Allegheny County Department *of* Human Services

RFP Response Form

Process and Outcome Evaluation of Hello Baby, a Tiered Intervention Program for New Parents and Their Babies

PROPOSER INFORMATION

Proposer Name: The Urban Institute

Authorized Representative Name & Title: Rachel Conway, Senior Director, Office of Grants, Contracts, Purchasing, and Pricing

Address: 500 L'Enfant Plaza SW, Washington, DC 20024

Telephone: 202-261-5409

Email: rconway@urban.org

Website: www.urban.org

Legal Status: ☐ For-Profit Corp. ☒ Nonprofit Corp. ☐Sole Proprietor ☐Partnership

Date Incorporated: March 30, 1968

Partners and/or Subcontractors included in this Proposal: Chapin Hall at the University of Chicago

How did you hear about this RFP? *Please be specific.* Emily Putnam-Hornstein

REQUIRED CONTACTS

	Name	Phone	Email
Chief Executive Officer	Sarah Rosen Wartell	202-833-7200	swartell@urban.org
Contract Processing Contact	Victoria Gee	202-261-5514	vgee@urban.org
Chief Information Officer	Khuloud Odeh	202-261-5759	kodeh@urban.org
Chief Financial Officer	Nani Coloretti	202-261-5985	ncoloretti@urban.org
MPER Contact*	Julie Schor	202-261-5862	LHPOPS@urban.org

* *MPER is DHS's provider and contract management system. Please list an administrative contract to update and manage this system for your agency.*

BOARD INFORMATION

Provide a list of your board members as an attachment or in the space below.

See attached for full Board of Trustees list.

Board Chairperson Name & Title: Click here to enter text.

Board Chairperson Address: Click here to enter text.

Board Chairperson Telephone: Click here to enter text.

Board Chairperson Email: Click here to enter text.

REFERENCES

Process and Outcome Evaluation of Hello Baby, a Tiered Intervention Program for New Parents and Their Babies

Provide the name, affiliation and contact information [include email address and telephone number] for three references who are able to address relevant experience with your organization.

Please do not use employees of the Allegheny County Department of Human Services as references.

Noelle Simmons, Deputy Director, Economic Support and Self-Sufficiency, San Francisco Human Services Agency, Noelle.Simmons@sfgov.org, 415-557-5753

Caitlin Cross-Barnet, PhD, Social Science Research Analyst, Center for Medicare and Medicaid Innovation, Caitlin.Cross-Barnet@cms.hhs.gov , (410) 786-4912

Bonnie Hommrich, SLI Government Solutions, bhommrich@sligov.com, (502) 523-0578

PROPOSAL INFORMATION

Date Submitted 1/31/2020

Amount Requested: Total direct cost $509,209; total cost $910,250

CERTIFICATION

Please check the following before submitting your Proposal, as applicable:

☒ I have read the standard County terms and conditions for County contracts and the requirements for DHS Cyber Security, EEOC/Non-Discrimination, HIPAA and Pennsylvania's Right-to-Know Law.

☒ By submitting this Proposal, I certify and represent to the County that all submitted materials are true and accurate, and that I have not offered, conferred or agreed to confer any pecuniary benefit or other thing of value for the receipt of special treatment, advantaged information, recipient's decision, opinion, recommendation, vote or any other exercise of discretion concerning this RFP.

Choose one:

☐ My Proposal contains information that is either a trade secret or confidential proprietary information and I have included a written statement signed by an authorized representative identifying those portions or parts of my Proposal and providing contact information.

OR

☒ My Proposal does not contain information that is either a trade secret or confidential proprietary information.

ATTACHMENTS

Please submit the following attachments with your Response Form. These can be found at http://www.alleghenycounty.us/dhs/solicitations.

- MWDBE documents
- Allegheny County Vendor Creation Form
- 3 years of audited financial reports
- W-9

REQUIREMENTS

Please respond to the following and submit only one Response Form, even when proposing both Process and Impact Evaluations. If you are only proposing to perform one of the Evaluations,

leave the other section blank. Each Evaluation will be scored separately and the maximum score for each Evaluation is 135 points.

All Proposers should complete Section A. Complete this section only once, even if you are proposing both Evaluations. To score each proposed Evaluation, your score from Section A will be added to your response to the Evaluation-specific sections. (Each Evaluation-specific section is worth a possible 80 points).

Process and Outcome Evaluation of Hello Baby, a Tiered Intervention Program for New Parents and Their Babies

<u>Section A – to be completed by all Proposers.</u> Complete only once even if proposing both a Process and an Impact Evaluation. Your response to this section should not exceed 5 pages.

Mission and Commitment (15 points)

1 **Describe why you feel that you are the best candidate for this opportunity and how the Hello Baby initiative fits within your mission**

As partners on this evaluation, The Urban Institute (Urban) and Chapin Hall's Center for State Child Welfare Data (the State Center) are uniquely positioned to conduct the most thoughtful and thorough evaluation of Hello Baby. The two organizations have a two-decade relationship working together on large child welfare projects. Current joint projects include Planning the Next Generation of Evaluations for the John H. Chafee Foster Care Program for Successful Transition to Adulthood (Chafee), and Supporting Evidence Building in Child Welfare, an ACF-funded project to conduct rigorous evaluations of interventions for the child welfare population. Together, we bring a nationally recognized team of experts with decades of experience evaluating child and family interventions using implementation science and both experimental and quasi-experimental designs. Both organizations have a mission to conduct research to improve the lives of vulnerable children and families.

Each organization brings nationally recognized researchers with extensive experience leading and conducting both process and impact evaluations of child maltreatment prevention and home visiting programs. Members of the proposed team have served DHS on numerous child welfare projects, such as a study of the impact of Allegheny's Family Support Centers (FSCs) on maltreatment investigations. Through that body of work, we are familiar with the demographic features of the County's municipalities and neighborhoods. We understand the historic and present scope of DHS's mission to comprehensively serve Allegheny's residents, especially the most vulnerable. We are well-versed in DHS's data resources, organizational structure, and business processes. And we share a commitment to discover effective ways to engage and serve families likely to come into contact with the child welfare system.

Organizational Experience (40 points)

2 **Describe your organization's strategy experience in conducting large scale implementation studies**

Urban has extensive experience conducting large, complex implementation studies of home visiting and child welfare programs. For example, Dr. Sarah Benatar led the implementation study of Welcome Baby, a home visiting program in Los Angeles that provides education and support for pregnant women and mothers of newborns. Urban conducted annual site visits, a 3-year longitudinal in-home survey of participants and a comparison group of women and their children, and matched participant records to MediCal files to look at healthcare outcomes. In a later expansion of the program, we conducted a 10-day site visit and held individual and small group semi-structured interviews with 94 key informants, including First 5 LA staff, training and technical assistance providers, and Welcome Baby providers. Findings from the study guided program improvement and mid-course corrections.

Urban also recently completed a multi-site evaluation of the Children's Bureau's demonstration providing housing for child welfare-involved homeless families (known as SHARP). The implementation study examined how each site's program model functioned over five years and included analysis of service integration and systems change. We used program documents and yearly interviews with program planners, administrators, partners, and frontline staff and leveraged existing information from the grantees' federal progress reports. Dr. Bridgette Lery served as Co-Principal Investigator for the local evaluation in one site. That implementation study employed a continuous quality improvement (CQI) framework as a strategy to manage the complex nature of the intervention.

Process and Outcome Evaluation of Hello Baby, a Tiered Intervention Program for New Parents and Their Babies

From its earliest days, State Center staff have been at the forefront of large-scale implementation and evaluation studies. That history starts with the very first randomized family preservation studies in the mid to late 1990s and continues to this day in locations as diverse as New York City, Tennessee, Texas, and California. In New York City and Tennessee, State Center staff served as the Title IV-E waiver evaluators. Both evaluations involved the scale-up of two evidence-based interventions. For each, State Center staff managed both the process and outcome evaluation. In Tennessee, the scale-up for both programs were multi-site, requiring State Center staff to work closely with stakeholders to understand implementation on the ground in each site's context. In Texas, State Center staff supported implementation and evaluation during a statewide roll-out of Community-based Care by providing state-of-the-art analytics for service targeting and outcome monitoring.

3 Describe your organization's experience collecting data.

Urban's direct data collection experience provides an understanding of the challenges and nuances of data collection in particular communities or with government agencies. We develop, test, and implement interviews and focus groups of public agency staff, nonprofit organization activities, and clients receiving services that contribute to needs assessments, case studies, and ethnographic or population studies. We are especially strong in designing, administering, and analyzing surveys. We have collected phone and web-based surveys of agency staff (e.g. child welfare, housing, SNAP) using Qualtrics software. We have also designed and analyzed in-person surveys such as baseline and one-year follow-up surveys of families in the treatment and control groups of the SHARP evaluation, with content capturing child and family well-being, neighborhood quality and housing stability, and other outcomes.

State Center staff also bring extensive experience with data collection, most aptly illustrated in a first-of-its-kind time use survey of child protection services workers in New York City. To carry out that work, staff organized 29 focus groups across all geographic areas of the City and a cross-section of administrative functions and personnel role types. The input from those focus groups was used to craft a survey that included nearly 2,000 role-specific questions about the work involved in CPS investigations, service referrals, court work, and placement. State Center staff were deeply involved in building analytical files, cleaning the data, and reconciling data anomalies with various stakeholder groups including ACS leadership. Demonstrating innovation, all the collected data were linked to children through a worker assignment key that connects children, workers, and survey responses, creating a unique ability to tie worker time use responses to the outcomes for children served by that specific worker.

4 Describe your organization's experience conducting action research

Urban's work is directed at improving processes through the voices of those engaged in and affected by those processes. For example, our current evaluation of the Family Unification Program includes a Community of Practice, bringing together the housing and child welfare practitioners to exchange ideas on program implementation. In several studies, we have developed connections within communities to ensure the voices of program participants are reflected in the questions we ask and the conclusions we draw. These have included community advisory boards, community forums, data walks, and other community events.

To the extent that action research refers to a disciplined process of inquiry conducted by and for those taking the action, State Center staff regards nearly all of its research as action research. We are deeply committed to the idea that evidence-building requires a hand-in-hand partnership between practitioners and researchers, working together to solve some of the world's most vexing problems. Our work organizing the CQI evaluation framework reflects what we know about knitting together rigorous science and practical utility. We strive to provide formative and summative judgements in a timely, actionable manner. For example, State Center staff worked with the Harlem Children's Zone (HZC) to help leaders articulate their theory of change and ascertain the extent to which their service pipeline conforms with their model. State Center staff helped organize HCZ's effort to integrate an overarching evidence-informed improvement

strategy, and met with senior leadership and program staff to learn about the organization's priorities from the people who carry out the work.

5 Describe your organization's experience analyzing large, complex quantitative data sets

The Urban and State Center teams are widely recognized experts in managing and analyzing large, complex data sets related to children and their families. State Center staff have been at the forefront of innovation in this area since 1979, which is longer than any other child welfare research organization. The work started with foster care, child protection, and social services data in Illinois and continues today across the human services: Medicaid, TANF, juvenile justice, mental health, schools, and criminal justice. The range of our research speaks directly to that experience. From such standard child welfare outcomes as permanency, reentry, and recurrence of maltreatment to multi-service families to the impact of supply induced demand on racial and ethnic disparities, few organizations can match the breadth and depth of our experiences. We have more than 30 years of experience doing geographic or spatial analysis. We have written our own predictive risk models for placement following a substantiated allegation of maltreatment and placement stability, both of which have been deployed in Tennessee in the last four years. To support that work, programmers at the State Center have constructed their own machine learning algorithms that deploy the standard battery of refinements to the underlying methods including random forests and gradient boosting. In addition, we have experts in econometric methods, with experience in propensity score matching, the full range of linear and non-linear models, and random effects models. We also have extensive familiarity with both agent-based and system dynamic simulation models. Naturally, we have working knowledge pertaining to a complete range of statistical software including SAS, R, SPSS, STATA, HLM, and SuperMix as well as HTML, Java, JavaScript, and Python.

Most of Dr. Lery's twenty years of research experience has involved analyzing the administrative data of state and local child welfare and adjacent systems to measure trends, guide policy, evaluate programs, and target resources to the subpopulations that need them most. She employs longitudinal and multilevel analytic methods to understand how prior experiences affect future risks, and how the social context surrounding families influences their exposure to and experience with public systems such as child welfare.

In particular, Drs. Wulczyn and Lery have extensive experience organizing, linking, and analyzing Allegheny's administrative data, producing reports on foster care dynamics and making policy recommendations. One project involved creating a multi-system placement file, linking child-level records from several child-serving systems in Allegheny's data warehouse – child welfare, mental health, mental retardation, juvenile justice, and multi-system rapid response team – to understand children's out-of-home placement experiences within and across those systems.

6 Describe your organization's experience analyzing administrative data sets for evaluation purposes

Staff at both Urban and the State Center specialize in analysis of administrative data for program evaluations using experimental and quasi-experimental designs. For the SHARP evaluation, Urban used child welfare and homelessness administrative data from five sites to measure the impact of supportive housing on housing stability and child welfare outcomes. As part of its Chafee evaluations, Urban has used child welfare data from multiple jurisdictions and is currently analyzing administrative data from ten states, including Pennsylvania, for an outcome study of the Education and Training Voucher (ETV) program for youth aging out of foster care.

