, was also compared between families assigned to FAR and IR using weighted multiple logistic regression models fit to the data.[91] There were *no* statistically significant differences between the FAR or IR tracks in improvement of family needs at the 'moderate/much' versus the 'no/little' level.[92] However, several of the covariates were significantly related to the improvement outcomes. Age of youngest child was statistically significant for improvement in material needs ($\chi^2(1) = 4.42$, $p = 0.04$, OR = .94, CI [.89, 1.00]). Number of children was statistically significant for improvement in physical health needs ($\chi^2(1) = 9.05$, $p = .003$, OR = 1.64, CI [1.19, 2.26]). Number of caregivers was significant for improvement in mental health needs ($\chi^2(1) = 4.63$, $p = .03$, OR = .57, CI [.34, .95]). This indicates that, on average, the odds of higher improvement for a material need (i.e., moderate/much instead of no/little) are about 6% lower for each year increase in the age of the youngest child and the odds of higher improvement in a physical health need are about 60% higher for each additional child in the family. The results for number of caregivers in the mental health needs

[88] See Table M.8 in Appendix M for details on the logistic regression model.

[89] See Table M.9 in Appendix M for details on the logistic regression model.

[90] See Tables M.10 – M.15 in Appendix M for details on the logistic regression models.

[91] 'Improvement' was the dependent variable for each model and it could take a value of 'no/little' or 'moderate/much'. The logistic regression models were again run separately for each of eight needs. The analyses for each of the improvement models included *only* cases where the families' need had been addressed during the life of the case. For each of the eight regressions, the independent variables included DR track, county, number of children in the home, number of caregivers and age of youngest child. Non-significant demographic variables were dropped from the model. As with number of caregivers in the model for meeting social support needs (described above), this odds ratio is very poorly estimated.

[92] See Tables M.16 – M.23 in Appendix M for details on the logistic regression models.

model indicates that, on average, the odds of higher improvement in a mental health need are about 43% less if the family had two caregivers than if it had only one.

5.2.3. Meeting Family Needs

The family exit survey was used to answer the question, "do families assigned to FAR believe that more of their needs were met satisfactorily than do families assigned to IR?" A one-item yes/no question about whether the respondent felt their needs were met satisfactorily was used as the dependent variable for the logistic regression model. The model findings revealed no statistically significant difference between the tracks on respondent perceptions of having their needs meet satisfactorily, $\chi^2(1) = 3.02$, $p = .082$, OR = 1.47, CI[0.95, 2.28].[93] Overall, 74% (n = 2,322) of FAR families reported having their needs met satisfactorily compared to 66% (n = 1,152) of IR families, however these differences only approached statistical significance.

To answer the question, "are family characteristics associated with having needs met satisfactorily?" education, race, and income were added as independent variables with track and gender as control variables in the logistic regression model. The model findings indicate a relationship between income and perception of having needs met satisfactorily, $\chi^2(3) = 11.14$, $p = .011$.[94] This indicates that higher income is associated with having needs met satisfactorily. However, due to the non-response for the income variable, the relationship among the income categories was somewhat difficult to interpret. No other relationships were found to be statistically significant.

5.3. Family Engagement Outcomes

Family engagement was measured by family and caseworker responses captured in the family exit survey, case-specific report, and caseworker general survey. Specifically, questions were asked about family-centered caseworker practice, satisfaction with caseworker, feelings toward CPS, family involvement, and engagement characteristics.

5.3.1. Family-centered Practice

The family exit survey was used to answer the question, "are FAR caseworkers more likely to be perceived as having style and skills reflecting family-centered practice than are IR

[93] See Table N.14 in Appendix N for details on the logistic regression model. The logistic regression model included factors for DR track, county, and income. Only differences by income level were statistically significant ($\chi^2(3) = 8.61$, $p = .035$), with increasing income associated with higher perception of needs met.

[94] See Table N.15 in Appendix N for details on the logistic regression model. The logistic regression model included DR track, county, education, race, and income as predictors.

caseworkers?"[95] The perception of family-centered practice was compared between families assigned to FAR and families assigned to IR using weighted ordered logistic regression models fit to the data. The findings indicate that there is a statistically significant difference between the tracks on the family-centered practice scale scores, $\chi^2(1) = 6.49$, $p = .011$, OR = 1.59, CI[1.11, 2.27].[96] On average, families assigned to the FAR track had 1.6 times the odds of scoring the family-centered practice scale as "high" than did families assigned to the IR track.

The findings suggest that there is a relationship between caregivers' education level and the family-centered practice scale scores.[97] Specifically, as education level increases, the likelihood of reporting higher family-centered caseworker practice scores decreases. Respondents with "less than a high school diploma" had 2.9 times the odds, on average, of reporting higher family-centered practice scores than did caregivers with a "four year degree or higher."

5.3.2. Satisfaction with Caseworker

The family exit survey was used to answer the question, "is family satisfaction with the caseworker for families assigned to FAR different from families assigned to IR?"[98] The family satisfaction with the caseworker was compared between families assigned to FAR and families assigned to IR using weighted logistic regression models fit to the data. The logistic regression model findings indicate that there is a statistically significant difference between tracks on family satisfaction scores, $\chi^2(1) = 6.92$, $p = .0085$, OR = 1.61, CI[1.13, 2.29].[99] On average, FAR families had 1.6 times the odds of rating satisfaction with their caseworker as high versus low

[95] A family-centered practice scale was developed using seven items from the survey. These items focused on the caseworker listening to the family, understanding needs, discussing important things, considering the family's opinion, recognizing strengths, accessibility of the caseworker, and if the family was offered services in their preferred language. Items were re-coded as "positive" or "less positive" and a mean score was calculated for each respondent.

[96] See Table N.16 in Appendix N for details on regression model. The model included DR track, county, and gender as predictors. Only differences by gender were significant ($\chi^2(1) = 4.55$, $p = .033$, OR = 0.50, CI[0.26, 0.95]), with female respondents more likely to score the family-centered practice scale as "high."

[97] See Table N.17 in Appendix N for details on regression model. The ordinal logistic regression model included DR track, county, income, education, race, gender, and number of children in the home as predictors. Differences by education and number of children in the home were statistically significant. Perception of family centered practices were lower for respondents with higher education level, ($\chi^2(3) = 8.17$, $p = .043$) and lower for families with more children in the home ($p = 0.047$).

[98] A family satisfaction scale was constructed by utilizing questions that inquired about satisfaction with treatment, help received, and if the respondent would call CPS in the future. The questions were re-coded as "positive" or "less positive" and then computed a mean score.

[99] See Table N.18 in Appendix N for more detail on the regression model. The logistic regression model included DR track and county as predictors. Differences between counties were not significant.

than did IR families. The logistic regression model[100] indicates no statistically significant relationships between the family satisfaction scale and family characteristics, except for education, $\chi^2(3) = 7.85$, $p = .049$.[101]

5.3.3. Feelings Toward CPS

The family exit survey was used to answer the question, "are FAR families more likely to call CPS in the future than are IR families?"[102] The willingness to call CPS in the future was compared between families assigned to FAR and families assigned to IR using weighted logistic regression models fit to the data. The logistic regression model findings indicate a statistically significant difference between the tracks on this family engagement measure $\chi^2(1) = 6.99$, $p = .0082$, OR = 1.71, CI[1.15, 2.54].[103] On average, FAR families had 1.7 times the odds of being willing to call CPS in the future than did IR families.

The family exit survey was used to answer the question, "do families assigned to FAR have different feelings after the first time a CPS caseworker came to their home than do families assigned to IR?" Survey respondents were asked to identify from a list of twelve emotions how they felt after their initial meeting with their caseworker. A feelings scale[104] was included as the dependent variable in a linear regression model. The linear regression model revealed a statistically significant difference between the tracks on the feelings scale, $p = .012$,[105] with FAR families reporting higher scores by an average of 0.61 units on the feelings scale.

[100] Each model included education, income, gender, and race. Track and county were kept in the model as predictor variables.

[101] See Table N.19 in Appendix N for more detail on the logistic regression model. The logistic regression model had DR track, county, income, education as predictors.

[102] The "somewhat likely" and "not likely" categories were condensed and compared with a "very likely" response, which is consistent with how this item was coded in the family satisfaction scale.

[103] See Table N.20 in Appendix N for more details on the logistic regression model. The logistic regression model included DR track and county as predictors. Differences by county were not statistically significant.

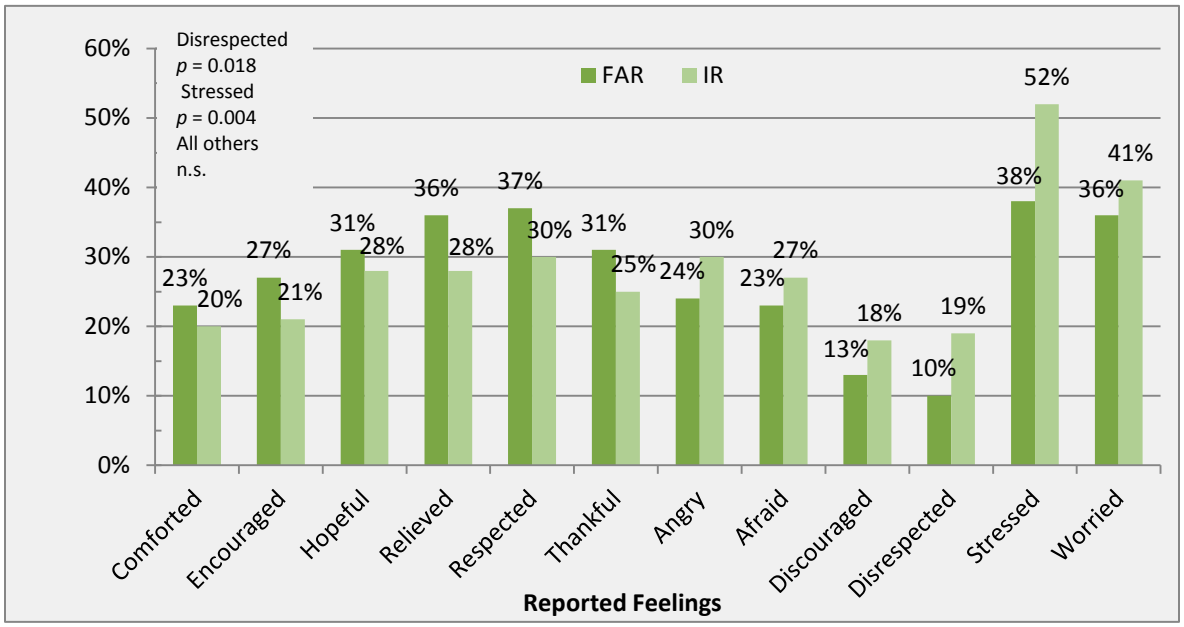
[104] A scale was created to explore differences between the tracks on an overall feelings score. The response options were recoded by assigning positive feelings a "1" and negative feelings a "-1", and a mean score was computed for each respondent to create the feelings scale. Negative feelings include feeling angry, afraid, worried, discouraged, disrespected, and stressed. Positive feelings include feeling comforted, encouraged, hopeful, relieved, respected, and thankful. Higher values on the feelings scale correspond to more positive emotions.

[105] See Table N.21 in Appendix N for more details on the linear regression model. The linear regression model included DR track and county as predictors. Differences among counties were not statistically significant.

The 12 individual feelings at first visit were compared between families assigned to FAR and families assigned to IR using weighted logistic regression models fit to the data.[106] The model revealed statistically significant differences between the tracks on reports of feeling disrespected, $\chi^2(1) = 6.27$, $p = .012$, OR = 0.49, CI[0.28, 0.86].[107] On average, IR respondents had two times the odds of feeling disrespected during their initial caseworker meeting than did FAR respondents. The model revealed statistically significant differences between the tracks on reports of feeling stressed, $\chi^2(1) = 8.53$, $p = .004$, OR = 0.55, CI[0.37, 0.82].[108] On average, IR respondents had 1.8 times the odds of feeling stressed after their initial meeting with the caseworker than did FAR respondents. The model did not reveal statistically significant differences between the tracks on reports of the other 10 feelings.[109] The Figure 5.7 displays weighted percentages of reported feelings by track.

Figure 5.7

Reported Feelings at First Visit by Track



[106] Each reported feeling was used as a dichotomous dependent variable in logistic regression models. Negative feelings were recoded to be consistent with the positive feelings and ease interpretation of the findings. For example, "angry" was modeled as "not feeling angry" with the exception of discouraged and stressed, which were changed for ease of interpretation.

