I. Background Information on Welcome Baby

Welcome Baby is a locally designed home visiting program for mothers and their newborns developed as part of First 5 LA's place-based community investment supporting low-income families with children ages zero to five broadly referred to as Best Start. The home visiting program focuses on education and support for pregnant women and mothers of newborns. Women are recruited into the program either prenatally or in the hospital after giving birth. The program includes up to nine engagement points for women and their infants, and provides breastfeeding support, guidance on infant health and development, and referrals to needed resources and services.

The model was first launched in 2009 in the “pilot” Best Start community—designated Metro LA—offering services to all women who gave birth at California Hospital Medical Center (CHMC) and lived within a five-mile radius of the hospital. The program is now being implemented in 13 additional Best Start communities in Los Angeles County. In 2013, local hospitals applied for grant funding through First 5 LA to provide Welcome Baby services either directly or through partnerships with a local community-based organization to women who gave birth at that hospital. As of December 2013, fourteen hospitals were selected among those that applied across the 13 new Best Start LA communities (as well as the pilot Metro LA community). Hospitals began hiring and training staff at different points in time, as early as May 2013, and began recruiting clients in July 2013.

A key difference between the Metro LA pilot and the current Welcome Baby initiative is that women will be screened postpartum in the hospitals where they deliver to determine their risk level and the intensity of the home visiting services they need. A modified version of the Bridges for Newborns screening tool identifies whether women are at high risk and in greater need of intensive services or whether they are low-to-medium risk. High risk women residing in Best Start communities will be offered enrollment in Select
Home Visiting services—one of three evidence-based home visiting models selected by the Best Start community, including Healthy Families America, Parents as Teachers, or Triple P. Women residing in a Best Start community who are not high risk (i.e., score below an identified cut-off) will be offered enrollment in Welcome Baby, which is lower intensity than the Select Home Visiting programs. However, high risk women who give birth at a designated Welcome Baby hospital who do not live in a Best Start community will be offered Welcome Baby Lite—a lower dosage of services up to two months postpartum (versus nine months for the full Welcome Baby model)—while lower risk women who live outside the Best Start communities will only be offered referrals to local community services.

In an effort to replicate Welcome Baby in a systematic way across the new communities, First 5 LA created a Welcome Baby “fidelity framework” that documents the key indicators that service providers must achieve to be implementing the home visiting program appropriately. First 5 LA also contracted for the creation of new data system, Stronger Families LA, that could be accessed by all parties involved in service delivery and oversight, including outreach specialists, hospital liaisons, program supervisors, parent coaches, and nurse home visitors. This new system will contain a multitude of data elements that can track the status of cases, family needs, services rendered, referrals made and completed, and immediate outcomes, such as child development, breastfeeding, and maternal mental health. Additionally, a new oversight entity—LA Best Babies Network (LABBN)—was retained to provide training to service providers and ensure services are implemented with high quality across all sites. These new components are expected to improve program performance and the quality of services provided as Welcome Baby expands.

Given First 5 LA’s large investment in home visiting, it desires evidence of how services are being implemented, what outcomes are associated with participation in Welcome Baby, and where improvements are needed to maintain a high level of fidelity to the model. This design report outlines options for conducting future evaluation research on the cross-site implementation and outcomes associated with Welcome Baby.

II. Purpose of the Study

The key purpose of any future cross-site study of Welcome Baby would be to: 1) evaluate implementation across L.A. County to inform program improvement/mid-course corrections; and 2) build the evidence base regarding the outcomes achieved by program participation.
A. Study Goals

The cross-site evaluation could potentially include two study components: an implementation evaluation and an outcomes evaluation. The two components would be composed of distinct, yet complementary, tasks and could address the following goals:

**Implementation Evaluation:**

1) To examine how the implementation of Welcome Baby develops and changes over time, and the associated costs.
2) To examine the extent to which providers implement Welcome Baby in a manner consistent with First 5 LA’s fidelity framework and the ways in which providers are working to meet fidelity.
3) To describe how clients participating in Welcome Baby experience the program.
4) To examine who participates in Welcome Baby—their characteristics and risk profiles—and determine whether Welcome Baby is reaching its target population.

**Outcomes Evaluation:**

1) To examine the short-term and long-term outcomes demonstrated by Welcome Baby participants.
2) To understand if and how variation in Welcome Baby service delivery (i.e., the service model, level of fidelity, timing of program entry, dosage, and service quality) correlates with client outcomes.
3) To examine variation in outcomes based on client risk profiles and characteristics (e.g., by marital/partnership status, race/ethnicity, education level, home language/English proficiency).

Both evaluation components would address the issue of program fidelity—the former from the perspective of implementation and the latter as a program performance outcome.