State Center staff have provided consultation to child welfare agencies around program evaluation and has long relied upon administrative data to evaluate child welfare services. State Center staff have analyzed administrative data for the federally funded evaluations of family preservation programs, an evaluation of a large evidence-based intervention scale-up in New York City (Child Success New York City), and a cluster randomized trial of child representation for the Children's Bureau. They also recently completed seven Title IV-E Waiver evaluations.

Process and Outcome Evaluation of Hello Baby, a Tiered Intervention Program for New Parents and Their Babies

For one detailed example, 18 months ago, leadership in Tennessee recently asked us to evaluate their longstanding Youth Villages Intercept program. To do that, we had to integrate five distinct data sets: CPS records, CANS assessments, placement records, and worker assignment records from the state of Tennessee. and encounter/services records from Youth Villages. Data in hand, we then had to design a quasi-experimental study that comports with the standards set forth by the Title IV-E Clearinghouse. In summary, we solved a number of problems the affect observational studies of program impact. For the comparison group, we implemented many-to-many exact matching, which is the method of choice, all else being equal. To manage other confounds, we used the geographic data to control for placed-based differences in referral and placement probabilities. We also linked children to caseworkers to manage any bias workers may exercise in the decision-making process vis-á-vis service/placement referrals. For the study of sustained effects, given the comparison group had no treatment start and stop dates, we devised a method for imputing those dates and adding them to the comparison group records.

Working together, Drs. Wulczyn and Lery used Allegheny's child maltreatment and foster care data to conduct quantitative evaluations of the Family Support Centers, Family Group Decision Making, High-Fidelity Wraparound, Inua Ubuntu, and the Systems of Care Initiative, among others, to examine their effects on key child welfare outcomes. In particular, the Family Support Centers evaluation used individual-level encounter data from the FSCs linked to CPS records to describe patterns of FSC service use and to examine the relationship between service use and subsequent maltreatment investigation.

7 Describe your organization's experience utilizing qualitative research methods

Many Urban projects, and virtually all evaluations, involve site visits, staff/administrative interviews, focus groups, case studies, and/or record/content analysis (e.g., Welcome Baby, SHARP, Chafee, Family Unification Program, Partners for Fragile Families, and the Young Parents Demonstration). The team assembled for this evaluation has expertise designing and conducting mixed methods evaluations of programs for low-income pregnant women, new mothers and babies, parents involved with the child welfare system, and children. Qualitative researchers on the team have designed and conducted case studies of Medicaid, CHIP, early childhood home visiting and maternal and child health programs in all 50 states, the District of Columbia, and several U.S. territories.

We design and conduct semi-structured interviews with various stakeholder groups ranging from government officials to providers to community members, as well as focus groups with unique and diverse sets of consumers or program beneficiaries. When appropriate, we've conducted semi-structured observations of client-provider interactions or waiting room settings. Findings from qualitative work are systematically analyzed to offer insight on how well programs are working, where they fall short, and beneficiary experience.

State Center staff have also had a long history of qualitative work, with nearly every project conducted in NYC in the last 20 years having a qualitative research component. Consistent with action research, State Center staff have developed data capture tools in close collaboration with stakeholder groups that include the individuals who take part in the study themselves. Their process includes a rigorous pre-test/revise cycle that strengthens the work product and the insights gained.

8 Describe your organization's experience developing or adapting existing measures, preferably in the home visiting, maltreatment or early childhood fields

Urban and State Center staff have both developed and adapted measures for the home visiting, maltreatment, and early childhood fields. For the evaluation of Welcome Baby, Urban adapted validated measures of child development, safety, and maternal well-being for an in-home longitudinal survey of families enrolled in the program and a comparison group. Surveys were pre-tested at each wave, and adapted for age-appropriateness. We've also used and adapted measures designed to assess the career trajectories of home visitors, the workplace characteristics that affect their job satisfaction, and their long-term career goals. Urban co-authors the National

Home Visiting Resource Center's annual *Home Visiting Yearbook*, contributing analyses of the current state of home visiting in the United States. State Center staff have experience working with Prevent Child Abuse America (i.e., Healthy Families America) on the development of a cross-site integrated home-visiting data base, a project that gives us experience with data collection mandates for the MIECHV reporting requirements.

In addition, the Urban and State Center team have used administrative data to estimate the occurrence of maltreatment in a variety of ways. Administrative data only capture reported maltreatment, thereby not counting instances that go unreported and over-counting instances where the allegations are unfounded. We addressed the over-count problem in the State Center's evaluation of Allegheny's FSCs by measuring investigation rates, i.e., reports that, based on information presented at the hotline, require an investigation. At times, a more conservative measure of maltreatment might be called for, such as a substantiated allegation or an allegation that rises to the level of requiring a case opening. Lery and colleagues reported such metrics in their local SHARP evaluation. (Reference for this study and others throughout this proposal are available upon request.)

9 **Describe your organization's experience securing funding for evaluations, and provide ideas and strategies for working with DHS and other partners to secure additional evaluation funding.**

For over 50 years, Urban has demonstrated experience securing funding for research and evaluations on social and economic issues. In 2018, Urban raised nearly $85 million in grants and contracts.

The national focus on child abuse prevention is having a renewed moment. The Family First Prevention Services Act, while not addressing primary prevention, puts a spotlight on prevention broadly. Foundations are supporting evaluations of programs aimed to reduce the need for child welfare involvement, particularly those that fit the Family First criteria for federal reimbursement. We can capitalize on this national focus, working with DHS to pursue additional evaluation funding from local and national foundations.

For a few examples, Arnold Ventures released a rolling solicitation for *Randomized Controlled Trials to Evaluate Social Programs Whose Delivery Will Be Funded by Government or Other Entities.* Arnold has funded five child welfare projects, including a replication study of the Child First home-visiting program. Hello Baby's focus on families at birth means that many young mothers may be involved, and The William T. Grant and Annie E Casey Foundations are interested in young people (through their mid-20s). The Doris Duke Charitable Foundation has a focus area on child maltreatment prevention.

Targeting these and other funders, possible proposals for Hello Baby include: (1) extending study of the priority group beyond the federal grant period to observe outcomes 12-months post-exit – a requirement to receive a "well-supported" rating in the Title IV-E Prevention Services Clearinghouse; (2) a study focusing on an important subpopulation that current funding can't support; and (3) an annual longitudinal survey of priority tier families to understand the extent to which Hello Baby engagement and services contributed to family well-being.

Process and Outcome Evaluation of Hello Baby, a Tiered Intervention Program for New Parents and Their Babies

<u>Section B – Process Evaluation</u>. If you are proposing to perform the Process Evaluation, respond to the items below. If you are not proposing a Process Evaluation, leave this Section blank and move to Section C. Your response to this section should not exceed 15 pages.

Understanding the Scope of Work (30 points)

1 **Provide an overview of your approach to the project tasks, including explanations and rationales for any suggested modifications.**

Our approach is centered around five tasks articulated in the RFP, plus one we have added: collaboration, documentation, assessment, monitoring, understanding, and feedback.

1. Collaborate with ACF, model developers, evaluation technical assistance providers, and service providers to refine the design and implementation of the evaluation. We propose a Continuous Quality Improvement Evaluation Framework (CQI-EF) that lends itself perfectly to collaboration by design. The evaluation team, in partnership with DHS, will develop and iterate a monthly implementation dashboard. Its purpose is to capture program inputs and outputs in order to quickly identify model fidelity problems and successes, such as enrollment progress in the priority and differentiated groups and frequency of home visits in the priority group. The dashboard will form the backbone for regular CQI meetings. The outputs of these meetings will inform the project partners about implementation progress, and conversely will be a venue for putting project partner input into practice.

2. Document participant experiences by tiered level. We plan to document participant experiences within each tier in two primary ways. We will conduct focus groups with participants and make efforts to recruit those who opt out to attend focus groups as well. At each point in their decision process, we will ask parents to reflect on their awareness of services and whether they see participation as beneficial to them as parents focused on the well-being of their children. Understanding how answers to these questions differ by tier is a central concern.

We will also use FSC data and the priority program's client management data to characterize participant experiences according to what services they received.

3. Assess implementation facilitators and barriers, and strategies to address those issues. Our implementation (process evaluation) takes a dual perspective: (1) implementation barriers, facilitators, and strategies and whether the program is 'installed' as planned and (2) how is the program delivered, given what was installed, as a function of facilitators, barriers, and strategies. The Stages of Implementation Completion (SIC) method offers a highly reliable, practical, and concrete way to pinpoint implementation milestones in child welfare programs. We will develop detailed SIC measures to answer the research questions about implementation (How did the process of care change and was the capacity developed to carry out the process?) and fidelity (How well did those changes play out?).

4. Monitor participant outcomes using administrative data and primary data. This task will fall primarily under the outcomes study.

5. Use the processes above to understand engagement, retention, and attrition, and the connection to outcomes.

Our proposed randomization strategy provides a mechanism for understanding whether decision-making considerations are correlated with enrollment, engagement, persistence, and outcomes. The process study will enrich those findings with insight about why families chose to engage, and persist (or not), and how that is connected to Hello Baby's implementation strategies and model fidelity.

6. Provide active feedback to stakeholders as input to the CQI process.

Process and Outcome Evaluation of Hello Baby, a Tiered Intervention Program for New Parents and Their Babies

Our orientation to evaluations of this sort combines rigorous methods, statistical analysis, and a keen awareness of how evaluation *must* inform the evolution of the program over time as implementation plans are modified by timely feedback. Our aim is to help DHS address program improvement built on close to real time evidence about outcomes. With stakeholders, we aim to develop a clear understanding of the theory of change, the process, quality, and capacity improvements that define Hello Baby relative to business as usual and the mechanisms that generate outcomes so that we can work with stakeholders to generate the evidence needed to adapt Hello Baby to its strengths and weaknesses. Although thumbs up/thumbs down evaluations are important to evidence-building in the long run, we see a more dynamic learning agenda in play in Allegheny County over the nearer term.

2 Describe your understanding of the Hello Baby target populations and risk factors

The Hello Baby program uses a predictive risk model (PRM) to identify new parents at substantial risk for future child welfare involvement and post neo-natal death. The program offers support for three specific populations, as determined by the PRM: low risk families; families that have a higher level of need than those deemed low risk; and priority level families, who have the greatest level of need. Previous work in Allegheny County demonstrates that uptake remains low in the face of a strong network of prevention services, suggesting families at the greatest risk of harm are not being adequately served. This underutilization is thought to be a consequence of many factors, including geographic proximity to service centers, transportation challenges, social isolation, lack of outreach, and limited support networks.

Moreover, research suggests that key correlates of child abuse include untreated maternal depression or mental illness, substance use disorder in caregivers, age of the mother, and intimate partner violence in the home, as described in the Hello Baby services RFP. As such, families in the priority group may face numerous and overlapping service needs that the peer/social worker teams will help address. How they do so, using the COACH model, as well as the extent to which families participate in and benefit from such services will be important to capture in the process evaluation.

Finally, up to 40% of priority families will be active in child welfare at the time of birth and more will likely become active within the evaluation timeframe. Other families will already be enrolled in an evidence-based home visiting program, and may not want to be inundated or burdened with another program or services. Variability of child welfare and other program involvement in the target populations offers an opportunity to leverage connections to existing systems, while also ensuring that families are not inundated or receiving duplicate services.

3 Describe your understanding of the field of primary prevention of maltreatment, as well as issues specific to Hello Baby.

Primary prevention of maltreatment is directed at the general population on a universal level, attempting to avoid maltreatment in the first instance. As such, all members of a given community have access to and can potentially benefit from the offered prevention services. Although designed to address secondary prevention, the the Family First Prevention and Services Act (FFPSA) has generated a renewed emphasis on early intervention and on primary prevention services as well.

DHS's interest in prevention is not new. Over a decade ago, Wulczyn and Lery began partnering with DHS to evaluate several child welfare programs meant to prevent secondary and tertiary maltreatment (e.g., Inua Ubuntu, Systems of Care Initiative, Hi Fidelity Wraparound, Family Group Decision-Making). Our evaluation of the FSCs' impact on maltreatment investigation rates found promising results, namely that the FSCs appear to have a protective effect in communities. That study contributed to DHS's thinking about how it can optimize that resource for secondary and tertiary prevention, but also to deliberately expand the reach of FSCs for primary prevention.

Programs like Hello Baby take a universal approach to raising the awareness of the general public and normalizing expectations around seeking assistance for new parents. This type of primary prevention approach also offers a better chance of engaging a greater proportion of the highest risk families for child maltreatment, potentially decreasing the services necessary downstream for these families. Yet while primary prevention services may play a part in increasing the engagement of high-risk families, there is still reason to believe that many high-risk families may opt out of voluntary services. Hello Baby may not reach many of the highest risk families if they choose to opt out at the hospital, via postcard, or during subsequent engagement efforts by the peer/social worker teams. Capturing engagement and attrition, as well as when they both occur, will be a key part of both the process and outcomes evaluations.