[107] See Table N.22 in Appendix N for more detail on the logistic regression model. The logistic regression model included DR track and county as predictors. Differences among counties were not statistically significant.

[108] See Table N.23 in Appendix N for more detail on the logistic regression model. The model included DR track and county as predictors. Differences by county, ($\chi^2(4) = 14.28$, p = .0064), were statistically significant.

[109] See Tables N.24 – N.33 in Appendix N for more details on the logistic regression models.

5.3.4. Family Involvement

Table 5.11 summarizes caseworker and supervisor responses from the caseworker general survey about perceptions of differences between the FAR and IR tracks in regard to family involvement. The scale was 1 to 5 with 1 being 'much more likely with FAR' to 5 being 'much more likely with IR.' A response of 3 indicated a perception of 'no difference' between FAR and IR. The results indicate that caseworkers and supervisors believe that the FAR approach makes it somewhat more likely that families are present at initial assessment (*M* = 2.1), are cooperative (*M* = 2.3), and participate in decisions and case planning (*M* = 2.3).

Table 5.11

Perceptions of Differences between FAR and IR on Family Involvement (N = 119)

Question	Mean	Median	Min	Max
Family members present at initial assessment	2.1	2.0	1	5
Cooperation of caregivers/family members	2.3	2.0	1	4
Participation in decisions and case plans	2.3	2.0	1	4

The case-specific report was used to answer the question, "are there differences in the provision of support and resources by relatives/friends between families assigned to FAR and IR?"[110] The frequency of moderate/extensive and no/little support from relatives was compared between families assigned to FAR and IR using a weighted multiple logistic regression model fit to the data.[111] The difference between the FAR and IR tracks was statistically significant for receiving relative/friend support at the moderate/extensive versus the no/little level, $\chi^2(1) = 10.07$, $p = .002$, OR = 1.42, CI[1.14, 1.76].[112] This means that, on average, families assigned to the FAR track had 1.4 times the odds of receiving moderate or extensive levels of support from relatives/friends (instead of no or little support) than did families assigned to the

[110] The following items were used to answer this research question: "since the case opened, were relatives and friends outside the household involved in providing needed support and/or assistance to this family' (not at all, very little, moderately, extensively). The family support items were recoded to reflect higher or lower support. Responses of 'not at all' and 'very little' corresponded to 'no/little' support and responses of 'moderately' and 'extensively' corresponded to 'moderate/extensive' support. The actual sample for these analyses was 1,672 families, as no responses were missing.

[111] For the regression model, the independent variables included DR track, county, number of children in the home, number of caregivers, and age of youngest child. Age of youngest child ($\chi^2(1) = 56.65$, $p < .0001$, OR = .91, CI [.89, .93]) and number of children in the home ($\chi^2(1) = 18.76$, $p < .0001$, OR = .79, CI [.71, .88]) were statistically significant for the family support outcome. The odds ratio for age of youngest child in the relatives' support model indicates that, on average, the odds of a family receiving moderate/extensive support (versus no/little support) are about 9% lower for each year increase in the age of the youngest child. The odds of receiving moderate/extensive relatives' support are about 21% lower for each additional child in the family. All other demographic variables were not statistically significant and were dropped from the logistic regression model; county was retained in the model. The actual sample size for this analysis was 1,672 families.

[112] See Table M.24 in Appendix M for details on the logistic regression model.

IR track, according to caseworkers. Table 5.12 displays the weighted frequencies for family support by track.

Table 5.12

Weighted Frequency of Family Support (N = 4,996)

Track	No/Little Relative/Friend Support	Moderate/Extensive Relative/Friend Support	Total
FAR	1,414 (44.3%)	1,780 (55.7%)	3,194 (100.0%)
IR	940 (52.2%)	862 (47.8%)	1,802 (100.0%)
Total	2,353 (47.1%)	2,643 (52.9%)	4,996 (100.0%)

5.3.5. Engagement Characteristics

The case-specific report was used to answer the question, "are there differences in the improvement of engagement between the caseworker and families assigned to FAR or IR?"[113] Five family characteristics were mentioned in the survey: (a) cooperativeness, (b) receptivity to help, (c) engagement with the caseworker, (d) uncooperativeness and (e) difficult behavior. The frequency of improvement for these characteristics was compared between families assigned to FAR and IR.[114] There was a statistically significant difference between FAR and IR families in the odds of improvement for cooperation, $\chi^2(1) = 11.14$, $p < .0001$, OR = 2.13, CI[1.35, 3.35], receptivity to help, $\chi^2(1) = 16.61$, $p < .0001$, OR = 2.36, CI[1.54, 3.61], and engagement, $\chi^2(1) = 25.97$, $p < .0001$, OR = 3.10, CI[1.96, 4.89]. There was a marginally significant difference for difficult behavior, $\chi^2(1) = 3.71$, $p = .054$, OR = 1.70, CI[.98, 2.92]. There was no significant difference between FAR and IR families in the odds of improvement for the 'uncooperativeness'

[113] Caseworkers were asked to 'rate the characteristics of the family members at the first time you met with them' and '…if you met with members of the family more than one time, rate these characteristics the last time you met with them.' The initial-contact and final-contact family characteristics were recoded so that the family was rated 'moderate/extensive' or 'no/little' for all five characteristics. Responses of 'not at all' and 'very little' corresponded to the no/little description of the characteristic and responses of 'moderately' and 'extensively' corresponded to the moderate/extensive level of the characteristic. Then, an *improvement* variable was created for each of the five characteristics. For the three positively-worded characteristics (cooperativeness, receptivity and engagement) the improvement variable was coded 'yes' if the family was rated as 'no/little' at initial contact with the caseworker *and* 'moderate/extensive' at final contact. Any other combination of responses (i.e., 'moderate/extensive' at initial and 'no/little' at final contact *or* no change between initial and final contact) was coded as 'no' for the improvement variable. For the two negatively-worded characteristics (uncooperative and difficult) the improvement variable was coded 'yes' if the family was rated as 'moderate/extensive' at initial contact with the caseworker and 'no/little' at final contact. Any other combination of responses was coded as 'no.' The actual sample size for this analysis was 1,627 families, as 45 responses were missing.

[114] Weighted chi-square frequency analysis was used. All five dependent variables (cooperative, receptive to help, engaged, uncooperative and difficult) could take values of 'no' or 'yes.' For these analyses, the independent variable was FAR versus IR track assignment. Due to the fact that these analyses were chi-square analyses and not generalized linear models, it was not possible to control for county or for the demographic variables number of children in the home, number of caregivers and age of youngest child.

characteristic.[115] On average, FAR families had 2.1 times the odds of showing improvement in cooperation over the life of the case, 2.4 the odds of showing improvement in receptivity to help, 3.1 times the odds of showing improvement in engagement, and 1.7 times the odds of showing improvement (i.e., a reduction) in difficult behaviors than did IR families. These results show that, based on caseworker responses, families assigned to the FAR track experienced increases in several characteristics that demonstrate engagement with CPS.

Table 5.13 summarizes the frequency distribution of the family engagement 'improvement' characteristics. Note also that only 91 cases showed *worse* results between initial and final contact for one or more of the five characteristics. Thus, the vast majority of families classified as not improved for the improvement variable were cases where there was *no change* in behavior between initial and final contact – either positively or negatively. However, less than ten percent of families reported improvement for any of the five family engagement characteristics.

Table 5.13

Weighted Frequency of Improvement for Family Engagement Characteristics (N = 4,889)

Characteristic	Improved	Not Improved	Total
Cooperativeness			
FAR	263 (8.4%)	2,885 (91.6%)	3,148 (100.0%)
IR	72 (4.1%)	1,669 (95.9%)	1,741 (100.0%)
Receptivity to help			
FAR	324 (10.3%)	2,824 (89.7%)	3,148 (100.0%)
IR	81 (4.7%)	1,660 (95.3%)	1,741 (100.0%)
Engagement			
FAR	332 (10.5%)	2,816 (89.5%)	3,148 (100.0%)
IR	64 (3.7%)	1,677 (96.3%)	1,741 (100.0%)
Decreased Uncooperativeness			
FAR	145 (4.6%)	3,003 (95.4%)	3,148 (100.0%)
IR	59 (3.4%)	1,682 (96.6%)	1,741 (100.0%)
Decreased Difficult Behavior			
FAR	164 (5.2%)	2,984 (94.8%)	3,148 (100.0%)
IR	55 (3.2%)	1,686 (96.8%)	1,741 (100.0%)

5.4. Caseworker Satisfaction Outcomes

Caseworker satisfaction was measured by responses to the caseworker general survey and feedback provided during the focus groups by caseworkers and supervisors who worked in

[115] See Table M.25 in Appendix M for details on the weighted chi-square frequency analysis for all family engagement improvement outcomes.

the five counties during the initial and ongoing implementation of DR in Colorado. To explore satisfaction with DR in their counties, supervisors and caseworkers were asked for an overall assessment of their job satisfaction, intention to remain in the field, and opinion of DR almost two years into implementation. They also were asked to reflect on their satisfaction of training, supervision, coaching, workload, services, and the relationship between FAR and IR practice.

5.4.1. Overall Satisfaction

Overall, caseworkers reported moderate levels of job satisfaction and satisfaction with differential response.[116] Table 5.14 summarizes caseworker and supervisor job satisfaction and retention intentions from the caseworker general survey. For these questions, the scale was 1 to 5 with 1 being 'very dissatisfied' and 5 being 'very satisfied.' The mean level of satisfaction was 3.5, which indicates a slightly positive level of satisfaction among the caseworkers and supervisors concerning their child welfare jobs. For retention, the mean was 3.4, which indicates that caseworkers and supervisors were somewhat more likely to stay in the field because of the introduction of FAR.

Table 5.14

Caseworker Job Satisfaction and Retention (N = 119)

Question	Mean	Median	Max	Min
How satisfied are you with your current child welfare job?	3.5	4.0	1	5
Has the introduction of FAR made it any more or less likely that you will remain in this field of work?	3.4	3.0	1	5

Table 5.15 summarizes caseworker and supervisor satisfaction with FAR in their own county based on responses to the caseworker general survey. The scale was 1 to 5 with 1 being 'not at all satisfied' and 5 being 'completely satisfied.' The mean response was 3.5, which indicates that caseworkers and supervisors were somewhat satisfied with FAR in their county.[117]

Table 5.15

Caseworker Satisfaction with FAR (N = 119)

Question	Mean	Median	Max	Min
Overall, how satisfied are you with the FAR program in your county?	3.5	4.0	1	5

[116] These are descriptive results so cannot be generalized beyond the sample of 119 caseworkers and supervisors who responded to the survey.

[117] These are descriptive results so cannot be generalized beyond the sample of 119 caseworkers and supervisors who responded to the survey.

5.4.2. Training Satisfaction

The training during initial and ongoing DR implementation was received with both appreciations and worries. These varying views extended to caseworker's perceptions of the training sequence and scope, application of training, and specific trainings. Caseworkers expressed some worries related to the volume of training, especially first- year training requirements, which were described as "difficult to complete." Although many caseworkers expressed a desire for more training on the DR approach and FAR casework, they also stated that it would be "hard to fit in more" trainings. Thus, one caseworker suggested that, "before we say you need to have all these other trainings, we need to look at the ones we're already requiring." Another caseworker noted that attending these trainings earlier saved time in trying to figure out Trails for FAR cases. Other FAR caseworkers reported that their work on family assessment response began before they received any "official training." Some reported that, due to the lack of formal training beyond the initial assessment, informal training often occurred between caseworkers.

It appears that IR and FAR caseworkers received different training due to their role in the DR project. While IR caseworkers reported receiving the Signs of Safety™ trainings when DR was first implemented, they reported very few augmented trainings, mentoring, or coaching opportunities. One IR caseworker reported feeling "disconnected" from the tools, concepts, and language associated with FAR. Another IR caseworker stated that they "haven't had a lot of support around how to bring this practice to traditional investigations." Interestingly, many FAR caseworkers expressed the opinion that the tools used in the FAR track would be "really beneficial for investigative [work]" and that IR caseworkers should receive more training with regard to the DR approach.

As highlighted in the Year 1 Site Visit Report, caseworkers were generally satisfied with the effectiveness of specific trainings. The Signs of Safety™ training was received positively by caseworkers and was thought to be helpful. Most caseworkers expressed that the practice-focused trainings were more helpful. Caseworkers commented that both the half-day and full-day trainings conducted by state personnel were well-received. However, caseworkers agreed that the full-day training is preferable due to the expanded amount of time to cover the material and to get additional questions answered.