B. Research Questions

A cross-site evaluation could address the following research questions:

**Implementation Evaluation:**

1) To what extent have organizational factors at sites affected implementation of Welcome Baby? For example:
   a. Overall operations at the hospital, work climate, culture, communication, leadership’s support for Welcome Baby
   b. Integration of Welcome Baby into existing routines and programming for infants and families
2) What factors account for variability in sites’ ability to reach fidelity to the Welcome Baby model?
3) What are the program-level costs of implementing Welcome Baby? How do costs change over time?
   a. Start-up costs
   b. Costs related to ongoing delivery of services
4) To what extent have TA providers effectively built staff capacity through training and technical assistance?
5) To what extent do staff feel knowledgeable, skilled in and positive towards the Welcome Baby model?
6) How are sites, TA providers and F5LA learning from each other and developing networks for sharing lessons learned? To what extent are sites sharing and connecting to Best Start Community Partnership work?
7) What are the experiences of staff in implementing the Welcome Baby model? For example, what are experiences with:
   a. Outreach and enrollment
   b. The Modified Bridges for Newborns screening tool
   c. Implementing the WB curriculum during engagement points
   d. Identifying community resources and making referrals
   e. Supervisory processes
   f. Client-centered approach
8) What are the experiences of participants in receiving Welcome Baby?
9) What are the demographic characteristics of Welcome Baby participants, and do they differ from what was expected (e.g., higher or lower risk, different demographic mix, etc.)? What groups is Welcome Baby not reaching?

Outcomes Evaluation:

1) What features of service models and implementation (i.e. program infrastructure/hospital run vs. hospital + CBO; program fidelity; service quality/client-centered approach) are associated with key child and family outcomes?
2) Is there a difference between outcomes for women with various dosage levels (post-partum WB only, prenatal + post-partum WB, WB-lite, variation in engagement points completed due to clients)?
   a) Which engagement points are associated with positive outcomes? Are some visits more important than others for certain outcomes (e.g., breastfeeding, child development, discipline strategies)?
   b) How many engagement points are needed to demonstrate positive outcomes? Is there a minimum number of engagement points needed?
3) How do client characteristics (i.e., marital/partnership status, race/ethnicity, education level, home language/English proficiency) correlate with child and family outcomes?
4) What are the correlations between client risk profiles, referrals made, whether services (i.e., mental health, substance abuse, etc.) were received, and child and family outcomes?
   a) What is the reason why services were not received?
III. Design Options for an Implementation Study

A multi-year evaluation of Welcome Baby in the Metro LA Pilot Community included a rigorous qualitative component, including case studies of implementation, focus groups with home visiting consumers and providers, and structured observations of Welcome Baby nurse and parent coach home visits. Building off of this effort, a preliminary study of early implementation of Welcome Baby in 10 new Best Start Communities was started in March 2014, and will use key informant interview results to describe the experiences of First 5 LA officials, training and TA providers, and Welcome Baby hospitals and community-based organizations in launching the initiative during its first year. (A final report from this preliminary study is expected in June 2014.)

A future cross-site implementation evaluation could continue to examine the implementation of Welcome Baby across all 14 Best Start LA communities in the coming years. One option would be to examine a single year of implementation—after all 14 communities have Welcome Baby up and running—at the start of the evaluation. This study could focus on identifying the steps local provider sites took to launch their Welcome Baby programs, build partnerships and internal capacity, hire and train staff, recruit clients, and provide home visiting services. Topics could include challenges with adhering to fidelity, screening women reliably with the Modified Bridges for Newborns tool, and recruiting a sufficient number of women.

However, hospitals began to hire and train staff and recruit clients at different time periods over the course of approximately one year from 2013 into 2014—forming several cohorts of Welcome Baby hospitals—and thus the sites may continue to be at very different stages of implementation during the early years of the evaluation. Therefore, an alternative option would be to examine the ongoing implementation during the first several years of implementation.

Once the initial issues with program start-up have been addressed, the evaluation could focus on program performance and success in reaching program goals. Topics could include how programs have adjusted to collecting and using data in the new Stronger Families LA database and how that database is utilized; additional training that program staff need to successfully serve families’ needs; areas where the Welcome Baby model is working well or not; and challenges with linking women to the community services they need (such as mental health services that may be difficult to find).

The design options for the implementation evaluation include both primary data collection and analyses of secondary data from administrative records and the U.S. Census. These options are listed in Table 1 and each is described in this section.
A. Primary Data Collection Options for an Implementation Study

Potential evaluation options requiring primary data collection include: case studies of Welcome Baby program implementation (based primarily on key informant interviews); focus groups with clients; structured client observations; surveys of Welcome Baby staff; and client satisfaction surveys. These options could be used in combination or separately, and conducted once or repeatedly during the course of the evaluation to measure change during the different stages of implementation.