Moreover, the primary prevention approach of Hello Baby raises unique challenges around the PRM score and potential stigmatization of some families. First, some families may opt out bedside at the hospital because they do not want their information used to target them for eligibility. Second, because the PRM requires certain administrative information to calculate a risk score, children not known to the system or who don't have enough data in the Data Warehouse will not receive a score. According to a 2019 report about Allegheny's other PRM, the AFST, approximately 10 percent of referrals were not generating a risk score, highlighting children who may be at risk of maltreatment but missed by the PRM. Third, while no system can remove the chance of falsely screening-in children based on their risk score, the PRM does provide the potential for increasing stigma if scores are not closely guarded and call screeners are not adequately trained on the possibility of false positives and confirmation bias. Both the impact and process evaluations will play a key role in understanding how the use of the PRM to target families plays a role in who opts in or out, and why.

4 **Describe the existing evidence base in terms of key features of effective implementation and outcomes for primary prevention programs.**

Implementation

In a 2011 brief, Deborah Daro outlined a set of principles demonstrated to be effective across diverse disciplines and service delivery systems. These lie at the core of effective interventions, and include: a strong theory of change that demonstrates clear pathways, specific outcomes, and curriculum content; recommended duration and dosage that is systematically applied to all enrolled in services; a clear, well-defined target population, as well as clear eligibility criteria and outreach/engagement strategies; a strategy for guiding staff in balancing content delivery and respect for the family's culture and circumstances; a method to train staff on model delivery; a supervisory system to support direct service staff; reasonable caseloads that allow direct service staff to accomplish core program objectives; and the systematic collection of process implementation data to ensure standards of model fidelity are met (Gateway, 2011).

These principles are echoed in a review of 500 quantitative studies that evaluated preventative intervention programs, in which the authors identified five pillars that significantly influence the implementation of a prevention program. These pillars include *community characteristics, provider characteristics, innovation characteristics, the prevention delivery system* (i.e., features related to organizational capacity) and the *prevention support system* (i.e., training and technical assistance). In order to implement an intervention successfully, variables in all five pillars must interact in a constellation to support successful implementation (Durlak & DuPre, 2008).

Outcomes

It is often difficult to establish that a primary prevention program did in fact reduce maltreatment directly, because it is hard to know whether the program served parents who would have maltreated without the program. (Hello Baby intends to minimize this problem using the PRM to target families.) Instead, most programs seek to affect an interim outcome – either individual or community-level – that in turn may reduce the likelihood of maltreatment. Programs focused on individual-level outcomes include parenting programs and therapy, which seek to change attitudes, knowledge, or behavior. The notion of collective parenting is generally not favored in

our society (Daro, 2016), yet some approaches do draw this causal path to maltreatment, seeking to boost social capital and collective efficacy in a community – constructs measured only at an aggregate level – in order to create the broad social supports that can help parents protect and raise children (Kim, Gloppen, Rhew, Oesterle, & Hawkins, 2015).

5 Identify any logistical challenges and describe how they will be addressed or minimized.

We expect it may be difficult to reach individuals who decline to participate in Hello Baby. The outcomes study will use the Data Warehouse inputs for the PRM to understand what characterizes families who opt out. But it will also be important to hear directly from some families who opt out in order to gain insight into why. Though we anticipate this challenge, we have extensive experience reaching out to and engaging individuals who have declined program participation. We will make multiple contacts, offer an incentive, and ensure that all research staff involved in this effort are adequately trained in respectful client engagement to maximize the success of our efforts. Though the best chance of gleaning insight into why families opt-out might be a short survey (see question 8 below), we will consider hosting a focus group with individuals who have recently opted out of the program. These opportunities are sometimes more appealing because of the increased incentives. In addition, we have found that some people like the opportunity to provide more unstructured feedback. As described in the data collection section, we will report engagement numbers on the implementation dashboard and discuss them at CQI meetings to identify problems quickly.

As with any cross-system effort, we anticipate logistical challenges during implementation. The research team has extensive experience evaluating programs that involve cross-system coordination. We know what challenges to expect and how to set up CQI checks to identify even the problems we don't expect. For example, we will need to work with the hospital and DHS to craft a specific protocol for giving the research team information about those who opt out, and the protocol may differ for those who opt out at the hospital verses later by postcard. As described in the data collection section, we will develop a regular data transfer process with DHS to help keep track of this task and identify problems early.

6 Describe your experience partnering with other researchers and stakeholders

Our research team is always aware of the benefits of drawing on the expertise of others. For some examples:

- Urban and the State Center have partnered for two decades on child welfare projects.
- Bridgette Lery partnered with the researchers who developed Allegheny's AFTS and PRM, on a PRM project in California. Lery is planning a new project with Putnam-Hornstein and Vaithianathan in Douglas County, Colorado related to their PRM.
- Urban partnered with technical assistance providers, the Children's Bureau, and local evaluators at all five sites to evaluate a five-year federal demonstration project testing a supportive housing program for child welfare-involved families who were homeless. Lery was a local evaluator at one of the sites.
- Urban recently completed a large evaluation of enhanced prenatal care programs for Medicaid beneficiaries, funded by the Center for Medicare and Medicaid Innovation (CMMI), in which we partnered with three other firms that had complementary skill sets. This work received exceptional reviews and a special commendation by CMMI leadership. Sarah Benatar led the process study.

We value the partnerships developed with stakeholders throughout the research process. These relationships are critical to the success of program evaluations in particular because they are the gateway to honest and clear information that informs evaluation conclusions. Our priorities in these relationships rest on the promise of confidentiality, the explicit acknowledgement that these projects are not and should not be designed to point fingers at any individual or agency in

particular and are instead about benefiting the providers and beneficiaries for whom the programs are designed to help. Additionally, ongoing recognition of the additional burden that this work can heap on community stakeholders, and efforts to minimize that burden while highlighting the ways in which this research can be beneficial, are at the center of all of our stakeholder interactions.

An important stakeholder group consists of community members. When doing research in communities on its members it is essential to value the needs and concerns of that community. Operating under the principle of "nothing about us, without us," we will take the time to understand the preferences and needs of the communities affected by Hello Baby – particularly the differentiated and priority programs – through our planned process study methods such as key informant interviews with FSC and DHS staff. This approach will always benefit the research, and consistently informs our implementation and impact evaluation work.

Technical Approach (30 points)

7 Please describe your organizations conceptual framework that addresses the components of the process and/or impact evaluation

The major innovations being tested in Hello Baby are (1) the tiered approach to engaging families in services for newborns and their parents, and (2) a persistent, intensive engagement model for the priority population. The process study will describe and measure the extent to which these two innovations were implemented as planned in each tier. We use our Continuous Quality Improvement (CQI) approach to evaluation. CQI is an iterative process of planning a project activity or strategy, implementing it, evaluating it, and then using immediate evaluation results to improve ongoing planning, implementation, and evaluation. Sometimes referred to as "Plan-Do-Study-Act" (PDSA), this approach combines rigorous evaluation methods with the real-world contingencies associated with implementing complex social programs such as Hello Baby.

Introducing a new intervention requires investments in three major areas: the *process* of care, the *quality* of care, and the *capacity* of the providers to deliver process and quality with fidelity. (Wulczyn, Alpert, Orlebeke, & Haight, 2014). We use this frame as a way to conceptualize the "Do" phase – what gets done, exactly? We will collect data in all three areas to document whether processes are being followed, services are being delivered as intended and whether the service providing entities are adequately staffed to meet the needs of participants.

At a high-level, the Hello Baby process involves engagement, service coordination, and case management offered to families who agree to participate. Needs are established with the PRM and/or through referral. To make that happen, there needs to be contact with the identified families followed by work with the family to establish their need for help, a referral to *high quality* access, and preferential access to the target population. Closer to the ground, the *process of care* is triggered when families are provided with in-hospital information on a universal basis. For families that opt-in, there is an assessment of need that is tied to the PRM, at least in part. Families at the lower end of the risk-scale are referred to Family Support Centers. As appropriate, they can expect a visit and a link to services. Families that fall into the highest risk category will be nudged toward priority services using strategies informed by the COACH model. Because DHS and the provider of priority services will have an assessment of risk from which they will be working, the approach taken has to avoid coercion. If a family decides to take the next step, a needs assessment and service plan goals will follow. So long as families persist, the priority services will be in place for three years. Support for families will be unconditional. Operationally, the Priority Service providers will build two-person terms with domain specific knowledge together with clinical supervisors (i.e., *capacity*). The *quality* components touch on reflective supervision, a do-what-it-takes attitude plus the skills needed to undertake motivational interviewing informed by domain specific knowledge that includes the latest evidence pertaining to risk and safety, parenting, trauma, and child development.

Process and Outcome Evaluation of Hello Baby, a Tiered Intervention Program for New Parents and Their Babies

The data collection section specifies how we connect DHS's priority process study domains for each tier to this process/quality/capacity framework and plan to capture the information necessary to answer the research questions.

8 Please describe your evaluation design and approach to sampling families

Our CQI evaluation framework includes identifying and measuring process, quality, and capacity investments that help Hello Baby take shape during implementation – the PDSA "Do" phase, as described earlier. In addition to DHS's priority research questions identified in the RFP, we propose to address some others, for example:

- Will the universal component of Hello Baby create a normative standard, making more parents comfortable with the service?
- How many engagement attempts and how long does it take to engage families in the differentiated and priority groups? For the priority group, how long does it take for families to reach key milestones?
- How do mothers react to being told in the hospital that their history and administrative data will be used to determine eligibility for Hello Baby?
- How do participants and navigators experience the process of care when a participant enrolls in another program, such as an evidence-based home visiting program, but wishes to retain the Navigator relationship? Who is lead? Is there coordination and what does it look like?
- Do efforts to engage and retain families in the differentiated and priority groups affect the level of service offered by certain programs? Does the time and effort it takes to engage families reduce the time case managers can give to other clients? Does it reduce the volume of clients a provider can handle?
- Has the mix of services offered changed as a result of Hello Baby? Do certain communities see more service enhancements than others?

Our methods for answering these questions and the other process study research questions in the RFP are described next.

We will conduct focus groups with participants from each of the three tiers of Hello Baby. We will rely on DHS, FSCs, and Hello Baby staff to help with recruitment via flyers or if they prefer, they can reach out to their clients directly. In our experience, having a trusted individual introduce the concept is more successful than cold calling.

The groups will be designed to solicit information on their experiences enrolling in the program, satisfaction with the services being provided and the staff providing them, whether Hello Baby is meeting their needs, and what concerns they have or changes they might recommend. We expect to host three focus groups per visit, prioritizing the priority tier. We will reach out to each of the tiers in the pilot community with the goal of recruiting between 10-12 participants for each group. We will offer participants a $50 incentive and a light meal. All groups will be facilitated by a skilled moderator.

We will conduct key informant interviews with staff who introduce the program in the hospital, Hello Baby providers, both existing and newly hired staff at community FSCs, and staff from agencies that are the source of Hello Baby referrals. We anticipate working with DHS to identify potential stakeholders. Key informant interviews will assess implementation facilitators and barriers such as training (process and quality of care), capacity, and infrastructure needs (capacity of care). Interview protocols will be designed to ensure systematic collection of data and tailored for different stakeholders to include targeted lines of questioning.

Lastly, we will use a structured observation method during the CQI meetings proposed below to gather information about implementation progress, particularly related to cross-system coordination.

We also propose to seek additional funding to collect survey data. A brief survey of individuals offered the Hello Baby program at hospital bedside could capture mothers' reactions to how the

program was offered, why they chose to participate or decline, whether they are still involved with the program and why or why not. The full sample could consist of the birth list of individuals who were approached in the hospital to ensure an intent-to-treat group that includes those who opt-out at any point. We would recommend special focus on the priority tier but depending on DHS's interest in learning about how the engagement effort is being received by the broader population, we could offer sampling strategies stratified by universal verses differentiated tiers among the 13,000 births per year in Allegheny County. We would offer two modes for the short survey: 1) via a phone call and 2) via text message, and we would offer respondents a small incentive for participating.

9 Describe your approach to data collection and measurement

The Stages of Implementation Completion (SIC) method offers a highly reliable, practical, and concrete way to pinpoint implementation milestones in child welfare programs (Saldana L. C., 2012). The process study will use this model to measure activities within eight stages of implementation that span three broadly accepted implementation phases (Pre-implementation, Implementation, Sustainability) (Saldana L. , 2014; Saldana L. C., 2012). The aim is to capture how the programmatic inputs that characterize the process of care and the capacity to deliver it differs each for the differentiated and priority groups relative to the universal group (implementation), and the quality of those inputs (fidelity).

Table 1 provides a high-level view of the data collection methods we will use to answer the research questions within each of DHS's stated domains of interest. For brevity, we combine the tiers and collapse the research questions into the topics. We will rely on a variety of data sources to collect the quantitative information required under the SIC model. One source will be the client management system that DHS will provide to the Hello Baby service provider to capture information about key areas such as outreach, engagement, and implementation.