5.4.3. Supervision Satisfaction

Although caseworkers also received individual supervision, they commented mostly about the move toward a group supervision model. While all five counties utilized group supervision, it appears that there were differences with regard to the logistics, participants, and

format of the approach. Some counties held group supervision every other week for a couple of hours, while other counties alternated weeks between individual supervision and group supervision. Either way, most caseworkers were satisfied with group supervision. One caseworker stated, "I think consult [group supervision] is great for getting ideas and next steps." A FAR caseworker said, "I'm learning from everybody else's learning" and appreciated that the group supervision process provides feedback and allows for a "transfer of learning" between caseworkers.

The counties used their RED team process as another venue for group supervision. One caseworker stated, "I was so impressed when I started working here the first time I saw a RED team…the amount of thought that goes into how to assign things [cases] and the amount of thought that goes into a caseworker presenting a case and asking for input and direction is really impressive and really helpful." Supervisors appreciated that RED teams created more acceptance of diversity in perspectives and shared responsibility. Supervisors also reported that caseworker initiative and critical thinking increased with group supervision and RED teams.

Despite the positive feedback, several worries were expressed regarding the group supervision model. One worry was that it provided less ownership of a decision or case direction for caseworkers. Another worry was the "time-consuming" nature of group supervision, especially if a caseworker didn't have a case to present. Caseworkers reported that between group supervision, RED teams, and CPT, they felt like they were in meetings all of the time. In addition, caseworkers reported a lack of management consensus on how to approach cases in a group supervision model. According to caseworkers, the lack of clarity and consensus on how to approach cases negatively impacted timely decision-making, which "slow[ed] the whole process down for families."

5.4.4. Coaching Satisfaction

Overall, caseworkers were satisfied with the coaching they received from supervisors, state staff, and other caseworkers, so much so, that they indicated a need for more coaching opportunities. Some caseworkers reported that there was a "very small amount" of coaching in their counties and that many caseworkers weren't involved in those opportunities. It also seems that coaching activities suffered from the same issue of inconsistency identified with regard to group supervision activities. One caseworker stated, "I think coaching is great; but, again, I think that there is quite a vast array when it comes to what coaching you get, the level, the type, and how it fits and suits who you are…I wouldn't say there's a great deal of consistency." Another caseworker stated, "I think coaching's a great idea. I'd love to see it evolve into something that's a little more powerful than what I received."

Shadowing, particularly the ability to see their practice from somebody else's point of view, was reported as helpful. Numerous caseworkers noted that internal systems of shadowing were developed for training new caseworkers. The main worry was expressed by caseworkers that transferred from IR to FAR within the past year, as there seemed to be a gap in shadowing opportunities, which left some caseworkers feeling uncertain how to adjust their casework practice. Furthermore, it appears that FAR caseworkers typically sought out other caseworkers to answer questions about FAR casework rather than approaching supervisors, who were perceived as having less practice experience with FAR. Others had hoped for more shadowing opportunities with caseworkers more familiar with FAR.

5.4.5. Workload Satisfaction

Overall, caseworkers were less satisfied with their workloads as a result of DR implementation. Caseworkers outlined worries specific to both FAR and IR workloads. The perceived increase in caseload for FAR workers was partly attributed to inexperience with the family assessment response and uncertainty with the new documentation and assessment protocols. Without the randomizer in place to assign FAR eligible assessments to the IR track, all FAR eligible assessments were being assigned to the FAR track, increasing the FAR workload. FAR caseworkers discussed the difficult and exhaustive nature of managing multiple family members at one time during meetings as compared to speaking with family members on an individual basis. FAR caseworkers expressed worries related to new caseworkers having to carry "extremely high" caseloads. As a result, it appears that some new FAR caseworkers did not last long in their positions.

The overwhelming worry expressed by IR caseworkers was that of burnout, due to the intensity of having all high risk assessments assigned exclusively to them. One IR caseworker recounted that previously, "you might have some rough ones, a couple bad ones during the month, but you'd have a good percentage of easy, small mark, low risk [assessments]." IR caseworkers also received emergency assessments–those that necessitate an immediate response–exclusively. IR caseworkers expressed that turnover is a concern due to the strain it puts on remaining IR caseworkers because of the amount of time it takes to hire, replace and train a new IR caseworker.

5.4.6. Service Satisfaction

Table 5.16 summarizes caseworker and supervisors responses from the caseworker general survey about service satisfaction in their communities. For finding services, the scale was 1 to 5 with 1 being 'do not agree' and 5 being 'strongly agree.' The mean was 3.4, which indicates that caseworkers and supervisors could usually find services to help keep children safe

in their home. For working with service providers, the scale was 1 to 5 with 1 being 'do not agree' and 5 being 'strongly agree.' The mean was 3.5, which indicates that caseworkers and supervisors reported that it was somewhat easy to work with service providers in their communities.

Table 5.16

Service Satisfaction (N = 119)

Question	Mean	Median	Max	Min
Can usually find services in my community that can help keep children safe in their home	3.4	3	1	5
It is easy to work with most of the service providers in my community	3.5	4	1	5

5.4.7. Satisfaction with Relationship between FAR and IR

FAR caseworkers were generally satisfied with the relationship between the FAR and IR tracks. Specifically, FAR caseworkers noted a shift in IR practice from an incident-focused approach to assessing whether the incident can be mitigated through services and the family's social supports. FAR caseworkers also observed a shift in concern towards the family and the impact of stressors. From the supervisor perspective, the FAR track increased flexibility and creativity and created a different agency-wide philosophy. Specifically, they perceived a greater focus on family engagement and family-centeredness by all caseworkers.

By contrast, IR caseworkers contend that they still "engage the same way" with families. One IR caseworker stated "my practice really hasn't changed…I was doing it before they put that name on it." Another IR caseworker commented, "I don't think my practice has changed significantly…it basically supports what I've always felt like I needed to do as an investigational worker." Another IR caseworker noted that both FAR and IR caseworkers have the same end goal for the child "to not get hurt again, or not be neglected again, so we're ultimately out there to do the same thing, and we still have to engage the families, we still have to figure out how's this not going to happen again, whether it's incident driven or not." Some IR caseworkers believe that the only differences between the tracks were procedural (e.g., lack of a finding, the initial approach, documentation in Trails) and the tools being utilized (e.g., Signs of Safety™).

5.5. Community Buy-in Outcomes

The relationship between the five counties implementing DR in Colorado and their respective communities was explored in regard to awareness and understanding of DR,

community outreach and training, formal involvement of community stakeholders with DR, and community support and resources.

5.5.1. Awareness and Understanding of DR

In the second year of DR implementation, community stakeholders reported that they learned more about DR, and that explanations from DHS made more sense now than they had previously. Some acknowledged the emotional benefits of working in a manner that is supportive of the family, while others expressed ongoing concerns about whether children were safer in a family assessment response. However, the consensus was to reserve their final opinions until the evaluation for Colorado was completed.

From the supervisor perspective, more community stakeholders were aware of DR during the second year of implementation. They observed that the more the community was involved, the more they understood and accepted the system reform. Relatedly, supervisors believe that community stakeholders with less involvement tended to be more resistant to DR. For example, they attributed concerns from certain law enforcement agencies about the safety of children to them not having enough information about or involvement with DR.

5.5.2. Community Outreach and Training

Community stakeholders involved in the Dependency & Neglect (D&N) court system discussed the outreach provided to their professionals by the counties. One CASA representative said the counties were "very open to sharing information with us regarding the challenges and the successes, and how we're doing at making the practice better, using information that we've been gathering and evaluating." Public health representatives reported that DHS conducted in-services for public health staff regarding mandatory reporting and that their staff attended RED teams regularly. Law enforcement representatives reported that they brought county DHS trainers into their patrol meetings and that they were invited to observe RED team meetings. Head Start representatives reported that DHS held trainings at their site and are "always available to us if we have questions or concerns."

Overwhelmingly, community stakeholders reported taking advantage of the opportunity to attend county-sponsored DR trainings, which highlighted "what the differential response model looks like compared to the traditional investigative model" from the perspective of DHS, families, and partner agencies. School district representatives reported that the trainings were "pretty informative," especially about the evaluation component of the project. Mental health providers encouraged other mental health professionals to attend the trainings "so that everybody's hearing the same thing." Community stakeholders also were given the opportunity to sit in on RED teams to gain a better understanding of how referrals were assigned.

5.5.3. Formal Involvement of Community Stakeholders

The formal involvement of service providers, law enforcement personnel, and judicial representatives was explored during the focus groups.

5.5.3.1. Service providers

Community service providers discussed how DR positively impacted their relationships with DHS, as well as with families. Domestic violence (DV) service providers reported that they were invited to attend home visits with caseworkers where there was suspected DV to offer services to the non-offending parent, and that this practice "seems to have been really helpful." DV service providers appreciated the increased accountability and greater focus on identifying strengths in DV situations related to the DR approach. One service provider commented, "offending parents are being held accountable for their actions." Mental health and substance abuse service providers reported that they saw an increase in referrals for services, and believe the DR approach helped to expedite services to families. Mental health service providers also discussed working more with DHS to connect families to the most effective and appropriate services. They noted that families are often involved with multiple providers; therefore, another role for the agency was to coordinate services to ensure that families were receiving the necessary services.

Head Start representatives reported that they were participating in case planning with families more than they had with the traditional investigative response, and were trying to do more supportive and preventative work. One participant stated, "I think we've been a little bit more informed and involved with FAR than we have been otherwise…that's been good for us to be able to communicate with the staff at DHS and to really work together a little bit more." In some counties, public health representatives did not perceive a significant change to their role or work in child welfare as a result of DR. They stated that service provision was consistent, but that with families in the FAR track, there tended to be more buy-in. In other counties, it was reported that nurses were "more comfortable reporting because they know that they're going to be a part of that process, working with a caseworker…our [Public Health's] relationship has really grown with the caseworkers…we're working much more closely."

5.5.3.2. Law enforcement

Law enforcement representatives reported that, at the time of implementation, they expressed concerns about the conflict between a non-investigative response and the responsibilities of law enforcement agencies. For these representatives, this conflict remained and their overall opinion was unchanged. An illustrative comment was, "it's not up to us to influence the way DHS does their job; but, I think, generally, we're kind of on different ends of the spectrum as far as our appreciation for the FAR process." Another common refrain was

that, “with IR there was a lot more interaction between Human Services and law enforcement. There’s less now.” A particular concern is the usefulness of law enforcement personnel participating in CPT or RED Teams because of a perceived lack of involvement in providing services for the FAR assessments that comprised the bulk of these meetings.

Other law enforcement representatives conveyed a different perspective on the impact of the DR project. They reported that, previously, it has been difficult to get caseworkers to come out when families needed help, but nothing could be done to help legally. However, interactions with caseworkers were much improved. One respondent commented, “I feel like patrol is a lot more satisfied with everybody’s response and the things people are willing to come to the table and talk about”, while another stated, “we have a good team of people and I have no doubt that we’re going to work together.” For these law enforcement professionals, there was a perception of greater consistency in how cases were handled by caseworkers.

5.5.3.3. Judicial representatives

Overall, judicial representatives expressed a differing opinion from their law enforcement colleagues. One D&N court employee stated, “when we first heard about it [DR], there were a lot of questions…are people going to be held accountable…are cases going to be filed…there were a whole bunch of questions about that and, as we see in practice, I believe it probably is making a tremendous positive impact on the families who are a part of it.” Others expressed hope that filings would go down due to the implementation of the DR approach, but do not believe there was such a decrease. One court professional stated, “there are going to be some families, I think, that are just going to make it to the court system regardless of their response from DHS and we’re still seeing those come.” Although the number of cases going to jury trial and/or being adjudicated was about the same, it was noted that hearings were often less adversarial and that the cases going to court were the more difficult types of cases. One GAL commented that the cases being filed on belonged in the court system, more so than they had previously.