Table 1. Design Options for an Implementation Study of Welcome Baby

<table>
<thead>
<tr>
<th>Data Collection Methods</th>
<th>Potential Frequency/Timeframes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary Data Collection</strong></td>
<td></td>
</tr>
<tr>
<td>Case studies of Welcome Baby programs</td>
<td>Annual or semiannual</td>
</tr>
<tr>
<td>(through including key informant interviews)</td>
<td></td>
</tr>
<tr>
<td>Focus groups with clients</td>
<td>Once during course of evaluation, alternate years, or annually</td>
</tr>
<tr>
<td>Structured client observations</td>
<td>Concurrent with case study site visits (if shadowing home visitors only); ongoing during course of evaluation (for more rigorous sub-options)</td>
</tr>
<tr>
<td>Surveys of Welcome Baby staff</td>
<td>Once during course of evaluation, alternate years, or annually</td>
</tr>
<tr>
<td>Client satisfaction surveys</td>
<td>Once or twice during course of evaluation</td>
</tr>
<tr>
<td><strong>Secondary Data Analysis</strong></td>
<td></td>
</tr>
<tr>
<td>Analysis of client characteristics and</td>
<td>Ongoing during course of evaluation</td>
</tr>
<tr>
<td>program fidelity using Stronger</td>
<td></td>
</tr>
<tr>
<td>Families LA Database</td>
<td></td>
</tr>
<tr>
<td>Analysis of hospital data records and</td>
<td>First or second year of evaluation; one year to complete</td>
</tr>
<tr>
<td>Stronger Families LA Database to</td>
<td></td>
</tr>
<tr>
<td>determine program reach</td>
<td></td>
</tr>
<tr>
<td>Analysis of Welcome Baby costs</td>
<td>First or second year of evaluation; one year to complete</td>
</tr>
</tbody>
</table>

1. Case Studies of Welcome Baby Programs in the Best Start Communities

An important component of any implementation evaluation would be case studies of Welcome Baby program implementation. Case studies would provide an understanding of the design and implementation of Welcome Baby in the 14 communities (or more as communities are added) and enhance other evaluation findings by providing insight into
the different service models being implemented, adherence to fidelity, and the barriers or challenges program providers are facing. Given the diversity of the Best Start communities and the addition of the Bridges to Newborns risk screener, the case studies could explore the characteristics of the clients being served in Welcome Baby (as well as those screening into the Select Home Visiting programs), whether provider are reaching their target populations, and who the program is not reaching. The findings from the case studies would point to elements of the programs that may be correlated with success, which can be further examined quantitatively in an outcomes evaluation. This study component would generate early and regular feedback to First 5 LA, and provide context for interpreting the evaluation’s other qualitative and quantitative findings.

Case study methods would begin with a detailed review of program documents, and then focus on key informant interviews with program directors, providers, and other staff. These interviews, optimally, would be held in person during site visits but could also be conducted by telephone. Key informants could be interviewed once per year, or twice a year if more intensive monitoring of implementation was desired. For example, half the Welcome Baby provider sites could be visited in the fall and the other half in the spring so each is visited once a year, or all sites could be visited in both fall and spring. The sites that began implementation the earliest could form cohort 1 and be interviewed first, whereas the late starters could form cohort 2 and be interviewed later when they are at a similar point in implementation. The benefit of interviewing once a year is that some time would elapse between interviews to allow for the providers and program managers to report on experiences during the past year. More frequent interviews could track some subtle changes that might go unreported and could ultimately feed into program planning, but they would be more costly and may not produce findings that are very distinct from what could be gathered on an annual basis.

Key informant interviews would involve the use of tailored, semi-structured interview protocols to guide in-depth conversations with key stakeholders, allowing for the systematic collection of information on home visiting implementation, challenges, and successes. The scope of the interviews would concentrate on Welcome Baby implementation in the Best Start communities, but could also be designed to include the implementation of Select Home Visiting and Welcome Baby Lite (the latter of which is offered to only high risk women residing outside of the Best Start communities). Key informants could include Welcome Baby program directors and clinical supervisors; Welcome Baby providers (including hospital liaisons, outreach specialists, parent coaches, and nurses); hospital administrators; representatives from the Best Start community partnerships; and First 5 LA staff responsible for overseeing Welcome Baby, Select Home Visiting, and Best Start community partnerships.
2. Focus Groups with Welcome Baby Clients

Another design option that could be built into the case study approach is focus groups with Welcome Baby clients. Focus groups have the ability to enrich other evaluation components in several ways, by providing valuable detail about the experiences and concerns of mothers (and potentially their partners) receiving Welcome Baby. Insights from the focus groups would likely also enhance interpretation of other case study findings. Perhaps most important, focus groups could bring the perspectives of beneficiaries themselves into the evaluation, complementing findings from any quantitative analysis included in an outcomes evaluation.