Table 1. Process Study Domains by Data Collection Method

Categories per RFP	Key Informant Interviews	Focus Groups	Structured Observations	Client Management System	Participant Survey (if funded)
Outreach	X	X		X	X
Engagement	X	X		X	X
Implementation	X	X	X	X	
Program Response		X			X
Resources	X	X			
Referrals	X	X		X	

We will develop detailed SIC measures to answer the questions within the categories. Broadly, the measures will identify the *process of care* components of Hello Baby, and the *capacity* changes necessary to deliver the program with fidelity. For example, one distinguishing element of the priority program is that two-person teams – a family engagement specialist (peer), and a trained social worker – will be trained to engage and serve families. These teams will form the basis of the priority group intervention, so it will be critical to know whether the teams are fully staffed up to meet the need; when they trained in the COACH model; and did each member perform their key functions? Staffing and training fall under SIC Stage 4: Staff are Hired & Trained, which specifies milestone dates and the agents involved. Table 2 shows this example. Measuring the extent to which the teams perform their functions falls under later SIC stages, and the full table will be developed during the planning phase.

Table 2. Example of SIC Stages, Activities, and Involved Agents (Priority Tier)

Stage	Activity	Involved Agent

4	Staff Hired & Trained	Date service provider contract began Date each peer staff hired Date each social worker hired Date COACH training held	Service Provider Agency, Practitioner
5	Adherence Monitoring Processes in Place	Date client management system training scheduled Date client management system training held Date of first CQI meeting	Agency, Practitioner
6	Services Begin	Date first Priority Program family enrolled Date second Priority Program family enrolled	Practitioner, Family

The evaluation team, in partnership with DHS, will develop and iterate a monthly implementation dashboard about the priority program drawing from the client management system and other sources as needed such as the FSC database. Its purpose is to capture program inputs and outputs in order to quickly identify model fidelity problems and successes, such as enrollment progress and frequency of home visits. The dashboard will form the backbone for regular CQI meetings held at DHS or an FSC (evaluators will participate by phone).

Importantly, the CQI team will promote cross-sector coordination. It will address the data entry and evaluation requirements, produce and discuss implementation progress, and will develop information for DHS to review and use for project decisions. The research team, DHS, and the Hello Baby service provider will be the core members of the team, and other agencies such as hospitals and FSCs will be drawn in as appropriate. This is a key team and we recommend that it meet at least monthly.

Table 3 shows in more detail exactly how we break down the elements of the intervention into its implementation dimensions, (i.e., *process/quality/capacity*). Those dimensions are mapped to specific research questions and the data sources we plan to use to answer the questions. For brevity, the table is limited to one priority program element – two-person teams consisting of a family engagement specialist (peer) and a trained social worker. The fuller evaluation plan will include the other program elements such as low caseloads and reflective supervision with case consultation.

Table 3. Example of Measurement Approach for One Priority Program Element

Program Element and Implementation Dimension	Research Questions	Data Sources
Two-person teams - family engagement specialist (peer) and social worker		
Capacity	Did all teams consist of peer/social worker duo?	Informant interviews, Client Management System
	Are there enough of them to meet demand?	Client Management System
Quality	Were all staff trained in the COACH model?	Informant interviews, Focus groups
Process	Did peers perform their key functions?	Informant interviews, Focus groups

Interviews and Focus Groups

Senior researchers at Urban will train all data collectors on reliability before collecting data using standard training protocols. Two data collectors (one senior lead interviewer and one research analyst) will conduct a semi-structured interview with each key informant and lead focus groups. Our past research studies and experience conducting interviews in person will facilitate building rapport with families and stakeholders to get their buy-in.

We will develop and pilot test a "moderator guide" and key informant protocols based on our review of existing instruments for home visiting evaluations such as Welcome Baby. The guides will include questions mapped to the key research questions and hypotheses as well as probes tailored to the interviewee's role.

Key informant interviews and focus groups will be conducted in English (or Spanish, when preferred). During the interview and focus groups, the research assistant will take close-to-verbatim typed notes and audio record responses with the respondent's permission to assist with note cleaning. Assistants will destroy all audio recordings in accordance with the study data security plan approved by Urban's Institutional Review Board. After each interview, the interview pair will debrief and write a one-page summary memo that captures key information about the informant or focus group and their responses. We will share the summaries within the Urban team to facilitate discussion around emerging themes.

Site visits

The research team anticipates conducting one in-person site visit each year of the evaluation (in addition to a kickoff meeting during the planning phase). The kickoff meeting will serve to discuss the project and the evaluation with the key stakeholders. In a subsequent visit during Year 1, we will interview hospital, DHS, FSC, and Hello Baby staff to obtain information about early implementation progress and challenges, including recruitment and enrollment into the differentiated and priority tiers.

The Year 2 site visit will take place after the implementation data collected using the SIC framework suggest that the program for the priority population has reached Stage 7: Ongoing Services, Consultation, Fidelity, Monitoring, and Feedback. On that visit, we will hold focus groups with the priority target population – both participants and those who declined. We will also conduct group interviews with FSC, DHS, and Hello Baby provider staff to learn their perspectives on the program's implementation ability to address family needs. The Year 3 site visit will have a similar purpose to capture implementation progress and fidelity, including interviews and focus groups with key staff among the partner organizations and participant/non-participant families.

Surveys (not in current budget)

With additional funding and if it interests DHS, we could develop a survey instrument for all individuals offered Hello Baby (i.e., everyone approached bedside, including those that decline or drop out at any stage) to address our study questions but will employ a similar approach to protocol development that we take with interviews and focus groups by starting with a construct map, including items from existing validated instruments, pilot testing, and developing data collection protocols.

During the planning phase, the research team will collaborate with DHS, the evaluation TA provider, ACF, and the cross-site evaluator to review the measures and refine the evaluation plan, ensuring that the plan successfully evaluates Hello Baby's most up-to-date logic model and theory of change. We also plan to complete the IRB approval process and begin data collection during the planning phase.

10 Describe your analytic approach for primary qualitative and quantitative data and proposed use of the Data Warehouse

The research team will first map information from the data sources to the Hello Baby logic model, and focus analytically on examining the program against the logic model, or "falsifying"

the logic model, a process to determine if the logic model represents what is happening on the ground. The team would verify that in practice the program is enrolling each tier's target population, providing the specified services, and achieving the outputs specified in the model.

We will combine information from the various data sources in order to do this. The SIC method allows us to quantify *whether* and *when* key aspects of implementation occurred. This is useful not only because it will pinpoint challenges and delays quickly to inform program adjustments, but the measures can also be leveraged in the outcomes study to empirically measure how fidelity affects the outcomes.

Qualitative Data Analysis: Immediately following focus groups and key informant interviews we will compile notes, comparing them to recordings for consistency, to prepare for preliminary analysis with NVivo. We will code and analyze data focused on specific topics corresponding to relevant domains. For example, to inform process analyses, we could prioritize coding on information collected on implementation versus effectiveness. We will have one consistent coder and will regularly check the work to ensure reliability and resolve any disputes about what the data say.

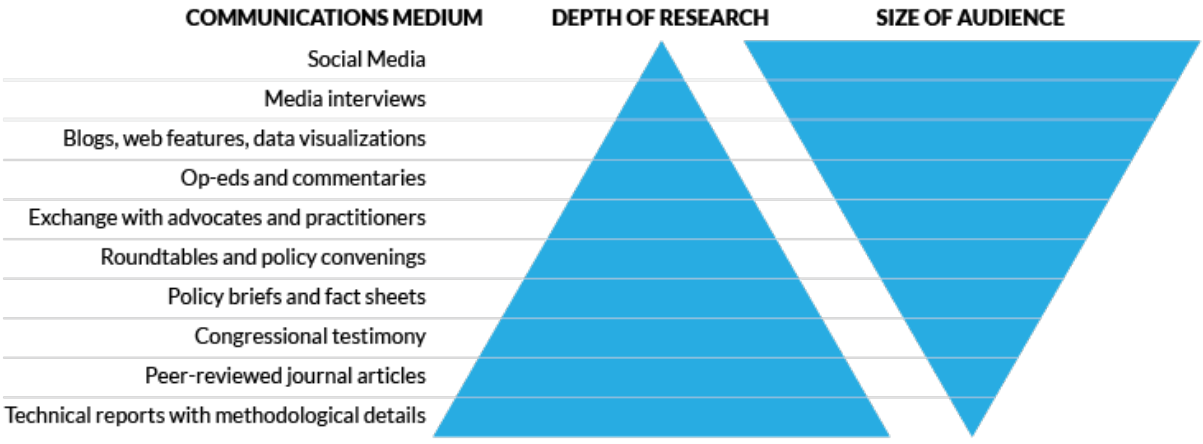
Survey Data Collection and Analysis: If funding is secured for a survey, survey data will be recorded in Qualtrics to facilitate our ability to analyze the data quickly. We will conduct descriptive analyses of the survey data, stratified by tier and program receipt. Findings from the survey will also be used to inform the areas of inquiry focused on during qualitative data collection efforts. For instance, if we learn that individuals opted out of the program because they did not understand what it would offer them, we will be certain to include questions about participant perceptions of initial engagement and what could have been improved.

Administrative Data

Administrative Data Analysis: The outcomes study will make primary use of the data warehouse, while the process study will focus on collecting qualitative data using the above methods as well as quantitative information from the Hello Baby client management system.

11 Describe your approach to data dissemination and use of findings

The research team has a great deal of experience tailoring documents to fit the targeted audience. As standard practice, Urban researchers use varied vehicles to promote research, including our web site (www.urban.org), our *Urban Wire* blog (www.urban.org/urban-wire), e-newsletters, and social media channels. Where appropriate we draw on the expertise of the communications team to incorporate specialized media techniques, including data visualizations, digital storytelling, and other innovative strategies designed to reach a wider range of audiences.



Meeting the goals and needs of our funders and maintaining the highest level of client satisfaction are central to our mission. Research reports, briefs, and other large deliverables undergo an extensive internal review process for quality assurance. Draft final reports are

reviewed at multiple levels, including peer review for substance and quality. Key deliverables are reviewed by project senior advisors to ensure the research meets high substantive and technical standards. An on-staff editor copyedits and formats all reports and briefs and ensures technical reports comply with government regulations on release.

The proposed research team benefits from close connections in the field of primary prevention research and extensive experience disseminating research to multiple audiences including federal, state, and local policymakers; researchers; technical assistance providers; and practitioners and frontline staff working with children and families. Wulczyn and Lery have presented the evaluation findings from Allegheny's FSC initiative to a variety of audiences, such as the National Family Support Network and earlier this month at the Society for Social Work and Research.

Proposed products

After each site visit, we will deliver a memo to DHS with a summary of findings, and feedback and recommendations based on those findings. Such memos might address: what to refine in the engagement protocol for the priority tier and other implementation challenges. We will contribute to DHS's semi-annual progress reports to ACF and will issue a final evaluation report. We will also seek to publish one public-facing product each year on the Urban and the State Center websites.

In addition to the proposed products, we plan to support the project team at grantee meetings in Washington, DC. We will use these meetings as opportunities to disseminate results and to network and collaborate with other researchers in the grant cluster.

We will seek opportunities to disseminate findings to key audiences, including policymakers and practitioners, through briefs, webinars, and blog posts with organizations in our early childhood network. In addition, we will submit proposals to present study findings at a range of conferences that include a focus on child well-being, such as the Society for Social Work and Research and the ZERO TO THREE National Conference.

12 **Please identify potential contextual and methodological challenges and your approach to addressing or minimizing them**

One methodological challenge will be reaching families who opt out, perhaps particularly among the priority group. A major objective of Hello Baby is to discover what it takes to engage families, so we will put extra emphasis on contacting families who opt out and offering them $50 compensation to participate in an interview, survey, or focus group.

A conundrum that Hello Baby raises involves its voluntary nature. This includes both the decision to opt-out of the PRM and the decision to opt-out of differentiated or priority services. The decision to opt-out of the PRM may be sensitive to how the information is conveyed – literally the bedside manner of the hospital outreach worker. It will be important for the process study to capture how this is done, including the wording used, and report back to the project team early and often about how many opt-outs are occurring both at the hospital and via postcard. We plan to facilitate that feedback by arranging frequent data transfers from DHS (birth universe) and the hospital outreach team (opt-in/out data) to the research team for processing and reporting at the CQI meetings. We also plan to link those sources to the data warehouse and hospital records about the births to describe what characterizes families who opt-out. Finally, we will attempt to find out why families opt out, both via a (potential) survey sent to everyone approached at the hospital (and anyone who may be missed), and by including a short multiple choice question on the opt-out postcard sent home with the mothers, asking the reason(s) for opting out.

Opting out of offered differentiated or priority services poses another conundrum. By definition, many of babies that the PRM identifies for the priority group will go on to be placed into foster care by age five, and more will have some interaction with child welfare. When that happens, how will DHS consider the fact that the family refused Hello Baby? While not directly related to

the evaluation, we raise the issue to note that information collected during the process study might inform DHS's plans for how to address the matter.

Another challenge will be for DHS, FSCs, and service providers to meet the expected increased demand for services under Hello Baby. What will happen if a Hello Baby family (in any tier) is referred to a substance abuse treatment program in their community, but either there is no availability or the program is of low quality? DHS may want to consider adding more slots, improving the quality of existing ones, or both. The evaluation will employ stakeholder interviews to estimate to what extent the mix of services (quality and capacity) changes over time, and whether certain communities see more service enhancements than others.