5.5.4. Community Support and Resources

The case-specific report was used to answer the question, “are there differences in the provision of support and resources by the community between families assigned to FAR and IR?”[118] The frequency of moderate/extensive and no/little support from the community was

[118] The following item was used to answer this research question: ‘were no-cost neighborhood/community resources (i.e. churches) used to assist this family’ (not at all, very little, moderately, extensively). The family and community support items were recoded to reflect higher or lower support. Responses of ‘not at all’ and ‘very little’ corresponded to ‘no/little’ support and responses of ‘moderately’ and ‘extensively’ corresponded to ‘moderate/extensive’ support. The actual sample for these analyses was 1,672 families.

compared between families assigned to FAR and IR using a weighted multiple logistic regression model fit to the data.[119] The difference between the FAR and IR tracks was statistically significant for receiving higher levels of community support at the moderate/extensive versus the no/little level, $\chi^2(1) = 15.89$, $p < .0001$, OR = 1.72, CI[1.32, 2.24].[120] This means that, on average, families assigned to the FAR track had 1.7 times the odds of receiving relatively higher levels of support from the wider community (instead of no or little support) than did families assigned to the IR track. However, as displayed in Table 5.17, there was no or little community support reported by caseworkers for almost 79% of families.

Table 5.17

Weighted Frequency of Community Support (N = 4,996)

Track	No/Little Community Support	Moderate/Extensive Community Support	Total
FAR	2,424 (75.6%)	770 (24.1%)	3,194 (100.0%)
IR	1,518 (84.2%)	284 (15.8%)	1,802 (100.0%)
Total	3,942 (78.9%)	1,054 (21.1%)	4,996 (100.0%)

[119] For the regression model, the independent variables included DR track, county, number of children in the home, number of caregivers and age of youngest child. Age of youngest child was statistically significant for the community support outcome ($\chi^2(1) = 4.36$, $p = .04$, OR = .97, CI [.94, 1.00]). The odds ratio for age of youngest child in the community support model indicates that, on average, the odds of receiving moderate/extensive community support are about 3% lower for each year increase in the age of the youngest child. All other demographic variables were not statistically significant and were dropped from the logistic regression model.

[120] See Table M.26 in Appendix M for details on the logistic regression model.

6. Cost Study

The cost study was designed to answer the following research question, "are there differences in initial, follow-up, and overall costs between families assigned to FAR and IR?" The cost study also detailed the start-up costs for the project. The cost study design had three main components:

Start-up Costs: This component estimated the level of effort expended for program installation and initial implementation of DR at the county and state levels in Colorado. The start-up costs were divided into recurring and one-time levels of effort to provide counties interested in replication with an estimate of the amount of resources needed to implement DR. The start-up costs included the time period from February 1, 2010 (beginning of grant) to November 30, 2010 (end of "pilot" phase).

Initial Costs: This component estimated case-level costs for families assigned to the FAR track or IR track during their "initial involvement" with the child welfare system in the five counties. The initial involvement represented an uninterrupted span beginning with the initial report, continuing through the assessment (for FAR cases) or investigation (for IR cases), and ending with the ongoing service period (if applicable). Based on days to last family contact, the length of initial involvement was 60 days for FAR and 35 days for IR. The initial period for the FAR cases included any costs associated with the assessment in addition to any ongoing case services provided to the family as long as the agencies' involvement with the family was not interrupted (e.g., case closure). The initial period for the IR cases included any costs associated with the investigation in addition to any costs associated with the opening of a traditional child welfare case to deliver services under a treatment plan. The initial costs included personnel costs associated with caseworker contacts, service costs (recorded in Trails), and OOH placement costs.

Follow-up Costs: This component estimated the case-level costs for families assigned to the FAR track or IR track during the follow-up period with the child welfare system in the five counties. The follow-up period included subsequent involvement with a family for 365 days after the initial involvement period ended. The follow-up costs included personnel costs associated with caseworker contacts, service costs (recorded in Trails), and OOH placement costs.

6.1. Start-up Costs

The start-up time period for the CCDR comprised ten months from February, 2010, through November, 2010. Although the project officially started in February, 2010, when the State of Colorado received the grant funding from the QIC-DR, it actually began months earlier

when the proposal was developed by the state with assistance from the five participating counties and the evaluation team. The level of effort associated with proposal development and initial meetings of state and county staff was not included in the start-up costs. Furthermore, the level of effort associated with the state legislative action required to allow the five counties to participate in the QIC-DR grant and demonstrate the eight core elements outlined by the project was not included. Because this level of effort will not be incurred by other counties if they implement DR, they are not relevant to the start-up costs calculation.

The evaluation team estimated the level of effort of key participants during meetings, presentations, and trainings in addition to grant-related costs for the following activities:[121]

1. Designing and developing the Colorado DR model;
2. Developing and providing communication, outreach, and training to the community and other agencies;
3. Developing the training approach and providing the original training (including the time of presenters and the time of recipients);
4. Updating state and county policies and procedures;
5. Developing forms and procedures for program monitoring;
6. Establishing, developing, and implementing state/county workgroups.

To collect data for the start-up costs component, the evaluation team designed a data collection template to gather information for all eligible start-up activities on the following: date; type of meeting, presentation, or training; number of hours; travel time; and number of personnel. The project director, who was involved with the implementation of DR from the beginning of the start-up period, completed the template. The activities tracked for level of effort were categorized as (1) recurring costs from meetings, presentations, and trainings; or (2) one-time costs from training development and updating policies and procedures. The level of effort for meetings was based on between-county meetings, which were tracked by the PD, but did not include county-specific meetings, which were not tracked. The tracked meetings were for the leadership team and the various workgroups including screening, data/evaluation, intake/services, and family assessment response. The level of effort calculations also included presentations made by the agency to community stakeholders. The trainings tracked were screening trainings, initial site trainings, Signs of Safety™ trainings, and Trails trainings (webinars). It was not possible to calculate the level of effort by type of staff position.

[121] This list is a modification from Appendix A of the QIC-DR cost study guidance document.

Table 6.1 displays the level of effort, as measured in hours and full-time equivalent (FTE), for the five counties in the CCDR based on staff time spent attending meetings and trainings during the start-up period of the project. During the ten months of the start-up period, the five counties spent 1,557 hours (.91 FTE) in meetings, ranging from 422 hours in Larimer to 90 hours in Fremont. The variability in level of effort for meetings was based on several factors, including the number of personnel participating (e.g., Larimer tended to have the most people on the workgroups) and travel (e.g., most of the meetings were in Golden, so Jefferson County employees rarely had to drive, while Garfield staff typically drove six hours round-trip to attend meetings and Fremont often participated by phone).

The five counties spent 13,012 hours (7.57 FTE) in trainings, ranging from 3,945 hours in Arapahoe to 931 hours in Fremont. Again, the variability is mostly related to the size of the county agencies and travel considerations. Overall, the five counties spent 14,569 hours or 8.47 FTE participating in meetings and trainings during the start-up period. The FTE was very consistent among the three larger counties and fairly representative of each county's staff size and child welfare population.

Table 6.1

Level of Effort for Start-up Period

County	Meetings		Trainings		Total	
	Hours	FTE*	Hours	FTE*	Hours	FTE*
Arapahoe	364	.21	3,945	2.29	4,309	2.51
Fremont	90	.05	931	.54	1,021	.59
Garfield	394	.23	1,370	.80	1,764	1.03
Jefferson	287	.17	3,486	2.03	3,773	2.19
Larimer	422	.25	3,280	1.91	3,702	2.15
Total	1,557	.91	13,012	7.57	14,569	8.47
*FTE was calculated based on 1,720 hours for ten-month start-up period						

At the state level, the PD and select staff spent 480 hours (.28 FTE) facilitating and participating in meetings, presentations, and site-visits along with 248 hours (.14 FTE) delivering trainings for a total recurring level of effort of 728 hours or .42 FTE. The PD and staff spent 160 hours (.09 FTE) developing trainings and 80 hours (.05 FTE) updating policies and procedures for a total of 240 hours or .14 FTE one-time level of effort. This one-time level of effort would not be repeated in a statewide expansion of DR, although there would be some time spent on refining the trainings and working with the legislature on rule. Overall, the PD and select state staff had a total level of effort (recurring and one-time) of 968 hours or .56 FTE during the start-up period. It should be noted that the level of effort for the PD and state staff was only for project meetings, presentations, trainings, training development, and updating policies and

procedures; it did not include other work-related functions and responsibilities within CDHS that occurred during the start-up period.

6.2. Cost Study Methodology

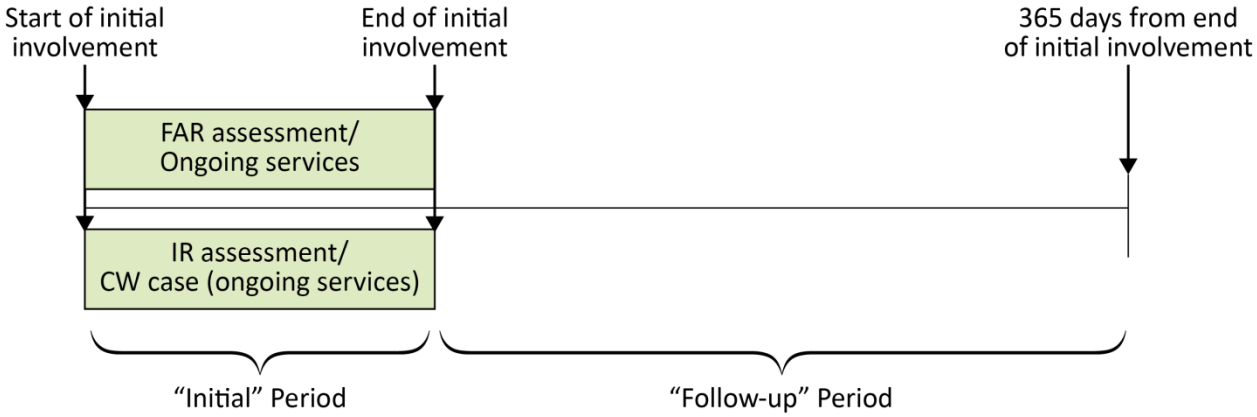
The following describes how the sample and data collection period were defined and obtained and how costs for caseworker contacts, service provision, and out-of-home placement were estimated, calculated, analyzed, and reported for the initial and follow-up time periods.

6.2.1. Cost Sample

The sampling frame has two levels: counties and cases. Originally, a random sampling approach was used to reduce the burden on county financial personnel in regard to collecting non-Colorado Trails related service costs for study families. However, only 7% of the 400 sampled cases had child welfare costs recorded outside of Trails. Thus, it was decided that the cost study analysis would include all 4,996 study cases to allow for greater statistical power in detecting cost differences between the tracks. The cost sample included cases randomly assigned from December 1, 2010, through February 29, 2012, which comprised the complete 15-month study period.[122] Figure 6.1 displays the timeline for the cost study.

Figure 6.1

Cost Study Timeline



[122] There were 89 cases (72 FAR and 17 IR), in which the initial involvement had not ended by the time of cost data collection in April, 2013; however these families were still included in the cost analyses. There were 301 cases (163 FAR and 138 IR) that did not have a complete one-year follow-up period; however these families were still included in the cost analyses. Overall, 390 cases (7.8% of the total sample) either had an open initial involvement or less than a one-year follow-up period at the time of data collection. Of these 390 cases, 60.3% were assigned to the FAR track and 39.7% were assigned to the IR track, which mirrors the overall percentages of the two tracks in the cost sample (63.9% FAR and 36.1% IR).

6.2.2. Cost Data Collection

As decided by the cross-site evaluation team, the cost data collection was limited to child-welfare-related costs associated with caseworker contacts, service provision, and OOH placement. There were no concerns with the completeness, quality, or consistency of these cost data by county or track that would cause problems for the analyses. However, data on contracted service costs paid for outside of child welfare funding (e.g., TANF payments from block grants given to the counties) were not readily available and were not collected. Although there is no reason to believe that these costs would differ between the two tracks, the service costs for both tracks should be considered a conservative estimate without the inclusion of non-child-welfare-related service costs.

Estimating costs for caseworker contacts constituted the biggest challenge for the cost study for several reasons. First, the data field for duration of contacts is not mandated in Trails, so there were no case-specific duration data. However, there were sufficient duration data to allow for an estimation of contact duration.[123] The Trails duration data were also reviewed by an expert panel consisting of caseworkers and supervisors from each of the five counties to ensure its accuracy. Second, the way in which caseworkers record contacts in Trails is somewhat different between the IR and FAR tracks, requiring a correction to the contact data to adjust for this difference.[124] Third, only caseworker time related to direct contact with families was accessible, which did not include time spent on tasks such as documentation. Lastly, the time spent by screeners, RED teams, supervisors, and administrators to manage FAR and IR cases was not included in the estimation of caseworker contact costs. As a result, the caseworker contact costs for both tracks should be considered a conservative estimate. Costs for caseworker contacts were estimated by the following approach:

1. Data on the number of contacts with the family were collected from Trails. Contacts were categorized as either FAR, IR, or traditional child welfare case depending on the status of the case when the contact was recorded.

[123] The duration data were collected from the same time period as the main study, but not exclusively from the five participating counties, which allowed for an adequate sample size to be generated.

[124] In Trails, contacts recorded in traditional investigations only allow for a single participant/client to be added to a single contact record. However, the design of the family assessment response contact in Trails allows multiple participants/clients to be added to a single contact record. In order to account for the discrepancy in the two designs and make the contacts comparable between the two tracks, the evaluation team adjusted contacts to be on a per-participant contact basis. For example, in a IR case in which the caseworker interviewed three children, this would be recorded as three separate contact entries in Trails and would represent three participant/client contacts. In a FAR case, if all three children were interviewed together, it would be recorded as a single contact with the three children as participants/clients, but would actually represent three participant/client contacts for the purposes of the study.

2. Data on the duration of contacts were collected from Trails. Duration data were collected for FAR, IR, and traditional child welfare case contacts.
3. For each case, the number for each type of contact was multiplied by the duration for each type of contact to get an estimate of caseworker time per case.
4. Salary data were collected from the County Financial Management System (CFMS) because these data were not available in Trails. Specifically, the average caseworker salary and benefit cost during the study time period were calculated by each county. A "loaded" hourly rate was computed by dividing salary and benefits costs by the number of hours worked in a year. The same rate was applied to both FAR and IR cases because there was no data element in Trails to indicate whether a caseworker carried FAR or IR cases. Table 6.2 displays the "loaded" hourly rates for caseworkers from the five counties. The hourly rates ranged from a low of $25.40 in Jefferson County to a high of $33.60 in Garfield County.
5. The caseworker contact cost was computed by multiplying the estimated number of hours per case by the "loaded" hourly rate for that county.

Table 6.2

Mean Caseworker "Loaded" Hourly Rate

County	Hourly Rate
Arapahoe	$26.48
Fremont	$27.19
Garfield	$33.60
Jefferson	$25.40
Larimer	$28.02

Cost data for Core Services and out-of-home placements for all individuals in a family were collected from Trails and used in the calculation of initial and follow-up costs. These data were linked by a household ID to allow for analyses at the case level for an overall cost per family.

6.2.3. Cost Data Analysis

Initial, follow-up, and overall costs were analyzed descriptively for caseworker contacts, contracted services, and OOH placement as well as inferentially to compare relative costs between FAR and IR cases. For the analyses, the cost data were weighted to account for county-level differences in the ratio of case assignment to the two tracks. The same weights used for the administrative data in the outcome analyses were also applied for the cost analyses. To ensure that the inferential statistical tests were not in violation of the assumption of normality, the actual costs were given a logarithmic transformation to account for a small

percentage of "outlier" cases which had extremely high OOH placement or service costs.[125] Although these outlier cases are not representative of the experience for most families in the study, they were retained in the cost study because they are actual costs accrued by the counties and a reality in child welfare agencies. The logarithmic transformed values, which were more normally distributed and much less positively skewed, were then entered as dependent variables into a general linear regression model with track, county, and the three demographic variables that varied between FAR and IR families (number of children in the home, number of caregivers in the home, and age of youngest child) serving as covariates. To determine which covariates should be in the final model, the least significant covariate was removed (one at a time) until only statistically significant covariates (if any) remained in the model.

6.3. Initial Costs

The descriptive results for the initial involvement[126] cost analysis are presented by track for: (1) mean duration of contacts; (2) mean weighted contacts per case; (3) mean weighted contact duration per case; (4) mean caseworker hourly rate; and (5) mean weighted contacts cost, service cost, and OOH placement cost per case. These cost comparisons were not tested for statistical significance between the two tracks.

As displayed in Table 6.3, the mean duration for face-to-face contacts were 77 minutes for FAR contacts, 53 minutes for IR contacts, and 89 minutes for CW case contacts. [127] The mean duration for phone contacts were 14 minutes for FAR contacts, 19 minutes for IR contacts, and 15 minutes for CW case contacts.

[125] A constant (half of the smallest non-zero value, in this case 3) was added to each initial, follow-up, and overall cost value and then the natural logarithm of each new value was calculated.

[126] The initial involvement represented an uninterrupted span beginning with the initial report, continuing through the assessment (for FAR cases) or investigation (for IR cases), and ending with the ongoing service period (if applicable).

[127] To obtain a more precise estimate of caseworker time, contacts with families were categorized as either FAR, IR, or traditional CW case. It is possible that a case assigned to FAR or IR could experience any or all of these contact types during the initial involvement period. During the initial involvement for FAR families, an IR contact occurred if the assessment was re-tracked, while a traditional CW case contact occurred if the assessment was transferred to a case during the initial involvement period or a HRA occurred. High risk assessments were categorized as a CW case contact because there were insufficient duration data in Trails for the HRA contact type. During the initial involvement for IR families, a FAR contact occurred if there was a randomizer entry error and the assessment was assigned in Trails to the FAR track, while a traditional CW case contact occurred if the assessment was transferred to a case during the initial involvement period or a HRA occurred.

Table 6.3

Mean Duration (in minutes) for FAR, IR, and Traditional CW Case Face-to-Face and Phone Contacts

Type of Contact	FAR	IR	CW Case
F2F	77	53	89
Phone	14	19	15

As displayed in Table 6.4, the mean weighted initial contacts per case for FAR was 19.7, with 6.2 F2F contacts and 13.5 phone contacts. Initial contacts per FAR case were composed of 15.4 FAR contacts, 0.5 IR contacts, and 3.8 CW case contacts.

Table 6.4

Mean Weighted Initial Contacts per Case for FAR (N = 3,194)

Type of Contact	FAR	IR	CW Case	Total
F2F	5.0	0.2	1.0	6.2
Phone	10.4	0.3	2.8	13.5
Total	15.4	0.5	3.8	19.7

As displayed in Table 6.5, the mean weighted initial contacts per case for IR was 17.1, with 7.4 F2F contacts and 9.7 phone contacts. Initial contacts per IR case were composed of 11.5 IR contacts, 5.2 CW case contacts, and 0.4 FAR contacts.

Table 6.5

Mean Weighted Initial Contacts per Case for IR (N = 1,802)

Type of Contact	FAR	IR	CW Case	Total
F2F	0.1	5.4	1.9	7.4
Phone	0.3	6.1	3.3	9.7
Total	0.4	11.5	5.2	17.1

As displayed in Table 6.6, the mean weighted initial contact duration per FAR case was 679 minutes based on 486 minutes for F2F contacts and 192 minutes for phone contacts. The mean weighted initial contact duration per IR case was 634 minutes based on 464 minutes for F2F contacts and 170 minutes for phone contacts. The mean weighted initial contact duration for each family was multiplied by the caseworker hourly rate for the county in which they were assigned to the study.[128] The mean weighted initial contact costs (rounded to the nearest dollar) were $310 for FAR cases and $284 for IR cases.

[128] The following is an example of how initial contact costs were calculated. Consider a family assigned to the FAR track in Jefferson County that had five F2F and 10 phone contacts during their initial involvement. The five F2F contacts would be multiplied by 77 minutes (see Table 6.3) for a subtotal of 385 minutes and the 10 phone contacts by 14 minutes (see Table 6.3) for a subtotal of 140 minutes. The two subtotals would then be added together for a total initial contact duration of 525 minutes. The total initial contact duration of 525 minutes (8.75 hours) then would be multiplied by the caseworker hourly rate of $25.40 (see Table 6.2) for a caseworker initial contact cost of $222.25 for the family.

Table 6.6

Mean Weighted Initial Contact Duration (in minutes) and Contact Costs per Case for FAR (N = 3,194) and IR (N = 1,802)

Track	F2F Initial Contact Duration	Phone Initial Contact Duration	Total Initial Contact Duration	Initial Contact Costs
FAR	486.3	192.3	678.6	$309.90
IR	464.4	169.7	634.1	$283.87
Average	478.4	184.1	662.6	$300.51

As displayed in Table 6.7, the mean weighted initial service costs per case (rounded to the nearest dollar) were $238 for FAR cases and $157 for IR cases. These costs were calculated for all cases whether they received services during the initial involvement period or not.[129] The mean weighted initial OOH placement costs per case (rounded to the nearest dollar) were $259 for FAR and $99 for IR. These costs were calculated for all cases whether they received OOH placement during the initial involvement period or not.[130]

Table 6.7

Mean Weighted Initial Service and OOH Placement Costs per Case for FAR (N = 3,194) and IR (N = 1,802)

Track	Initial Service Costs	Initial OOH Costs
FAR	$237.53	$259.42
IR	$157.17	$99.37
Average	$208.55	$201.69

6.4. Follow-up Costs

The process for collecting, analyzing, and reporting data on follow-up caseworker contact, service provision, and out-of-home placement costs was the same as the approach used for the initial costs with one exception: the timeframe was a 365-day period after the end of the initial involvement.

[129] A total of 341 FAR cases (10.7%) and 96 IR cases (5.3%) received services paid for in Trails during initial involvement, giving mean weighted initial service costs of $2,219 (FAR) and $3,004 (IR) for these cases.

[130] A total of 52 FAR cases (1.6%) and 16 IR cases (0.9%) received OOH placements paid for in Trails during initial involvement, giving mean weighted initial OOH placement costs of $15,780 (FAR) and $12,089 (IR) for these cases.

As displayed in Table 6.8, the mean weighted follow-up contacts per case for FAR was 9.8 with 3.4 F2F contacts and 6.4 phone contacts. Follow-up contacts per FAR case were comprised of 2.8 FAR contacts, 0.0 IR contacts, and 7.0 CW case contacts.[131]

Table 6.8

Mean Weighted Follow-up Contacts per Case for FAR (N = 3,194)

Type of Contact	FAR	IR	CW Case	Total
F2F	0.9	0.0	2.5	3.4
Phone	1.9	0.0	4.5	6.4
Total	2.8	0.0	7.0	9.8

As displayed in Table 6.9, the mean weighted follow-up contacts per case for IR was 11.9 with 4.1 F2F contacts and 7.8 phone contacts. Follow-up contacts per IR case were comprised of 0.0 FAR contacts, 4.4 IR contacts, and 7.5 CW case contacts.

Table 6.9

Mean Weighted Follow-up Contacts per Case for IR (N = 1,802)

Type of Contact	FAR	IR	CW Case	Total
F2F	0.0	1.9	2.2	4.1
Phone	0.0	2.5	5.3	7.8
Total	0.0	4.4	7.5	11.9

As displayed in Table 6.10, the mean weighted follow-up contact duration per case for FAR cases was 379 minutes based on 284 minutes for F2F contacts and 95 minutes for phone contacts. The mean weighted follow-up contact duration per case for IR cases was 426 minutes based on 299 minutes for F2F contacts and 127 minutes for phone contacts. The mean weighted follow-up contact costs (rounded to the nearest dollar) were $172 for FAR cases and $189 for IR cases.

[131] During the follow-up period for FAR families, an investigation response contact could not occur because FAR cases were required to be assigned to the FAR track provided the subsequent referral was low or moderate risk, while a traditional CW case contact occurred if a subsequent CW case was opened or a high risk assessment occurred. During the follow-up period for IR families, a family assessment response contact could not occur because IR cases were required to be assigned to the IR track provided the subsequent referral was low or moderate risk, while a traditional CW case contact occurred if a subsequent CW case was opened or a high risk assessment occurred.

Table 6.10

Mean Weighted Follow-up Contact Duration (in minutes) and Contact Costs per Case for FAR (N = 3,194) and IR (N = 1,802)

Track	F2F Follow-up Contact Duration	Phone Follow-up Contact Duration	Total Follow-up Contact Duration	Follow-up Contact Costs
FAR	283.8	94.8	378.6	$171.70
IR	299.0	127.5	426.4	$188.68
Average	289.3	106.6	395.9	$177.83

As displayed in Table 6.11, the mean weighted follow-up service costs per case (rounded to the nearest dollar) were $107 for FAR cases and $120 for IR cases. The mean weighted service costs were calculated for all cases whether they received services or not during the follow-up period or not.[132] The mean weighted follow-up OOH placement costs per case (rounded to the nearest dollar) were $127 for FAR and $104 for IR. The mean weighted OOH placement costs were calculated for all cases whether they received OOH placement during the follow-up period or not.[133]

Table 6.11

Follow-up Mean Weighted Service and OOH Placement Costs per Case for FAR (N = 3,194) and IR (N = 1,802)

Track	Follow-up Service Costs	Follow-up OOH Costs
FAR	$106.78	$126.64
IR	$120.32	$104.37
Average	$111.66	$118.61

6.5. Overall Costs

Overall costs for caseworker contacts, service provision, and out-of-home placements, are reported with descriptive statistics and compared between the two tracks with inferential statistics.