Budget (10 points)

13 **Include a budget that reflects a realistic estimate of the costs associated with the evaluation (you may attach your budget as a separate document, which will not be included in the page count).**

Attached

14 **Provide a budget narrative that reflects a realistic estimate of the costs associated with implementing the Program.**

This budget represents our best estimate of the effort required to carry out a process and outcomes evaluation. We are prepared to adjust according to the expectations of DHS and ACF, as the planning period unfolds. The budget reflects the mix of expertise and experience we have assembled to constitute the team that can maximize productivity across the tasks. We believe the fundraising requirements over and above the funding available per the RFP will be easily accomplished, given the track record of Urban and the Center. We have developed a contingency budget of $300,000 direct cost that we can share with DHS, should we be selected.

The cost data is represented by year and/or task and detailed line items. The following notes correspond to each of the line items included in the Urban budget and are intended to demonstrate cost reasonableness. Our objective is to provide our Client with the best value for the funds available.

The budget includes a line item to show the total direct costs ($509,209). This was calculated by deducting the fee, on-site personnel indirect, subcontract (Chapin Hall) indirect, subcontractor administration, and general & administrative from the total estimated cost plus fixed fee amount.

DETAILED COST NOTES

PERSONNEL

Research Staff: This line item identifies Urban Institute staff costs either by name or by general labor category, budgeted to undertake the work on this project. The budget shows for each category of staff the projected corresponding cost.

Fringe Benefits: The Urban Institute provides a benefit package to attract and retain quality staff. That package includes annual, sick, and holiday leave, health insurance, a retirement plan, term life insurance, and statutory required benefits. Fringe benefit costs are calculated as a percentage of salary costs and will be computed and charged to the Project at the fringe rate calculated for each year of the project.

SUBCONTRACTORS

We have budgeted $407,276 for a subaward to Chapin Hall to lead the outcomes evaluation. Actual cost of the subcontract will be charged to the project.

TRAVEL

We have budgeted $12,557 over the 3 years of the project. This includes 3 round trips between San Francisco and Washington DC to attend grantee meetings, 4 round trips between

Washington, DC and Pittsburgh for a kickoff meeting and site visits, and 4 round trips between San Francisco and Pittsburgh for a kickoff meeting and site visits.

OTHER DIRECT COSTS

Printing, Postage/Delivery, and Telephone. These costs are budgeted based on prior experience. Charges are estimated based on labor budgeted in the proposal and will be charged at actual cost incurred under the project.

Rent, Researchers' Offices. The Urban Institute allocates rent expense monthly based on an algorithm applicable to the labor costs of on-site research staff specifically working on the research effort. This algorithm is based on anticipated rent expense and is reviewed annually. We calculate rent using an algorithm based on salary since one's salary at Urban directly relates to the size/square footage office staff are provided. Thus, each grant only contributes a pro-rata share of the total Urban Institute rent expense based on the amount of time research staff work on the project.

The result of the algorithm is labor cost multiplied by .0069.

Computer Network Services. Urban maintains an integrated computing environment which includes a Hewlett-Packard (HP) Proliant DL580 and an HP 3Par 7400 iSCSI SAN to handle heavy-duty research computing. This includes SAS statistical software, for projects with large computational or data requirements. Urban operates a fully integrated Windows Active Directory environment for file sharing and printing. Urban runs two Internet connections, a high-speed 1 Gbps fiber-optic connection from Cogent Communications, and a 100 Mbps connection from Lightower. Using the Border Gateway Protocol (BGP) on our Internet routers, we have automatic failover between the two Internet connections. This not only benefits Urban staff members, who often need to use the Internet in their work, but it also benefits those who visit the Institute's web servers or who send e-mail to the Institute.

Urban strives to preserve data integrity and security. A firewall monitors and evaluates all attempted connections from the Internet to our public web servers and our private network. An up-to-date anti-virus software runs on desktop PCs and servers. A spam filtering appliance quarantines spam messages and also performs virus and spyware checking. We implement other "best practices" for securing servers and desktop PCs.

Charges for Information Services are estimated based on two factors - labor and the number of hours budgeted in the proposal, and will be charged based on actual incurred compensation costs. Urban's IT services benefit not only the overall activity of the Urban Institute, but are customized to address the needs of each individual project. Urban tracks and charges its computer costs based on each project's usage.

Participant Compensation– We have budgeted $4,500 total for focus group incentives in each year of the project.

Catering – We have budgeted $900 total for catering costs for the focus groups to be held each year.

Books/Periodicals/Library Services. These costs are estimated based on the labor budgeted in the proposal and will be billed based to the project at actual cost.

Subcontract Administration: The Subcontract G&A is an integral part of the Institute's Indirect Cost Rate Structure and is audited and approved by our cognizant audit agency as reflected in our NICRA. Accordingly, the rate of 4.37% must be applied to all subcontract/ sub agreements issued by the Institute.

GENERAL & ADMINISTRATIVE

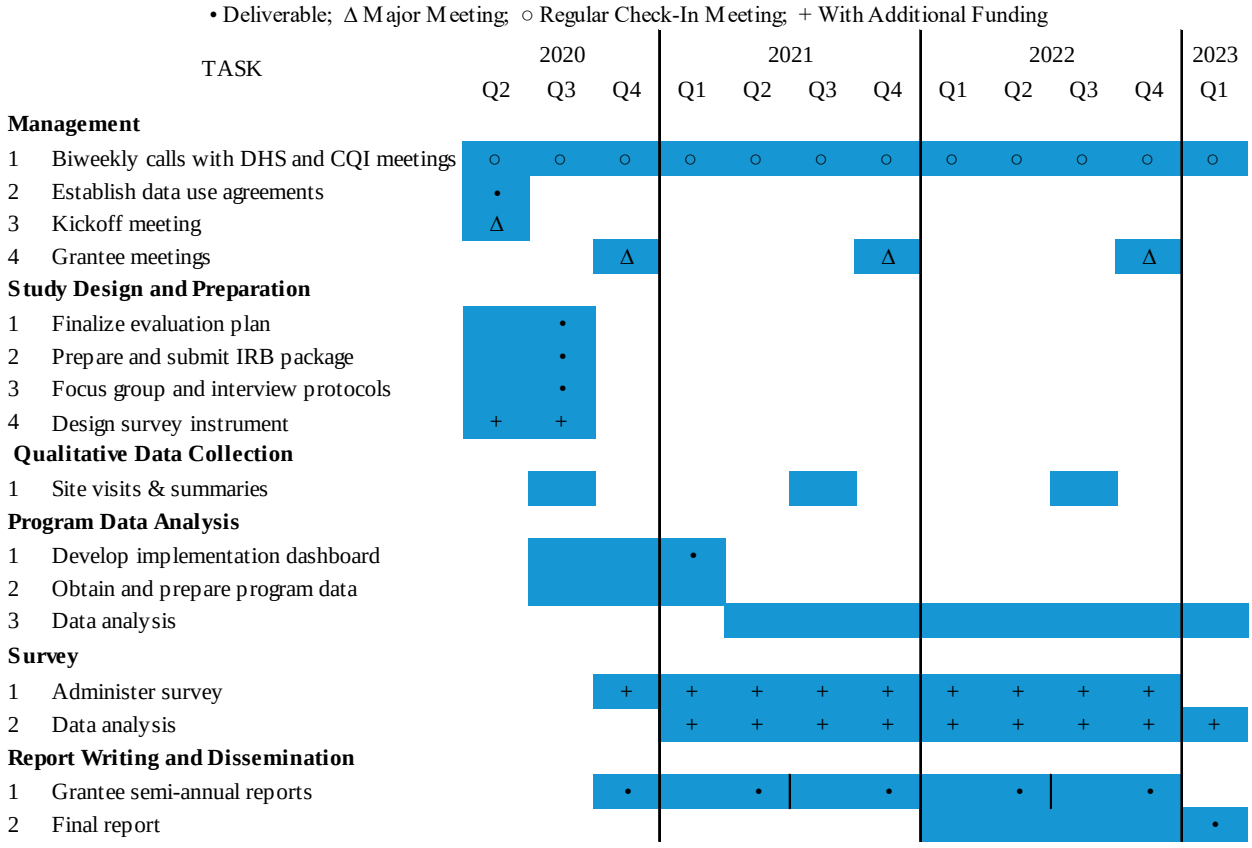
The Urban Institute's G&A rate is an integral part of our Indirect Cost Rate Structure and is audited and approved by our cognizant audit agency as reflected in our NICRA. Accordingly, the General & Administrative rate of 22.1% is applied to all federal project work prepared by the Institute.

FEE

Urban's Board of Trustees has, since its organization in 1968, authorized charging fee on work performed under contracts (but not under grants) for the purpose of maintaining working capital and paying for project-related expenses which are not reimbursable under the Federal Acquisition Regulations. It is the policy of the Urban to charge a fixed fee of 7% on cost reimbursable contracts.

Project Management (10 points)

15 Provide a timeline, with all associated deliverables, that demonstrates how the project will be completed within the project period.



16 Describe your procedures for maintaining quality control and timeliness and conducting regular reviews of the quality of data collected.

Our proposed CQI meetings with project partners form the backbone of our approach to quality control and the rapid identification of challenges. We use this platform in other projects to gain buy-in about the evaluation from service providers and other staff that have data entry or other data collection responsibilities. The regularity of these meetings builds rapport with project partners and in turn, we believe it motivates them to produce better quality program data. An implementation dashboard, which we will produce regularly, will drive the core meeting agenda items, allowing the team to troubleshoot problems early.

Our research team includes a strong project management function. Team members are specifically familiar with the complex timelines, deliverables, and communication protocols required under federal grants and contracts, and we use a cache of project management tools and methods to stay on track.

Process and Outcome Evaluation of Hello Baby, a Tiered Intervention Program for New Parents and Their Babies

<u>Section C – Impact Evaluation</u>. If you are proposing to perform the Impact Evaluation, respond to the items below. If you are not proposing an Impact Evaluation, leave this Section blank. Your response to this section should not exceed 15 pages.

Understanding the Scope of Work (30 points)

1 **Provide an overview of your approach to the project tasks, including explanations and rationales for any suggested modifications.**

Our project plan is organized around the five tasks outlined in the RFP plus one we have added:

1. Collaborate with ACF

We understand that as a significant funder, ACF has a significant investment in Hello Baby. We also understand the leadership at the Children's Bureau is particularly interested in primary prevention as a strategy for reducing impact of maltreatment on the children and families. To advance the national impact of the lessons learned from Hello Baby, close coordination with the national evaluation team will be essential.

2. Document participant experiences by tiered level.

This is one of the principal challenges of the Hello Baby initiative. Predictive risk modeling is an evolving technology and its benefits are not fully understood. The evaluation has to clearly establish how risk group classification can be used to expand service uptake, without coercion. Our approach is designed to track the decision process that families undergo. We use randomization of recruitment scripts to understand how a PRM in combination with targeted recruitment strategies affect the willingness of families to uptake services and persist. At each point in their decision process, we ask parents to reflect on their awareness of services and whether they see participation in services as beneficial to them as parents focused on the well-being of their children. Of course, understanding how answers to these questions differ by tiered level is a central concern.

3. Assess implementation facilitators, barriers and strategies to address those issues.

Our implementation (process evaluation) takes a dual perspective: (1) implementation barriers, facilitators, and strategies and whether the program is 'installed' as planned and (2) how is the program delivered, given what was installed, as a function of facilitators, barriers, and strategies. The Stages of Implementation Completion (SIC) method offers a highly reliable, practical, and concrete way to pinpoint implementation milestones in child welfare programs. We will develop detailed SIC measures to answer the research questions about implementation (How did the process of care change and was the capacity developed to carry out the process?) and fidelity (How well did those changes play out?).

4. Monitor participant outcomes using both secondary and primary data.

We propose a comparative time series analysis tied to administrative data using a *multi-level discrete time model* to address three challenges that come with monitoring participant outcomes: baseline variation in the expected rate of maltreatment, time varying changes in important independent variables (e.g., marital status, neighborhood of residence, number of children), and the timing of service relative to the risk of maltreatment and service uptake. Our approach identifies all children born, with careful tracking of children into tiered risk and service groups. The comparative time series analysis is also well-suited to the problem of detecting surveillance effects.

5. Use the above to understand engagement, retention, and attrition and the connection to outcomes

Our approach provides a robust, economical evaluation strategy. The proposed randomization strategy provides a mechanism for understanding whether decision-making considerations are correlated with enrollment, engagement, persistence, and outcomes. We will have a clear

understanding of the decision process and how, across tiered groups, those decision processes affect engagement, retention, attrition and outcomes.

6. Use the above to provide active feedback to stakeholders as input to the CQI process.

Our orientation to evaluations of this sort combines rigorous methods, statistical analysis, and a keen awareness of how evaluation *must* inform the evolution of the program over time as implementation plans are modified by timely feedback. Our aim is to help DHS address program improvement built on close to real time evidence about outcomes. With stakeholders, we aim to develop a clear understanding of the theory of change, the process, quality, and capacity improvements that define Hello Baby relative to business as usual and the mechanisms that generate outcomes so that we can work with stakeholders to generate the evidence needed to adapt Hello Baby to its strengths and weaknesses. Although thumbs up/thumbs down evaluations are important to evidence-building in the long run, we see a more dynamic learning agenda in play in Allegheny County over the nearer term.