[132] A total of 127 FAR cases (4.0%) and 73 IR cases (4.1%) received services paid for in Trails during follow-up, producing mean weighted service costs for cases that received services during follow-up of $2,651 (FAR) and $3,036 (IR).

[133] A total of 44 FAR cases (1.4%) and 16 IR cases (0.9%) received OOH placement paid for in Trails during follow-up, producing a mean weighted service costs for cases that received OOH placement during follow-up of $9,088 (FAR) and $7,445 (IR).

6.5.1. Descriptive Cost Results

As displayed in Table 6.12, the overall mean weighted costs per case (rounded to the nearest dollar) for FAR cases were $1,212 based on $807 for overall initial costs[134] and $405 for overall follow-up costs.[135] The overall mean weighted costs per case (rounded to the nearest dollar) for IR cases were $954 based on $540 for overall initial costs and $413 for overall follow-up costs. It should be noted again that these mean weighted costs include "outlier" cases which had extreme service or OOH placement costs. An alternative descriptive statistic for skewed data is the median. For overall initial costs, the median was $167 for FAR and $165 for IR. For overall follow-up costs, the median was $0 for FAR and $0 for IR (indicating that more than half of all cases did not have any follow-up costs). For overall costs, the median was $199 for FAR and $199 for IR.

Table 6.12

Mean Weighted Initial, Follow-up, and Overall Costs per Case for FAR (N = 3,194) and IR (N = 1,802)

Track	Initial Costs	Follow-up Costs	Overall Costs
FAR	$806.85	$405.12	$1,211.97
IR	$540.41	$413.37	$953.78
Average	$710.75	$408.10	$1,118.85

6.5.2. Comparative Cost Results

The transformed initial, follow-up, and overall costs were compared between families assigned to FAR and families assigned to IR. A weighted general linear regression model was fit to the data, where the natural logarithm of initial, follow-up, and overall costs were the dependent variables. Independent variables included track and county, and covariates included number of children in the home, number of caregivers in the home, and age of youngest child.

For the initial cost analysis, the number of children in the home was statistically significant, $t(4995) = 5.98$, $p < .001$, $\beta = 0.092$ and CI [0.062, 0.122]. There was a positive relationship between the number of children and initial costs, which indicates that as the number of children in the home increased, the amount of initial costs increased. The number of caregivers in the home was statistically significant, $t(4995) = 2.52$, $p = .012$, $\beta = 0.081$ and CI [0.018, 0.145]. There was a positive relationship between the number of caregivers and initial

[134] The overall initial costs were calculated by summing the initial contact costs (from Table 6.6) with the initial service costs and initial OOH placement costs (from Table 6.7).

[135] The overall follow-up costs were calculated by summing the follow-up contact costs (from Table 6.10) with the follow-up service costs and follow-up OOH placement costs (from Table 6.11).

costs, which indicates that as the number of caregivers in the home increased, the amount of initial cost increased. The age of youngest child was not a statistically significant demographic variable and was dropped from the model. The difference between counties was statistically significant, $\chi^2(4) = 79.69$, $p < .0001$.

The difference between the tracks was not statistically significant, $t(4995) = 1.46$, $p = .144$ $\beta_{IR} = 5.36$, CI[5.22, 5.50], $\beta_{FAR} = 5.41$ and CI[5.21, 5.60].[136] Thus, there was no difference between the two tracks on initial costs.

For the follow-up cost analysis, the youngest child age was statistically significant, $t(4995) = 5.50$, $p < .001$, $\beta = -0.035$ and CI [-0.048, -0.023]. There was a negative relationship between the age of the youngest child and follow-up costs, which indicates that as the age of the youngest child decreased, follow-up costs increased. The number of children in the home was statistically significant, $t(4995) = 3.67$, $p < .001$, $\beta = 0.126$ and CI [0.058, 0.193]. There was a positive relationship between the number of children and follow-up costs, which indicates that as the number of children in the home increased, the amount of follow-up costs increased. The number of caregivers in the home was not a statistically significant demographic variable and was dropped from the model. The difference between counties was statistically significant, $\chi^2(4) = 6.86$, $p < .0001$.

The difference between the tracks was statistically significant, $t(4995) = 5.18$, $p < .001$, $\beta_{IR} = 2.59$, CI[2.34, 2.85], $\beta_{FAR} = 2.23$ and CI[1.84, 2.62].[137] Thus, the follow-up costs for IR cases were significantly higher than the total follow-up costs for FAR cases. Because of the nature of the data for the cost study,[138] it is difficult to provide a quantitative effect size measure to indicate the magnitude of the difference between tracks on follow-up costs. However, the relative size of the t statistic ($t = 5.18$) indicates a substantial difference between tracks on follow-up costs, which implies that there may be practical significance to this finding.

For the overall costs analysis, the number of children in the home was statistically significant, $t(4995) = 6.64$, $p < .001$, $\beta = 0.127$ and CI [0.090, 0.165]. There was a positive relationship between the number of children and overall costs, which indicates that as the

[136] See Table Q.1 in Appendix Q for details on the regression model.

[137] See Table Q.2 in Appendix Q for details on the regression model.

[138] The cost data were weighted and transformed for this analysis. The regression model yielded estimated marginal means based on the transformed values of the cost data. These means could not be back-transformed into dollar amounts because too many cases had a zero value for follow-up costs. Furthermore, the model yielded estimated standard errors which could not easily be changed to standard deviations (for a Cohen's d effect size calculation) because of the data being weighted.

number of children in the home increased, the amount of overall costs increased. The youngest child's age was statistically significant, $t(4995) = 2.61$, $p = .009$, $\beta = -0.010$ and CI [-0.018, -0.003]. There was a negative relationship between the age of the youngest child and overall costs, which indicates that as the age of the youngest child decreased, overall costs increased. The number of caregivers in the home was not a statistically significant demographic variable and was dropped from the model. The difference between counties was statistically significant, $\chi^2(4) = 62.14$, $p < .0001$.

The difference between the tracks was not statistically significant, $t(4995) = 0.51$, $p = .611$ $\beta_{IR} = 5.86$, CI[5.73, 6.00], $\beta_{FAR} = 5.84$ and CI[5.63, 6.05].[139] Thus, there was no difference between the two tracks on overall costs.

[139] See Table Q.3 in Appendix Q for details on the regression model.

7. Discussion

This chapter summarizes the findings from the outcome, process, and cost evaluations, identifies limitations of the methods, data collection procedures, and data analysis techniques, provides implications for the practice and policy of DR in Colorado, and offers ideas for future research and evaluation of differential response.

7.1. Summary of Findings

Overall, the main findings from the outcome, process, and cost evaluations were positive in regard to child safety, family well-being, family engagement, caseworker satisfaction, community buy-in, and cost neutrality. The most promising finding was that there may be long-term child safety benefits and cost savings due to lower levels of re-involvement, over time, for families that receive a family assessment response. The lack of a statistically significant finding for the short-term safety outcomes was not surprising given that the DR system reform also enhanced traditional CPS practices in Colorado.

7.1.1. Child Safety

Based on administrative data findings, the outcome evaluation found no significant differences between the FAR and IR tracks on all of the safety outcomes examined. This includes referral within 365 days of initial referral, assessment within 365 days of initial referral, HRA within 365 days of initial referral, founded HRA within 365 days of initial referral, traditional child welfare case opening after initial involvement, and placement in out-of-home care after initial involvement. However, survival analysis findings suggest that FAR families were 18% less likely to have an HRA, over time, than were IR families. Overall, the child safety findings suggest that family assessment response is as effective in preventing harm as a traditional investigation response for low- and moderate-risk cases. The family exit survey results support the findings from the administrative data, as there were no differences in perceived improvements in child safety between FAR and IR families as reported by caseworkers.

The vast majority of study families had a low number of risk and safety concerns and limited prior involvement with the child welfare system. However, there were statistically significant associations found between the outcomes for families in both tracks and several risk factors. As displayed in Table 7.1, prior referral within 24 months of the initial referral was significantly related to five of the six outcomes, as families with a prior referral had a greater probability of being re-involved. The total number of prior referrals was significantly related to three outcomes, as families with more prior referrals had a greater probability of being re-

involved. Similarly, the number of prior referrals within 24 months of the initial referral was significantly related to two outcomes, as families with more prior referrals within 24 months had a greater probability of being re-involved. Mental health risk was significantly related to four outcomes, as families with a mental health risk had a greater probability of being re-involved. Risk total score was significantly related to five outcomes, as families with higher risk total scores had a greater probability of being re-involved.

Table 7.1

Summary of Relationships between Family and Case Characteristics and Safety Outcomes

Family/Case Characteristics	Within 365 Days of Initial Referral				After Initial Involvement	
	Referral	Assessment	HRA	Founded HRA	CW Case Opened	OOH Placement
Track	--	--	--	--	--	--
County	✓	✓	✓	✓	✓	--
Ethnicity of Youngest Child	✓	✓	✓	✓	✓	✓
Age of Youngest Child	--	✓	--	✓	✓	--
Caregiver Age	--	✓	--	--	--	--
Number of Children in the Home	✓	--	✓	--	--	--
Risk Level	--	--	--	--	--	--
Domestic Violence Risk	✓	--	--	--	--	--
Mental Health Risk	--	✓	✓	--	✓	✓
Substance Abuse Risk	--	--	--	--	--	--
Abuse Allegation	--	--	--	--	--	--
Neglect Allegation	--	--	--	✓	--	--
Risk Abuse Score	--	--	--	--	--	--
Risk Neglect Score	--	--	--	--	--	--
Risk Total Score	✓	✓	✓	✓	✓	--
Number of Safety Concerns	--	--	--	--	--	--
Prior Child Welfare Case	--	✓	--	--	--	--
Prior Referral in Past 24 months	✓	✓	✓	✓	✓	--
Prior Assessment in Past 24 Months	--	--	--	--	--	--
Number of Prior Referrals Past 24 Mo.	✓	✓	--	--	--	--
Number of Prior Assess. Past 24 Mo.	--	--	--	--	--	✓
Number of Prior Referrals (total)	✓	--	--	--	✓	✓
Number of Prior Assessments (total)	--	--	--	--	--	--

✓ indicates a statistically significant relationship

There were several family and case characteristics associated with improved safety outcomes. For example, the age of the youngest child in the home was negatively related to the probability of a subsequent assessment, HRA, and traditional CW case opening after initial involvement. Furthermore, families with an older caregiver were associated with a lower probability of a subsequent assessment. The ethnicity of the youngest child was significantly related to all six safety outcomes, as families with a youngest child of "other/unknown" ethnicity were less likely to be re-involved. In addition, the number of children in the home was related to several outcomes, although the findings were mixed for this characteristic.

Lastly, parent education level (as measured by the family exit survey) was related to several safety, well-being, and family engagement outcomes. Interestingly, as self-reported parent education level increased, the following decreased:

1. Perceived improvements in child safety
2. Perceived improvements in well-being as a result of CPS involvement
3. Ratings of caseworkers on the family-centered practice scale

Thus, it appears that parents with more education had a less positive perspective about the impact of DR than did parents with less education. This could be related to expectations about child welfare involvement, in that parents with a lower education level may be more likely to want (and need) services and resources from CPS than parents with a higher education level.

The comparability of safety outcomes between the two tracks can be interpreted in several ways. First, the expectation from some community stakeholders and child welfare practitioners in the five counties was that family assessment response would result in children who were less safe because of the different approach and the lack of a finding resulting from the non-investigatory track. This expectation was not realized and may indicate that the finding of fault for an allegation of abuse or neglect, on its own, is not an essential component for keeping children safe. Second, other community stakeholders and child welfare practitioners expected that children would be safer because of the higher levels of family engagement, service provision, and caseworker satisfaction that have defined FAR in other states. This too, at least in the short-term, was not realized and may indicate that DR is more of a system reform that impacts all aspects of child welfare practice and provides similar benefits to children, youth, and families no matter the track to which they are assigned.

7.1.2. Family Well-being

According to caseworker perceptions, FAR families were more likely to have material needs and mental health needs met than were IR families. However, there were no statistically significant differences between the tracks in meeting substance abuse needs, physical health

needs, parenting skills needs, domestic violence needs, educational needs or social support needs. As reported by caseworkers, there were no statistically significant differences between the tracks in the improvement of family needs, given that a specified need was met. This is not surprising given that once services are provided for an identified need, there should not be an expectation that they would work better for FAR families than for IR families. These findings were generally supported by results from the family exit survey, as there was no reported difference between the tracks on whether global family needs were met satisfactorily. When parents were asked directly about improvement in family well-being as a result of their involvement with CPS, again there was no difference between the tracks. Although this is somewhat surprising given the other findings from the study, it may be explained by the notion that parents from low- and moderate-risk cases (regardless of track assignment) already perceive their families to have a high level of well-being before CPS involvement.

7.1.3. Family Engagement

Overall, the findings suggest that FAR makes a difference in engaging families and enhancing their experience with CPS. Based on findings from the family exit survey, FAR families reported feeling more engaged than did IR families. For example, FAR families were more likely to score their caseworkers high on demonstration of family-centered practice skills than were IR families. Relatedly, FAR families were more likely to rate satisfaction with their caseworker as high. Lastly and maybe most telling, FAR families reported being more likely to call CPS in the future than did IR families.

These findings are aligned with the results from the case-specific report. Caseworkers reported that, compared to IR families, FAR families were more likely to show improvement in cooperation, receptivity to help, engagement, and a reduction in difficult behaviors. Furthermore, caseworkers observed that FAR families were more likely to receive higher levels of support from friends and relatives. This may be related to the emphasis in FAR training on involving family members in the resolution of child maltreatment issues (e.g., access to family meetings). As reported in the caseworker general survey, caseworkers and supervisors perceive that FAR families are somewhat more likely to cooperate, attend the initial assessment, and participate in decisions and case planning.

7.1.4. Caseworker Satisfaction

Based on the findings from the focus groups conducted during the two site visits, caseworkers and supervisors reported satisfaction with both the initial and ongoing implementation of differential response in their counties and were generally very positive about the Colorado DR model and the system change it engendered. Specifically, there was

satisfaction regarding training, supervision, coaching, services, and the relationship between FAR and IR practice. Caseworkers found the practice-focused trainings and Signs of Safety training to be helpful, and reported that the half-day and full-day trainings conducted by state personnel were well-received. Caseworkers appreciated group supervision for the valuable insight provided by diverse perspectives and peer-validated decisions. RED teams were viewed very positively in terms of coming to consensus on eligibility decisions in an inclusive and dynamic way and for allowing a transfer of learning. Caseworkers who received coaching from supervisors, state staff, or other caseworkers found it to be a positive experience. Furthermore, shadowing, particularly the feedback provided to caseworkers and the ability to see their practice from somebody else's point of view, was reported as helpful. From the supervisor perspective, the FAR track increased flexibility and creativity and created a different agency-wide philosophy with a greater focus on family engagement.

These qualitative findings are supported by results from the caseworker general survey, which indicated that caseworkers and supervisors were somewhat satisfied with FAR in their counties. Caseworkers and supervisors also reported a slightly positive level of satisfaction concerning their child welfare jobs and reported that they are somewhat more likely to stay in the field because of the introduction of DR in Colorado. Additionally, caseworkers and supervisors reported that they were able to find services to help keep children safe in their home and found that it was somewhat easy to work with service providers.

Although the overall level of satisfaction was high, caseworkers did express some worries about training, supervision, coaching, and workload. Similar to the first year of implementation, the sequence and scope of training made it difficult to integrate and apply the knowledge gained, especially for IR caseworkers who were perceived to receive less practice specific training content. The biggest challenges for the group supervision model were the logistics, participants, format, consistency, and amount of time required. While caseworkers found coaching and shadowing opportunities to be particularly helpful in examining their own practice and receiving input from supervisors and other caseworkers, the main challenge during the second year of implementation was providing sufficient opportunities for formal coaching. After random assignment ended, there was a perceived increase in workload for FAR caseworkers, as more assessments were being assigned to FAR, and the optimal staff and caseload size had not yet been achieved. The overwhelming worry expressed by IR caseworkers was that of burnout, due to the high intensity of the cases being assigned exclusively to them. This was supported by findings from the caseworker general survey, in that caseworkers reported that FAR had slightly increased their caseload size, overall workload, and paperwork.

7.1.5. Community Buy-in

Overall, community stakeholders reported enhanced connections with DHS, which established new avenues of working together to serve children and families within a DR system. Many stakeholders reported having tried to establish a connection in years past to no avail, and were "astonished" by the new connections they made with the county agencies. However, the biggest challenge of DR implementation in Colorado was for the counties and state to engage community stakeholders in the system reform. For example, there was a general lack of community awareness of the DR project and DR in general during the initial implementation period. This led to the most demonstrable change observed during the project, as the concentrated efforts to generate community buy-in for DR were perceived to be effective. Specifically, the counties engaged community stakeholders in training, especially through participation in CPT and RED teams. Agency partners appreciated the transparency and informative nature of the training opportunities. Another indicator of community buy-in was that families assigned to the FAR track were more likely than IR families to receive higher levels of support from the wider community as reported by caseworkers. That being said, the formal involvement of law enforcement, the courts, mandated reporters, and service providers in differential response is still evolving and faces challenges regarding roles, responsibilities, and expectations.

7.1.6. Contacts and Services

Based on findings from the case-specific report, families assigned to the FAR track received significantly more face-to-face and phone contacts, on average, than did families assigned to the IR track. This finding is supported by the administrative data used in the cost study, which indicated significantly more contacts during the initial involvement period for families in the FAR track. However, family exit survey findings contradict this, as families reported no significant differences between FAR and IR on number of face-to-face contacts. This could be attributed to a lack of recall on the part of parents about exact numbers of contacts received based on the lag time (anywhere from 2-6 weeks) between the case closure and the administration of the family exit survey. In addition, the definition of a "contact" likely differs between caseworkers and parents.

According to the case-specific report, families assigned to the FAR track were more likely to receive a service referral and an actual service than were families assigned to the IR track. Caseworkers also reported that families assigned to the FAR track were more likely to receive material needs and mental health services than were families assigned to the IR track. Although FAR families were more likely to report higher satisfaction with services received as compared to IR families, the results from the family exit survey indicate no differences between the tracks

in the reported number of services received or the type of services received. Again, this could be attributed to a lack of recall on the part of parents about exact numbers and nature of services received based on the lag time between the case closure and the administration of the family exit survey. In addition, the definition of a "service" likely differs between caseworkers and parents.

Based on findings from the case-specific report, families assigned to the FAR track were less likely to receive services within three weeks, on average, then were families assigned to the IR track. It is important to note that the number of families who experienced a delay in service provision was relatively small. According to caseworker perceptions, there was no difference between the tracks for level of service match or for level of service effectiveness. However, caseworkers and supervisors reported in the caseworker general survey that FAR families were somewhat more likely to receive services they needed and receive them more quickly. Referrals to other community services were also more likely for FAR families, according to caseworkers. A key takeaway is that a majority of families from both tracks reported receiving no services.

7.1.7. Costs

There was no statistically significant difference between the two tracks on initial costs. The initial mean weighted service cost per case for FAR cases was $807 with $238 for service costs, $259 for out-of-home (OOH) placement costs, and $310 for caseworker contact costs. The initial mean weighted service cost per case for IR cases was $540 with $157 for service costs, $99 for OOH placement costs, and $284 for caseworker contact costs.

The follow-up costs for IR cases were significantly higher than the follow-up costs for FAR cases. The follow-up mean weighted service cost per case for FAR cases was $405 with $107 for service costs, $127 for OOH placement costs, and $172 for caseworker contact costs. The follow-up mean weighted service cost per case for IR cases was $413 with $120 for service costs, $104 for OOH placement costs, and $189 for caseworker contact costs.

There was no significant difference between the two tracks on overall costs. The overall mean weighted cost per case for FAR cases was $1,212 with $807 for initial costs and $405 for follow-up costs. The overall mean weighted cost per case for IR cases was $954 with $540 for initial costs and $413 for follow-up costs. It is important to note that these mean weighted costs include outlier cases for which the service or OOH placement costs were extremely high, thus skewing the mean costs. Such cases impacted the costs for the FAR track more than the IR track.

7.2. Limitations

The limitations for the outcome, process, and cost evaluations are discussed to provide context for the key findings and study implications and lay a foundation for future research and evaluation in this area.

7.2.1. Outcome Evaluation

This project had many strengths, including the implementation of a randomized controlled trial that allowed for a comparison of outcomes between two equivalent groups; an evaluation that examined the processes, outcomes, and costs of a DR practice model; the integration of multiple data collection methods including surveys, focus groups, interviews, and administrative data; and the utilization of appropriate data analyses techniques such as controlling for sampling bias, non-response bias, pre-existing group differences, and outliers through weighting, data transformation, and survival analyses. However, the outcome evaluation was challenged by several limitations common to child welfare research, such as a reluctance to randomly assign caseworkers to cases, ability to generalize findings to other counties or states, utilization of non-validated instruments, and measurement error including non-response bias.

The evaluation team accounted for the influences of county, case, and family characteristics in examining safety outcomes. However, they could not account for the influence of individual caseworker practices on these outcomes, as caseworkers were not randomly assigned to serve FAR or IR cases. Some counties had dedicated FAR and dedicated IR caseworkers to exclusively handle cases in their respective pathways while other counties had caseworkers who had cases in both pathways. Caseworkers often vary in their skill level and experience, which could be a contributing factor to differences in outcomes between FAR and IR families.

This study was limited to five of the 64 counties in Colorado. Although participating counties varied in size, population density, and resident demographics, the study findings may not be generalizable to other counties in the state due to other contextual differences between participating and non-participating counties.

To capture the perspectives of the caseworkers and families involved in the study, the evaluation team used survey instruments employed in previous DR studies. However, these instruments have not been validated. The retrospective nature of the survey procedures could have resulted in some inaccurate responses. This may account for the discrepancy between the family exit data and caseworker case-specific survey results in regard to the number of face-to-face contacts and services each track was reported to have received. Furthermore, the

response rate for the family satisfaction survey was low, requiring the team to weight the data to adjust for survey non-response.

Lastly, the counties in the CCDR mutually agreed early on that the differential response implementation would be utilized to reform multiple parts of the child welfare system in their counties. Thus, numerous programmatic, procedural, administrative, supervisory, and practice reforms occurred in conjunction with the installation of two distinct tracks to respond to referrals of child abuse and neglect. As a result, it was somewhat difficult to isolate the relationship between the Colorado DR model and safety outcomes. That being said, the rigor of the RCT was useful in controlling for confounding variables, as it was assumed that these system reforms would equally impact practice in both the FAR and IR tracks. As a result, it is likely that the RCT was able to isolate the main difference between the tracks (i.e., no finding in FAR assessments and findings in IR assessments), which would allow for meaningful conclusions to be drawn regarding the safety, permanency, and well-being of children in the study.

7.2.2. Process Evaluation

The process evaluation had several limitations regarding the selection of participants, the scope of the site visits, and the assessment of implementation fidelity. Focus group sampling was purposive and convenient, and not randomized, which could have yielded a sample of individuals more supportive of DR. However, the results show a clear continuum of opinion on the practice. Another limitation was that families were not interviewed for the process evaluation. Although families were sampled for the family exit survey, they did not have the opportunity to share their opinions in a more open-ended way.

The two rounds of site visits allowed participants to reflect on the initial stages of DR implementation in Colorado. However, a third site visit in the last year was planned but not implemented because sufficient data for the process evaluation were collected during the first two rounds. As a result, the later stages of DR implementation were not captured through qualitative interviewing. However, ongoing conversations with state staff and county representatives suggest that attitudes had not significantly changed, as caseworkers and supervisors remained supportive of the system change with community stakeholders–with some exceptions–increasingly warming up to DR.

The implementation fidelity assessment for the FAR track also did not occur as planned due to resource and time limitations. Furthermore, the evaluation team did not explore whether IR caseworkers implemented the investigation response with fidelity. However, the fidelity matrix presented in Appendix L could serve as a valuable tool for the ongoing evaluation of DR in Colorado as the practice is adopted by new counties.