2 Describe your understanding of the Hello Baby target populations and risk factors

The Hello Baby program uses a predictive risk model (PRM) to identify new parents at substantial risk for future child welfare involvement and post neo-natal death. The program offers support for three specific populations, as determined by the PRM: low risk families; families that have a higher level of need than those deemed low risk; and priority level families, who have the greatest level of need. Previous work in Allegheny County demonstrates that uptake remains low in the face of a strong network of prevention services, suggesting families at the greatest risk of harm are not being adequately served. This underutilization is thought to be a consequence of many factors, including geographic proximity to service centers, transportation challenges, social isolation, lack of outreach, and limited support networks.

Moreover, research suggests that key correlates of child abuse include untreated maternal depression or mental illness, substance use disorder in caregivers, age of the mother, and intimate partner violence in the home, as described in the Hello Baby services RFP. As such, families in the priority group may face numerous and overlapping service needs that the peer/social worker teams will help address. How they do so, using the COACH model, as well as the extent to which families participate in and benefit from such services will be important to capture in the process evaluation.

Finally, up to 40% of priority families will be active in child welfare at the time of birth and more will likely become active within the evaluation timeframe. Other families will already be enrolled in an evidence-based home visiting program, and may not want to be inundated or burdened with another program or services. Variability of child welfare and other program involvement in the target populations offers an opportunity to leverage connections to existing systems, while also ensuring that families are not inundated or receiving duplicate services.

3 Describe your understanding of the field of primary prevention of maltreatment, as well as issues specific to Hello Baby.

Primary prevention of maltreatment is directed at the general population on a universal level, attempting to avoid maltreatment in the first instance. As such, all members of a given community have access to and can potentially benefit from the offered prevention services. Although designed to address secondary prevention, the Family First Prevention and Services Act (FFPSA) has generated a renewed emphasis on early intervention and on primary prevention services as well.

DHS's interest in prevention is not new. Over a decade ago, Wulczyn and Lery began partnering with DHS to evaluate several child welfare programs meant to prevent secondary and tertiary maltreatment (e.g., Inua Ubuntu, Systems of Care Initiative, Hi Fidelity Wraparound, Family Group Decision-Making). Our evaluation of the FSCs' impact on maltreatment investigation rates found promising results, namely that the FSCs appear to have a protective effect in communities. That study contributed to DHS's thinking about how it can optimize that resource

for secondary and tertiary prevention, but also to deliberately expand the reach of FSCs for primary prevention.

Programs like Hello Baby take a universal approach to raising the awareness of the general public and normalizing expectations around seeking assistance for new parents. This type of primary prevention approach also offers a better chance of engaging a greater proportion of the highest risk families for child maltreatment, potentially decreasing the services necessary downstream for these families. Yet while primary prevention services may play a part in increasing the engagement of high-risk families, there is still reason to believe that many high-risk families may opt out of voluntary services. Hello Baby may not reach many of the highest risk families if they choose to opt out at the hospital, via postcard, or during subsequent engagement efforts by the peer/social worker teams. Capturing engagement and attrition, as well as when they both occur, will be a key part of both the process and outcomes evaluations.

Moreover, the primary prevention approach of Hello Baby raises unique challenges around the PRM score and potential stigmatization of some families. First, some families may opt out bedside at the hospital because they do not want their information used to target them for eligibility. Second, because the PRM requires certain administrative information to calculate a risk score, children not known to the system or who don't have enough data in the Data Warehouse will not receive a score. According to a 2019 report about Allegheny's other PRM, the AFST, approximately 10 percent of referrals were not generating a risk score, highlighting children who may be at risk of maltreatment but missed by the PRM. Third, while no system can remove the chance of falsely screening-in children based on their risk score, the PRM does provide the potential for increasing stigma if scores are not closely guarded and call screeners are not adequately trained on the possibility of false positives and confirmation bias. Both the impact and process evaluations will play a key role in understanding how the use of the PRM to target families plays a role in who opts in or out, and why.

4 **Describe the existing evidence base in terms of key features of effective implementation and outcomes for primary prevention programs.**

Implementation

In a 2011 brief, Deborah Daro outlined a set of principles demonstrated to be effective across diverse disciplines and service delivery systems. These lie at the core of effective interventions, and include: a strong theory of change that demonstrates clear pathways, specific outcomes, and curriculum content; recommended duration and dosage that is systematically applied to all enrolled in services; a clear, well-defined target population, as well as clear eligibility criteria and outreach/engagement strategies; a strategy for guiding staff in balancing content delivery and respect for the family's culture and circumstances; a method to train staff on model delivery; a supervisory system to support direct service staff; reasonable caseloads that allow direct service staff to accomplish core program objectives; and the systematic collection of process implementation data to ensure standards of model fidelity are met (Gateway, 2011).

These principles are echoed in a review of 500 quantitative studies that evaluated preventative intervention programs, in which the authors identified five pillars that significantly influence the implementation of a prevention program. These pillars include *community characteristics, provider characteristics, innovation characteristics, the prevention delivery system* (i.e., features related to organizational capacity) and the *prevention support system* (i.e., training and technical assistance). In order to implement an intervention successfully, variables in all five pillars must interact in a constellation to support successful implementation (Durlak & DuPre, 2008).

Outcomes

It is often difficult to establish that a primary prevention program did in fact reduce maltreatment directly, because it is hard to know whether the program served parents who would have maltreated without the program. (Hello Baby intends to minimize this problem using the PRM to target families.) Instead, most programs seek to affect an interim outcome – either individual or community-level – that in turn may reduce the likelihood of maltreatment. Programs focused on

individual-level outcomes include parenting programs and therapy, which seek to change attitudes, knowledge, or behavior. The notion of collective parenting is generally not favored in our society (Daro, 2016), yet some approaches do draw this causal path to maltreatment, seeking to boost social capital and collective efficacy in a community – constructs measured only at an aggregate level – in order to create the broad social supports that can help parents protect and raise children (Kim, Gloppen, Rhew, Oesterle, & Hawkins, 2015).

5 Identify any logistical challenges and describe how they will be addressed or minimized.

We expect it may be difficult to reach individuals who decline to participate in Hello Baby. The outcomes study will use the Data Warehouse inputs for the PRM to understand what characterizes families who opt out. But it will also be important to hear directly from some families who opt out in order to gain insight into why. Though we anticipate this challenge, we have extensive experience reaching out to and engaging individuals who have declined program participation. We will make multiple contacts, offer an incentive, and ensure that all research staff involved in this effort are adequately trained in respectful client engagement to maximize the success of our efforts. Though the best chance of gleaning insight into why families opt-out might be a short survey (see question 8 below), we will consider hosting a focus group with individuals who have recently opted out of the program. These opportunities are sometimes more appealing because of the increased incentives. In addition, we have found that some people like the opportunity to provide more unstructured feedback. As described in the data collection section, we will report engagement numbers on the implementation dashboard and discuss them at CQI meetings to identify problems quickly.

As with any cross-system effort, we anticipate logistical challenges during implementation. The research team has extensive experience evaluating programs that involve cross-system coordination. We know what challenges to expect and how to set up CQI checks to identify even the problems we don't expect. For example, we will need to work with the hospital and DHS to craft a specific protocol for giving the research team information about those who opt out, and the protocol may differ for those who opt out at the hospital verses later by postcard. As described in the data collection section, we will develop a regular data transfer process with DHS to help keep track of this task and identify problems early.

6 Describe your experience partnering with other researchers and stakeholders

Our research team is always aware of the benefits of drawing on the expertise of others. For some examples:

- Urban and the State Center have partnered for two decades on child welfare projects.
- Bridgette Lery partnered with the researchers who developed Allegheny's AFTS and PRM, on a PRM project in California. Lery is planning a new project with Putnam-Hornstein and Vaithianathan in Douglas County, Colorado related to their PRM.
- Urban partnered with technical assistance providers, the Children's Bureau, and local evaluators at all five sites to evaluate a five-year federal demonstration project testing a supportive housing program for child welfare-involved families who were homeless. Lery was a local evaluator at one of the sites.
- Urban recently completed a large evaluation of enhanced prenatal care programs for Medicaid beneficiaries, funded by the Center for Medicare and Medicaid Innovation (CMMI), in which we partnered with three other firms that had complementary skill sets. This work received exceptional reviews and a special commendation by CMMI leadership. Sarah Benatar led the process study.

We value the partnerships developed with stakeholders throughout the research process. These relationships are critical to the success of program evaluations in particular because they are the gateway to honest and clear information that informs evaluation conclusions. Our priorities in

these relationships rest on the promise of confidentiality, the explicit acknowledgement that these projects are not and should not be designed to point fingers at any individual or agency in particular and are instead about benefiting the providers and beneficiaries for whom the programs are designed to help. Additionally, ongoing recognition of the additional burden that this work can heap on community stakeholders, and efforts to minimize that burden while highlighting the ways in which this research can be beneficial, are at the center of all of our stakeholder interactions.

An important stakeholder group consists of community members. When doing research in communities on its members it is essential to value the needs and concerns of that community. Operating under the principle of "nothing about us, without us," we will take the time to understand the preferences and needs of the communities affected by Hello Baby – particularly the differentiated and priority programs – through our planned process study methods such as key informant interviews with FSC and DHS staff. This approach will always benefit the research, and consistently informs our implementation and impact evaluation work.

Technical Approach (30 points)

7 Please describe your organizations conceptual framework that addresses the components of the process and/or impact evaluation

Ever at the forefront of innovation, Allegheny County's Hello Baby initiative tests whether a predictive risk model (PRM) opens a service pathway to families who are unlikely to seek help raising their children even though the evidence suggests they will have a difficult time doing so safely. The Department of Human Services would like to increase service utilization among all families, but especially among the highest risk families, all with the goal of reducing child maltreatment. Doing so is, however, harder than it sounds. Hello Baby services are voluntary. Because there is no way to compel families to take advantage of what Hello Baby offers, DHS has to rely on what amounts to the art of strategic persuasion.

To that end, Hello Baby offers a tiered range of services in general alignment with family risk profiles. The family of every newborn will get a light-touch message that promotes the help families might use to make raising a youngster less stressful. For the families most likely to find parenting really stressful, DHS wants to deploy an even more deliberate approach to service engagement paired with service access and more careful attention to what families need through robust case management.

DHS has opted to focus on the youngest among us. The emphasis on infants (and toddlers) represents a sound epidemiological choice. Early childhood is a profoundly important developmental period; service investments that offset the risk factors will likely have long-term developmental benefits. Moreover, infants and toddlers are, almost without exception, the children most likely to encounter the child welfare system through the child protective doorway. In the Family First context, interventions that reduce the risk of maltreatment for this group of children have enormous scale-up potential across the country. The question is whether the approach Allegheny DHS has chosen elevates service engagement and whether the services provided to engaged families lower the risk of maltreatment.

Although the operational details are anything but simple, the causal model at the heart of the Hello Baby initiative is really quite straightforward. Much like other jurisdictions around the country, Allegheny County realizes that maltreatment rates – the number of children abused and neglected per 1,000 children – would be lower if families availed themselves of the services located in their community. Therein lies the rub. To be a user of services, a family has to know about the service, understand they have a problem to solve, *and then* regard use of the service as in their best interest relative to the alternatives and their own sense of what the future holds. In the judgment of Allegheny DHS, too few families—especially families at the high-risk end of the continuum—engage with services long enough to lower the risk of maltreatment in a

meaningful way, an outcome of family decision making that leads to maltreatment rates that are above otherwise preventable levels.

As for a remedy to this longstanding problem, Allegheny DHS has come up with a nested strategy designed to 1) raise service participation rates among families on the one hand, and 2) improve outcomes among those families that take up the service on the other. In the words of Allegheny County, the "people who most need services are not using them" for reasons that have to do with access, limited knowledge, engagement, isolation, and false assumptions about the root causes of maltreatment. Consequently, in the Hello Baby design, there are three interventions: one that addresses the decision making of potential participants, a second that improves case management, and a third that improves the quality of the services enrolled families get.

The causal model underlying the first feature of Hello Baby is easy to spot. Let's assume that each family has a latent propensity to participate in a program. For some families, it is relatively easy to activate that propensity; for other families, convincing them that services have utility is more difficult. In the Hello Baby model, DHS is betting that deliberate, coordinated outreach will, in effect, persuade families to opt-in to the program. The outreach raises the family's awareness of assets in the community, a strategy that carries with it the hope that families will want to know more. The second leg of the Hello Baby model ties case management to a stronger connection between the family and perceived value of the services. Taken together, outreach and service quality are meant to raise the willingness of families to engage. If these phases of the intervention work, families should be more likely to opt-in and enroll in the program.

Of course, enrollment in a service is one thing; engagement with services long enough to benefit from those services is an altogether different matter. Because the Hello Baby services are essentially voluntary, whether to persist with services is a judgment the family has to make based on their calculated self-interest, all else being equal. It starts with awareness – I have a problem and could use help – but the decision ultimately hinges on the trade-off between time and effort expended using services and the benefits derived. Short of a clear benefit, the propensity to continue participation could well wane until it shuts down altogether because families simply do not or cannot see the benefit. To address this concern, Hello Baby expects to prioritize access to the best services the county has to offer, with the hope that the gap between a family's realized benefit and the expected benefit will narrow. If it does, there is reason to believe that families will persist with those services long enough to derive the benefits a family needs to lower the incidence of maltreatment.