7.2.3. Cost Evaluation

There were several limitations to calculating start-up costs, estimating caseworker time, and collecting service cost data. The start-up effort for the CCDR may appear to be more costly and labor-intensive than the other demonstration projects because of the lack of a foundation for developing and implementing the DR model prior to the grant. The start-up costs were collected by the project director for all trainings and meetings held at the state level. However, county-specific trainings and meetings that occurred during the start-up period were not collected, which led to an underestimation of the level of effort for county staff.

Estimating caseworker time was the cost study's biggest challenge for several reasons. The data field for duration of contacts is not mandated in Trails, so there were not case-specific duration data. There were sufficient duration data to allow for an estimation of contact duration. However, data on the duration of caseworker time for documentation were not available and were not included in the calculation of caseworker time. Based on findings from the caseworker general survey, there may be a difference in the amount of time needed to document a FAR assessment as compared with an IR assessment. Furthermore, the time spent by screeners, RED teams, supervisors, and administrators to manage FAR and IR cases was not included in the estimation of caseworker time, although no difference between the tracks was expected. One limitation to the collection of service costs is that data on contracted service costs paid for outside of child welfare funding were not collected.

7.3. Implications

The following practice and policy implications are based on the following key takeaways from the evaluation. First, FAR does no harm to the safety, well-being, and engagement of children and families. Second, a majority of child welfare practitioners, service providers, and community stakeholders were increasingly satisfied with and invested in the Colorado DR model and the DR system reform. Third, there were more caseworker contacts and services provided in a family assessment response at equally or potentially lesser costs.

7.3.1. Policy and Practice

Based on the totality of findings from the outcome, process, and cost evaluation, there was clear balance between the three main focus areas for DR in the five Colorado counties: (1) child safety, child well-being, and family engagement; (2) caseworker satisfaction and community buy-in; and (3) cost neutrality. The main policy implication is that there is evidence to support the expansion of FAR into new counties in Colorado that have sufficient staffing availability, service capacity, and community readiness. This is not to imply that families assigned to the IR track did not achieve good outcomes or that there is a lack of caseworker and

community support for an investigatory response. It is also important to stress that if the eligibility criteria for family assessment response changes drastically, these findings may not apply and further evaluation would be necessary.

Five primary practice implications emerged from the study findings. The first is for each county to seek out consensus on the eligibility criteria for family assessment response. The development of the Agency Response Guide yielded more sound eligibility decisions, which was evidenced by the three percent re-track rate for the five counties. This indicates that almost all families were appropriately assigned to FAR based on their level of risk, and that families presenting with higher risk were determined to not be eligible for FAR and thus were assigned to receive a high risk assessment. Furthermore, the RED team approach was viewed very positively by caseworkers and supervisors in terms of making eligibility decisions in an inclusive and dynamic way. However, some caseworkers and stakeholders continue to have reservations about certain types of assessments being assigned to FAR, urging ongoing review of the eligibility criteria. These concerns center on cases with domestic violence, certain types of substance abuse, and prior involvement with the child welfare system. Stakeholders also call for more consistency and transparency in the assignment of assessments, including the re-tracking of assessments.

The second implication is to continue enhanced screening by refining the information gathered at this decision point. Most community stakeholders supported enhanced screening and want to see even more information collected at the time of referral. The downside of the new screening procedures was the resistance of some mandatory reporters regarding the increased time and information required to make a referral. The effectiveness of the enhanced screening and RED team process is illustrated in the pre-FAR and post-FAR referral acceptance rates displayed in Table 7.2. Specifically, the pre-FAR referral acceptance rates for the five CCDR counties ranged from 42% in Larimer to 67% in Garfield with a pre-FAR overall acceptance rate of 52%. The post-FAR referral acceptance rates declined in all five counties with a range of 35% in Larimer to 55% in Garfield with a post-FAR overall acceptance rate of 44%. There was some concern that DR would result in an influx of low- and moderate-risk families who would have otherwise been screened-out. This did not occur in Colorado as the referral acceptance rate actually decreased for the five DR counties.

Table 7.2

Pre-FAR and Post-FAR Referral Acceptance Rates for CCDR Counties

Time Period	Arapahoe	Fremont	Garfield	Jefferson	Larimer	Overall
Pre-FAR (10/2008 – 9/2010)	54%	54%	67%	57%	42%	52%
Post-FAR (10/2010 – 9/2012)	42%	49%	55%	50%	35%	44%

The third practice implication is for counties to pay special attention to managing caseloads and workloads for both IR and FAR caseworkers. For IR caseworkers, there should be a concerted effort to provide them with additional supports to alleviate some of the burden around such activities as court and facilitated family meetings. In addition, supports are needed to address the secondary traumatic stress associated with working emergency and high-risk cases on a daily basis. For FAR caseworkers, there is a need for new functionality in Trails to assist them in the timely documentation of FAR practice, including track change. Although group supervision was viewed positively by both FAR and IR caseworkers, there is a need for more consistency in how it is delivered and more efficiency in the time required, as there should be a balance between receiving feedback in a group setting and attending to caseload and individual performance requirements. Furthermore, the supervisor role needs additional clarity especially as it relates to performance monitoring. These recommendations have implications for county performance measurement, caseworker retention, and continuity in the implementation of family assessment response.

The fourth practice implication is for more integrated and seamless service delivery between county DHS and community providers. Based on findings from the site-visit reports, there was a gap in available and accessible resources and services, which may result in disparities in the provision and utilization across counties and tracks. To keep up with the increasing demand for family-centered services, the state and counties must redouble efforts to increase service capacity by identifying external service providers and enhancing internal service offerings. There also is a need for greater consistency in service plans, more follow-up on service referrals, and better communication between county DHS and providers.

The final practice implication is to clearly define what is different and what is the same about FAR and HRA casework, roles, and responsibilities. For example, regardless of assigned track, caseworkers try to engage with the family and always look for strengths and resources. Perhaps one of the emergent practice shifts for caseworkers in both tracks is the adoption of techniques promoted by the implementation of solution-focused assessment strategies and safety organized practice. This leads to the larger issue of whether DR is simply a two-track system defined by making a finding or not making one, a system reform, a well-defined practice

model, or perhaps a combination of all three. If it is a practice model, there is a desire to see it implemented with more fidelity and a hope that the practice is given time to mature so its impact on child welfare outcomes can be truly understood. Furthermore, administrators are focused on sustaining the fundamental underlying paradigm shift, not just a two-track system.

7.3.2. Expansion and Adoption

The Office of Children, Youth, and Families is in the process of rolling out Governor Hickenlooper's Child Welfare Plan "Keeping Kids Safe and Families Healthy 2.0." This plan includes improvements to the Child Welfare Training Academy; new prevention initiatives such as SafeCare, Community Response, and Nurse Family Partnership – Augmentation; mobile technologies for caseworkers; funding reform; a statewide child abuse and neglect hotline and public awareness campaign; and consistent screening rules and practices (including statewide training for enhanced screening and RED teams). These child welfare reforms align with Colorado's efforts to expand Differential Response and should serve to strengthen the supports for this program.

There was great appreciation for the collaboration between the state and the DR counties in that no one was left out of the decision-making process and everyone offered input into the design and implementation of DR in Colorado. At the state level, lessons learned include establishing a leadership structure that involves teams of individuals rather than a single project director, which allows for continuity of policy and practice. In comparison to other child welfare initiatives in Colorado, the perception of stakeholders is that there are fewer champions for differential response. Therefore, CDHS staff at all levels should be recruited to serve as DR ambassadors and advocates to facilitate the broad-based community and agency support required of such a system reform.

At the county level, there are implications for staffing, training, coaching, supervision, budgeting, and community outreach. The primary consideration for new counties adopting DR is to have a clear understanding of the practice model’s impact on workload and plan accordingly for training and staffing to minimize the adverse effects of turnover and burnout. The self-selection of caseworkers and the informal process for assigning caseworkers to either the FAR or IR tracks fostered increased satisfaction of caseworkers, as their strengths and philosophical stances were better matched with their practice.

The training should be front-loaded to align with implementation of the practice and to limit the training scope to discrete components rather than the overall picture of DR. Caseworkers want training that is more case-specific and topic-focused, while community stakeholders suggest ongoing trainings and continued evaluation of DR in Colorado. It is

imperative that supervisors receive extensive training in the model to better support caseworkers and serve families. Furthermore, supervisors should go out with caseworkers on initial visits periodically to gain first-hand knowledge of the family assessment response.

Caseworkers repeatedly suggested that counties preparing to implement DR for the first time should place a heavy emphasis on coaching activities–specifically shadowing–while urging supervisors to focus more on mentoring. Although FAR caseworkers received the state-level and external coaching well, there were scheduling conflicts and perceptions that IR caseworkers were less likely to take advantage of the coaching opportunities. It is essential that all caseworkers are engaged in coaching and that their schedules and workloads are taken into consideration to maximize participation. Additionally, the coaching should be done by individuals with expertise in child welfare to provide the most efficient and effective experience for caseworkers. It should be noted that coaching is a significant component of the recently revised Colorado Child Welfare Training Academy curriculum.

New jurisdictions should use findings from the cost study to prepare their budgets for DR implementation while being prepared to make adjustments to labor and resource allocations as the practice evolves. Based on the difference between the tracks on follow-up costs, it can be inferred that FAR cases will be less costly than IR cases over time based on lower levels of re-involvement. This may ultimately create cost efficiency for family assessment response in achieving better outcomes at a lower cost. This long-term implication for county budgets should be taken into account when deciding to adopt DR.

Lastly and perhaps most importantly, new counties looking to successfully implement DR must earn the support and engagement of the community. To accomplish this, both DHS and partner agencies must commit to an open and honest dialogue about the DR model and its impact on child safety, service capacity, and stakeholder roles. Furthermore, DHS staff should continue to reach out to law enforcement, GALs, CASAs and child welfare agencies regarding child safety within a DR approach, finding common ground for supporting families and protecting children to alleviate the tension that remains.

These implications must be considered in light of concerns about how to allocate the increasingly limited resources in child welfare. Specifically, there are questions about whether it makes sense to target more services and resources to a low-risk population that is unlikely to be re-involved. Stated another way, some argue that scarce services and resources should be given primarily to higher-risk cases who are more likely to come back into the system. This study was not designed to address this concern, but did show that FAR families received more services and caseworker contacts than did IR families, but at a comparable cost. Furthermore,

there was a decrease in post-FAR referral acceptance rates, which shows that DR in Colorado is not expanding the population of low- and moderate-risk families served by child welfare. Based on the overall safety, process, and cost results from the study, the question must also be asked about whether the Colorado DR model could benefit high-risk families as well.

7.4. Future Research and Evaluation

The three-year program evaluation also revealed avenues for future research of differential response. The hope is that researchers and evaluators will build on this work and continue to generate new evidence on this system reform for the field. Future research on DR must include more of the family voice. A more participatory evaluation design would better include families in all aspects of the evaluation. The family exit survey had the right intention but was insufficient to engage families in the evaluation. Although different types of incentives and survey modalities were offered, it was not enough to produce a respectable response rate. Interviews and focus groups may be better approaches to encouraging family participation in the evaluation. More attention should be paid to the instruments used to capture family and caseworker perspectives. Validated instruments that measure family engagement should be considered, and new instruments should be evaluated to ensure reliable and valid measurement. Cognitive testing should also be conducted to gauge the appropriateness of survey instruments for the target child welfare population. Future cost studies should consider administering a data collection template at the beginning of the project to more accurately and efficiently capture level of effort and financial start-up costs.

Another area for researchers to build on is examining the relationship between race/ethnicity and safety outcomes. For example, there was a lower probability of all six safety outcomes for families where the ethnicity of the youngest child was either unknown or something other than Caucasian, Hispanic, or African-American. However, given that the vast majority of the children in this count are “unknown” rather than of another other ethnicity, this finding may have limited practical importance.

Researchers also should consider the following suggestions to enhance the study of DR: (1) randomization at the caseworker level to control for differences in experience and skill level or inclusion of a cluster variable to account for differences in caseworker practices; (2) incorporation of implementation fidelity assessments to measure potential contamination of treatment practices; and (3) longitudinal data collection and analysis to explore the long-term impact of family assessment response on child safety, family engagement, and follow-up costs.

Finally, the results of this evaluation should be widely disseminated to child welfare practitioners and stakeholders to provide a common reference point for dialogue on the impact of differential response on child safety, well-being, and permanency outcomes. As such, the DR initiative may ultimately provide an opportunity to redefine the reputation of child welfare in the eyes of parents, practitioners, policymakers, and the public.