8 Please describe your evaluation design and approach to sampling families

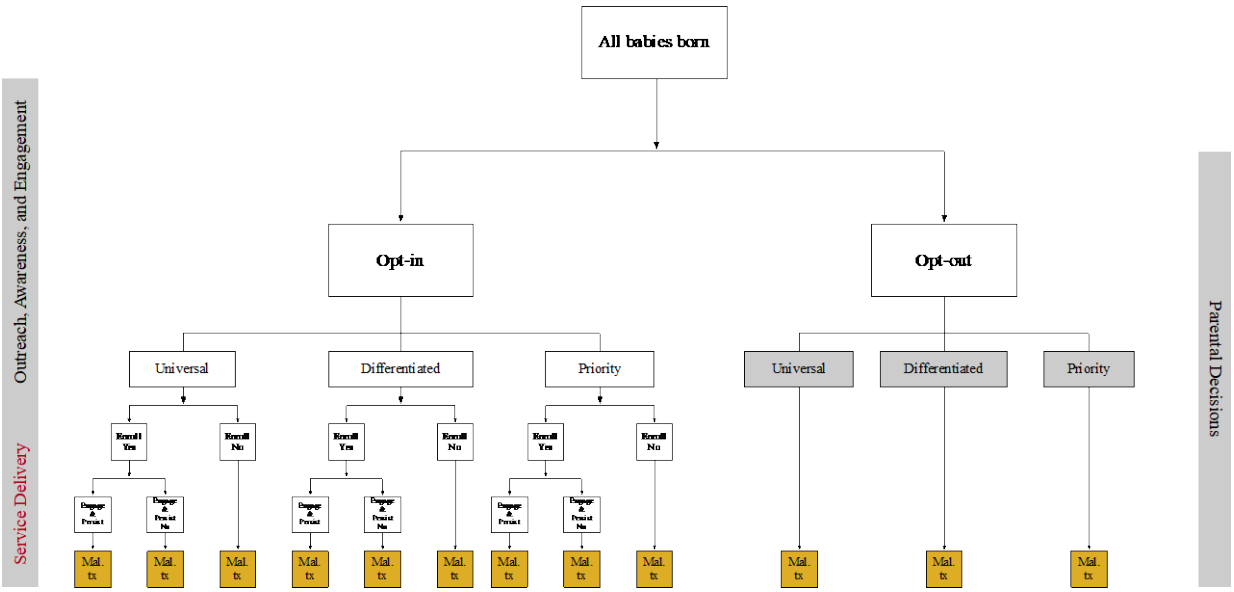
For the overarching evaluation design, we propose a quasi-experimental approach to test the impact Hello Baby has on maltreatment outcomes and a random control trial (RCT) of recruitment strategies. The quasi-experimental design (QED) would test whether Hello Baby overall has the intended impact on maltreatment rates; the random assignment of families with babies born during the target years would test whether different recruitment strategies influence take-up rates.

The rationale for our design choices are laid out graphically in Figure 1 below. In the diagram, we acknowledge that the starting point begins with *all* babies born in a given year. The parents of those children will make decisions about opting into Hello Baby following an invitation from DHS, a process that starts in the hospital. From opt-in, we follow children down pathways defined, in part by risk-group membership. Regardless of the risk group, it is possible for any group member to seek services or not. We define service use as enrollment and engagement with persistence. Much as opting in is a parental decision, enrollment and engagement with persistence represent parental decisions. The vertical bar to the right spans the portion of the map that pertains to parental decision-making. The specific decisions are highlighted in light red.

The vertical bar on the left represents the interventions. In the early stages of the decision-making, DHS hopes to upgrade outreach, awareness, and initial engagement. Success will boost opt-in and enrollment decisions across all three risk groups. Earlier we referred to this as persuasion: DHS will do what it can (short of coercion) to tilt parents towards opting in and enrolling. From there, the decision to engage and persist is motivated by service quality, which DHS hopes to improve through case management and other strategies plus the parent's sense that the support they're getting is worth their time and effort.

Lastly, as each family moves through their decision-making process, we have to focus on whether there is (ever) a report of maltreatment following the birth of the baby. Although we show the question of maltreatment as coming after a series of parental decisions, we know that maltreatment can come at any time post-birth. Ultimately, for each group shown, we want to know whether there was a report of maltreatment, whether the report was investigated, whether the investigation was substantiated, and when each of those happened. Depending on the availability of other data (e.g., data from hospitals regarding treatment for injuries), we are prepared to think beyond Allegheny's child protective services data for indicators of maltreatment in the covered population.

Figure 1: Design Map for the Hello Baby Evaluation



A key component of our design is the random assignment of families to groups who get different recruitment 'scripts.' Here, the script refers to the series of steps DHS takes to encourage parents to opt-in, enroll, engage, etc. The scripts might vary the wording of the invitation, the manner in which the invitation is presented (e.g., in person with SMS follow-up vs in-person only) or some other variation as decided by stakeholders after considering the operational implications. Every eligible family/child would be invited; only the script would vary, a design that would inform a key practice question: what is the best way to approach families and does the best approach depend on the risk group? The analysis of these data would be relatively simple. The dependent variable is the likelihood of moving further along the decision tree from opt-in to enrollment, from enrollment to engagement, and from enrollment to persistence. Random assignment addresses measured and unmeasured group differences. If the scripts themselves are designed with clear expectations regarding what stakeholders think will motivate parents, response rates that differ by script will provide key insights vis-á-vis enrollment and engagement.

As for our proposed evaluation of Hello Baby's impact on maltreatment, we do *not* see an opportunity to randomize families eligible for the priority services into treatment/no treatment groups post-enrollment. Similarly, for the universal and differentiated populations, because the

points of service in the community are so diverse, it would be operationally difficult to impose random assignment into treatment/no treatment groups. Regarding the priority service population, our perspective is based on the uptake estimates found in the RFP (160 families with service enrollment expectations well below that number), the heterogeneity of the eligible population in terms of their underlying vulnerabilities and the services they are likely to need (e.g., mental health service vs. substance abuse treatment), the exclusion of families with a high probability of placement, and the County's desire to serve everyone who is potentially eligible. Having said that, we are open to working with stakeholders on a random-assignment strategy if stakeholders see a way forward with that idea.[1]

In lieu of random assignment, a robust quasi-experimental design (QED) provides a reasonable alternative, especially given the extensive administrative data holdings controlled by DHS. Although there are multiple QED options, we think a comparative interrupted time series combined with some type of propensity score matching offers a practical starting point for the evaluation. The dependent variable in the model would be an agreed upon definition of maltreatment (e.g., a substantiated report of maltreatment). For the comparative time series, we would apply the PRM to an historical population of babies born before implementation begins so as to stratify families into like groups for the pre- and post-intervention time periods. To account for geographic differences in maltreatment rates – contextual effects – we would nest children into their neighborhoods and adjust the model results for differences in the level of social disadvantage and other features of the neighborhood such as service availability (Family Support Center vs. no Family Support Center, given the evidence that Family Support Centers affect reporting rates).

In the end, we aim to have a file that allows us to compare maltreatment rates (as defined) for babies born post-Hello Baby with a similar group of children from similar neighborhoods who would have been eligible for Hello Baby had those services been offered when those babies were born. In the most conservative test of the intervention, an intent-to-treat design (ITT) would examine maltreatment rates for all babies born during the Hello Baby implementation relative to comparable pre-Hello Baby birth cohorts. When stratified by universal vs. differentiated vs. priority, the ITT design would tell us about target group specific impact. For example, an ITT design on the families that enroll would tell us about the impact of Hello Baby regardless of engagement; an ITT design post engagement would tell us about the impact of Hello Baby regardless of persistence. For the effect of treatment on the treated (TOT), a propensity score match provides a solid foundation, especially if we are able to use what we learn about the ways in which families who enroll *and* engage are different from families who enroll and don't engage as part of the matching algorithm. This information would come from the random assignment paired with the administrative data.

Finally, a side benefit of the comparative time series analysis is the sustainability of the approach. Put simply, maltreatment of very young children will always be important. Going forward, DHS has to monitor what's taking place in the most robust manner possible. If the current data assets do that job well, the results will shape how DHS thinks about building out its data capture capability in the future. Establishing a self-sustaining evidence-building capability is one key to the long-term CQI efforts DHS will want to support.

Sample/Sample stratification

As a general matter, we see the sample for the Hello Baby evaluation as including all children born in the county, with a particular emphasis on births in the one or two target areas where the provider is working. We aim to cast a wide net because QEDs gain external validity with the

[1] Although we don't know the pattern in Allegheny County, in other urban areas around the country, a non-trivial fraction of children is reported to CPS authorities within 30 days of birth. Among other things, this means that for some families the events that lead to a maltreatment report happen quickly, before the family has been through the opt-in, enroll, and engage decision process, even if the goal is to have all of that happen relatively quickly. To the extent this dynamic is correlated with the risk groups, the reality affects whether an RCT is, practically speaking, feasible.

largest possible population. Specifically, we see using data for the children born in Allegheny for at least five years prior to the onset of Hello Baby at a minimum. Insofar as the marginal cost of adding years to the five is very small, we would recommend adding more years to the pre-Hello Baby data set because it would strengthen our ability to detect cohort, period, and surveillance effects on overall outcomes.

With this sample of children, we propose identifying their service trajectories through the administrative data. As already noted, among other benefits, this will tell us whether existing data support rigorous monitoring of outcomes within the CQI/EF framework. If these data need to be rounded out with additional data collection, we are prepared for that contingency as a strategic decision undertaken by stakeholders given the resources and priorities. The aim is to represent the service utilization trajectories for each family/baby pair over time, much as we did in our Family Support Center evaluation in Allegheny County. In this context, we see the service utilization as the behavioral manifestation of parental decision-making.

Given this sample of children, we think it is important to stratify the sample by maltreatment history. In the starkest possible terms, we have to ask ourselves what it means to prevent maltreatment. From the perspective of any given child, the question is whether *this* child has been maltreated. However, from the adult's perspective, the question is whether the adult/parent has maltreated this child or any of their other children. If the index child has yet to be maltreated but the parents have maltreated other children, providers of services are being asked to prevent the *recurrence* of maltreatment in the family. We strongly suspect that the baseline risk of maltreatment differs in the two populations, a factor that has to be taken into account in the evaluation design (e.g., how the matching variables are used). We also strongly suspect that clinical interventions have to take into account the distinction between families where there has never been abuse (even across two-generations) versus families where abuse happened at some point within the family system and the goal is to prevent a recurrence. The former case might be regarded as true prevention insofar as abuse has *never* been observed. All other cases involve the recurrence of maltreatment within the family system.

Although we regard measured service utilization as the manifestation of the parental decision process, we will not have evidence of the decision process itself – what motivates parents to enroll, engage, and persist. For this we propose a brief questionnaire targeted to a random sample of parents on both sides of the decision-process: enroll vs. not, engage vs. not, persist vs. not. Data collection will be timed as close to when the decision is made. If there is a clear decision point operationally (e.g., the provider has a clear point in the process when the family is asked for its decision), we can add a brief questionnaire that asks about the decision. The questionnaire would target awareness and the parent's judgment regarding the benefit of services to them, and their perception of risk. We are prepared to contemplate novel data collection schemes including SMS messages. Focus groups are also a possibility. The choices would be based on the practical realities of sample size at each decision point and other concerns as worked out with stakeholders.

One last sampling concern pertains to the areas where the program rolls out. We acknowledge that the initial the rollout decision may have been made by the time the evaluation starts. That said, given geographic differences in baseline maltreatment rates, there may be options for selecting the next area. We think there are strategic advantages tied to these neighborhood or context effects DHS may want to consider, especially if the baseline conditions within a given area affect the likelihood of success overall. In any case, we think it will be important to understand the geographic distinctions at baseline so that the pattern of variation is taken into account as a strategic consideration for program rollout and evaluation.

9 Describe your approach to data collection and measurement

In keeping with the action research paradigm, we think it is important to approach data collection and measurement from the perspective of the stakeholders. The role of the evaluator in action research is, we believe, to help the stakeholder group balance their sense of priorities as guided

by their theory of change, the practical realities of implementation, and the requirements of rigorous evaluation in an era of evidence-informed policy and practice. In the end, the evaluation design has to yield the highest quality evidence given the constraints on the ground. Our goal is to maximize the quality of evidence within the constraints of time, budget, and practicality, using transparent processes that actively engage the stakeholders.

We also think it is important to rely, to the extent possible, on existing data. As we have already noted, our strategy is guided by the desire to help public agencies and their stakeholders understand whether their existing data supports a rigorous CQI process post-formal evaluation. With that understanding in hand, our data collection and measurement strategy is organized around whether what is being proposed translates over the long-term into an approach DHS could adapt/adopt as part of its routine data collection, monitoring, and self-evaluation strategy.

Regarding data collection that involves contact with families, we want data collection to be simple with minimal burden placed on respondents, whether the respondents are professionals (paraprofessionals), stakeholders, or families.

10 Describe your analytic approach for primary qualitative and quantitative data and proposed use of the Data Warehouse

There are any number of ways data of the sort being collected might be analyzed. In our experience, we have found multi-level discrete time hazard models (MLDTM) as the most effective tool for analyzing binary (i.e., enrollment vs not; engage vs not) time-to-event data within a comparative time series framework. The reasons are two-fold. First, the MLDTM accommodates the nested structure of the data easily. As we have observed in virtually every study we have conducted, parameter estimates vary depending on geographically and/or administratively bounded areas. For example, we can expect the baseline maltreatment rates for identical populations to vary by neighborhood for reasons tied to the peculiarities of the neighborhood rather than the children (i.e., the effect of context). These differences have to be accounted for in the analysis of treatment effects. The MLDTM handles this situation easily and transparently. Second, many of the independent variables in the model of treatment effects vary over time: did the family stop or start receiving income support, change family composition, move from one part of the city to another? On either side of these changes, the underlying risk of maltreatment may differ. The discrete time model offers a transparent way to adjust the conditional probability of maltreatment (or engagement) given the absence or presence of factors that affect the outcome of interest. For example, if we measure persistence as the number of consecutive months for which a family attended a service, we can include the cumulative count as a descriptor linked specifically to the life course trajectory so that eight consecutive months of intervention in the eight months after the baby is born is distinguished from eight consecutive months of intervention in the eight months after the baby turns two. Because the risk of maltreatment differs during these developmental periods, the assessment of treatment affects has to place the onset of treatment and the end of treatment into a life course perspective. The MLDTM are, again, an easy transparent way for doing so.

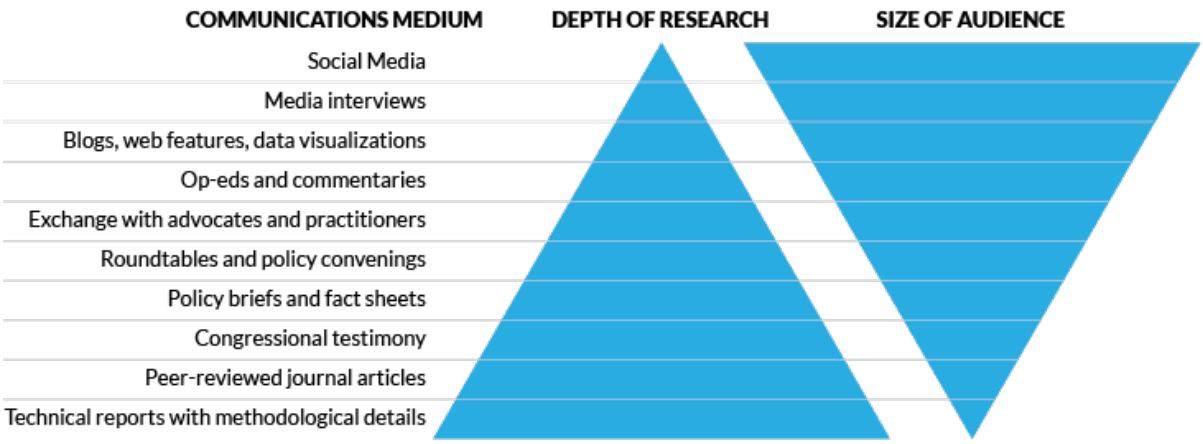
Regarding the analysis of qualitative data, we are guided by the same underlying philosophy – when something happens is often more important than if something happens. Questions posed to families have to be attuned to the family's own unique developmental stage and those of the children. The same can be said for programs and interventions. Interpretation has to be guided by a sense for the unfolding nature of a community's attempt to improve outcomes for children. That is another reason why a neighborhood or community perspective has to be brought to bear on the analytical strategy overall.

Although it does not extend to the data warehouse specifically, our experience with Allegheny's human services data is extensive (e.g., we have constructed integrated FSC encounter data, placement data, maltreatment data, and community data). In localities other than Allegheny, we have linked data pertaining to such programs as TANF, Medicaid, and education, to name a few. We have also used data from private agencies integrated with public agency data to evaluate the

interventions. Having built these types of files in the past the type, we are able re-purpose computer programs at considerable savings to the County. In short, our approach to administrative data is opportunistic and rooted in a long history of cutting-edge applications of those data to the type of study that will satisfy what DHS needs in this high-profile context.

11 Describe your approach to data dissemination and use of findings

The research team has a great deal of experience tailoring documents to fit the targeted audience. As standard practice, Urban researchers use varied vehicles to promote research, including our web site (www.urban.org), our *Urban Wire* blog (www.urban.org/urban-wire), e-newsletters, and social media channels. Where appropriate we draw on the expertise of the communications team to incorporate specialized media techniques, including data visualizations, digital storytelling, and other innovative strategies designed to reach a wider range of audiences.



Meeting the goals and needs of our funders and maintaining the highest level of client satisfaction are central to our mission. Research reports, briefs, and other large deliverables undergo an extensive internal review process for quality assurance. Draft final reports are reviewed at multiple levels, including peer review for substance and quality. Key deliverables are reviewed by project senior advisors to ensure the research meets high substantive and technical standards. An on-staff editor copyedits and formats all reports and briefs and ensures technical reports comply with government regulations on release.

The proposed research team benefits from close connections in the field of primary prevention research and extensive experience disseminating research to multiple audiences including federal, state, and local policymakers; researchers; technical assistance providers; and practitioners and frontline staff working with children and families. Wulczyn and Lery have presented the evaluation findings from Allegheny's FSC initiative to a variety of audiences, such as the National Family Support Network and earlier this month at the Society for Social Work and Research.

Proposed products

After each site visit, we will deliver a memo to DHS with a summary of findings, and feedback and recommendations based on those findings. Such memos might address: what to refine in the engagement protocol for the priority tier and other implementation challenges. We will contribute to DHS's semi-annual progress reports to ACF and will issue a final evaluation report. We will also seek to publish one public-facing product each year on the Urban and the State Center websites.

In addition to the proposed products, we plan to support the project team at grantee meetings in Washington, DC. We will use these meetings as opportunities to disseminate results and to network and collaborate with other researchers in the grant cluster.

We will seek opportunities to disseminate findings to key audiences, including policymakers and practitioners, through briefs, webinars, and blog posts with organizations in our early childhood network. In addition, we will submit proposals to present study findings at a range of conferences that include a focus on child well-being, such as the Society for Social Work and Research and the ZERO TO THREE National Conference.

12 **Please identify potential contextual and methodological challenges and your approach to addressing or minimizing them**

One critical issue is the consent to use administrative data in relation to opt-in/opt-out and other parental decisions. Because a QED is the most logical design for this study, it is important that the access to the *evaluation* data be as broad as possible. Parents, of course, have the right to limit access to data about them. Administrative data are, however, among the most important sources of the evidence needed to make genuine improvements in services and outcomes. In other words, parents stand to benefit from access, when services evaluation is the root purpose behind that access.

We see a clear distinction between the use of data for program reasons versus uses of the data for evaluation purposes. The former requires use of the data that is not per se protected – risk levels will be disclosed and action taken could affect clients directly. For the portion of the evaluation that relies on administrative data, confidentiality can be maintained because the uses of the data do not require disclosure of identifying information in the same way that those same data might be used by program staff to make decisions that affect the intervention.

We think, therefore, that the consent scripts and the role of the IRB in helping establish the line between research uses vs. program uses, is vital. In essence, one has to craft a consent script that allows a parent to say I restrict your access to information about me for purposes of the risk modeling and related uses, but does not restrict access to the data for purposes of evaluating the program. In the latter case these data are deidentified and managed with our usual confidentiality protections. If we deal directly with the source data and apply those data to analysis under the direct control of the evaluation team, we will make the argument to the IRB and others that the uses are distinct and therefore separable. We have, of course, the experience needed to work with the IRB to adapt novel approaches to data access, security, and client protection.

Budget (10 points)

13 **Include a budget that reflects a realistic estimate of the costs associated with the evaluation (you may attach your budget as a separate document, which will not be included in the page count).**

Click here to enter text.

14 **Provide a budget narrative that reflects a realistic estimate of the costs associated with implementing the Program.**

This budget represents our best estimate of the effort required to carry out a process and outcomes evaluation. We are prepared to adjust according to the expectations of DHS and ACF, as the planning period unfolds. The budget reflects the mix of expertise and experience we have assembled to constitute the team that can maximize productivity across the tasks. We believe the fundraising requirements over and above the funding available per the RFP will be easily accomplished, given the track record of Urban and the Center. We have developed a contingency budget of $300,000 direct cost that we can share with DHS, should we be selected.

Outcome Study

Task 1: Project Administration – this includes bi-weekly calls with DHS and other stakeholders, preparation of IRB materials, data sharing agreements, and routine data management tasks (extract, load, and manipulate). These activities are on-going over the course of the project.

Task 2: Write Research Plan – this task incorporates the development of the research plan for submission to ACF. It includes the development of an overall timeline for the projects, identification of data sources, and any decisions regarding methodological choices pertaining to the QED and the random assignment. Revisions to proposed design would be made here with stakeholders. This task is completed within 6 months.

Task 2.1: Data Capture (administrative data) – As implied, this task involves securing the administrative data and building the analytical files. The work is concentrated in the first project year.

Task 2.2: Data Capture (random assignment) – This task incorporates the work needed to carry out the random assignment for the study of parental decision processes. It occurs during the latter half of the first year and carries over into the next project year.

Task 2.3: Data Capture (original data collection) – This task incorporates the State Center's contribution to the survey work. This work coincides with the effort incorporated in the Urban budget.

Task 3: Data Analysis – The work completed under this task includes all data analysis and is on-going. It includes the effort needed to inform the CQI and implementation processes.

Task 4: Dissemination – This task covers the preparation of the interim and final reports and the effort needed to prepare to conduct webinars, prepare podcasts, and brief stakeholders as necessary.

Project Management (10 points)

15 **Provide a timeline, with all associated deliverables, that demonstrates how the project will be completed within the project period.**

TASK AND TIMELINE	Project Yr. 1												Project Yr. 2												Project Yr. 3											
Administration																																				
Communication with stakeholders (bi-weekly)	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Establish Data Sharing Agreements	x	x	x																																	
Submit protocol to Institutional Review Board	x	x	x																																	
Data Management	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
1. Write Research Plan																																				
Work with DHS to plan time frame for analysis	x	x	x	x	x	x																														
Work with DHS to define and identify data sources for outcome data from existing DHS data	x	x	x	x	x	x																														
Work with DHS to define and identify data sources for family level and individual level	x	x	x	x	x	x																														
Work with DHS to define and identify data sources for service utilization	x	x	x	x	x	x																														
Prepare and finalize research plan with implementation schedule	x	x	x	x	x	x																														
2. Data Capture - administrative data																																				
Collect data on utilization/dosage from participating agencies					x	x	x																													
Collect/verify additional data					x	x	x																													
Obtain and code administrative data					x	x	x	x																												
Data Linking							x	x																												
2. Data Capture - random assignment																																				
Design random assignment protocol					x	x	x																													
2. Data Capture - original data collection																																				
Interviews						x	x	x	x	x	x											x	x	x												
3. Data Analysis																																				
Analyze outcome data controlling for child and family characteristics					x	x	x	x	x	x	x		x	x	x	x	x	x	x	x	x	x	x	x		x	x	x	x	x	x	x	x	x	x	x
Fidelity analysis					x	x	x	x	x	x	x		x	x	x	x	x	x	x	x	x	x	x	x		x	x	x	x	x	x	x	x	x	x	x
Connect process study findings to outcome study					x	x	x	x	x	x	x		x	x	x	x	x	x	x	x	x	x	x	x		x	x	x	x	x	x	x	x	x	x	x
Integrate multivariate findings with additional data					x	x	x	x	x	x	x		x	x	x	x	x	x	x	x	x	x	x	x		x	x	x	x	x	x	x	x	x	x	x
Re-analyze data based on feedback					x	x	x	x	x	x	x		x	x	x	x	x	x	x	x	x	x	x	x		x	x	x	x	x	x	x	x	x	x	x
4. Dissemination																																				
Required deliverables																															x	x	x	x	x	x	x

16 **Describe your procedures for maintaining quality control and timeliness and conducting regular reviews of the quality of data collected.**

Original data collection largely falls under the purview of the process study so data quality procedures are discussed in the corresponding section. Regarding administrative data, we have well-established processes for building administrative data sets for research purposes. Those procedures include routine checks for out-of-range values, format errors and other anomalies. File building includes regular communication with DHS and other stakeholders regarding the contents of the files coupled with frequent discussions regarding file structure, the meaning of elements, and so on. Finally, at each stage of file building we produce a sample data set with selected records so that the data owner (e.g., DHS) can verify the rules applied and validate the results as aligned with in-house rules. The final step involves our submission of counts of cases (e.g., annual number of maltreatment reports) to compare with internal sources, with any discrepancy examined against processing rules until differences in counts are fully deconstructed and understood in their entirety. Final analytical files are used only after approval from the data custodian at the agency.