Focus groups could be conducted once during the course of the implementation evaluation, during alternate years of implementation, or even annually to track potential changes in client experiences once programs have reached their full capacity and are implementing with high fidelity. If conducted at multiple time points, the location of the focus groups could vary each time so that clients from each Welcome Baby site could be included at some point during the study. The focus groups would concentrate on women who receive the Welcome Baby, but could also include comparison groups of women, including those who receive Select Home Visiting and/or Welcome Baby Lite. Additionally, the focus groups could include women who have recently completed the program as well as women who are currently enrolled. To accommodate the language diversity of clients, focus groups would likely be held in both Spanish and English.

Groups of 8 participants are optimal for focus groups, but to ensure the participation of at least 8 participants per group, approximately 12 women would be recruited. Recruitment strategies could vary based on the different types of groups proposed and the communities in which they are being held. To help with recruitment and attendance, each parent could be offered a gift card incentive, and light food and refreshments. Other possible incentives that could be offered include on-site child care and reimbursement of transportation costs.

3. Structured Client Observations

To better understand the content and methods of care delivered by Welcome Baby providers, another design option could be to conduct structured observations of home visits. Observations could be conducted in several different ways depending on the purpose and desired level of intensity.

The first option could be to replicate the approach implemented in the Metro LA Pilot Community Evaluation wherein evaluators shadowed several home visitors as part of that evaluation’s case studies. Researchers would observe home visits delivered at different engagement points (e.g., during a prenatal home visit, onsite with an outreach specialist, during a postpartum home visit with a nurse or parent coach). Particularly sensitive visits such as the hospital visit would likely not be appropriate to observe. Structured
observations would only be conducted with the full consent of Welcome Baby clients and parent coaches, using an IRB approved process designed to afford Welcome Baby participants with as much privacy as possible. An observational tool would be used to record the content covered (e.g., breastfeeding education, health insurance enrollment), any assessments conducted (e.g., Ages and Stages Questionnaire, PHQ depression screener), materials shared or used during the visit, the quality of the interactions and any conversation that occurs between the parent and home visitor. This approach would give a snapshot of whether Welcome Baby protocols were followed and how they were implemented at each stage of implementation.

For a more rigorous examination of the quality of Welcome Baby implementation, and particularly, the execution of the family-center approach that is core to the model, there are several additional options for client observations.

Instead of including client observations as part of an annual case study, an alternative approach could be to conduct a more systematic assessment of the quality of the interactions between parent coaches and clients. Field staff would be trained to conduct ongoing observations of home visiting staff. Each home visitor could be observed 2-3 times a year during the multiple years of the evaluation as part of ongoing performance measurement. Observers could either use a validated tool such as the Home Visit Rating Scale (HOVRS; Roggman et al., 2008), which measures the quality of home visitor strategies and effectiveness at engaging the family during the home visits, or design a new tool specific to Welcome Baby to assess the quality of the family-centered approach. This option adds the benefit of having external evaluators observe and assess home visit implementation on a more regular basis to examine changes in staff performance over time. However, this option is likely to be significantly more costly in terms of staff time to collect the data. Moreover, this option requires field staff to be trained on the Welcome Baby model and protocols so they know how to reliably rate high-quality interactions within that context.

If such field observations prove to be too costly or intrusive to home visiting staff, another option would be to have Welcome Baby supervisors collect data during their periodic staff performance review. LABBN has designed a tool that supervisors can use to measure the family-centered approach with the purpose of providing feedback and training to the observed home visitor. This tool may be insufficient for evaluation purposes in which case that tool could be redesigned or a new tool could be developed that supervisors could use. The advantage of this option is that supervisors are experts on the Welcome Baby model and know how to identify strengths and weaknesses in staff approaches. The disadvantage is the added burden that may be required of supervisors to conduct the observations at regularly scheduled times. Additionally, supervisors may have a more biased perspective of their own staff and may have difficulties collecting the data systematically for the purposes of research.
A final alternative option could be to have home visiting staff (parent coaches and nurses) periodically (2-3 times per year) video-record their home visits. Staff would be given camcorders to bring with them on a randomly selected visit and, with the parent’s consent, the parent coach and/or nurse would set up the camcorder at the beginning of the home visit to capture both the mother and the parent coach/nurse. The home visit would be conducted as usual. The videos would then be sent to the research team. A reliably trained team of analysts would review and code the videos following an established coding scheme to measure the interactions between parents and home visitors. The benefit of this option is that it would impose less of a burden on data collectors and supervisors and be less intrusive during visits, since a second visitor would not have to be present. However, the video could alter the behavior of the home visitor and/or parent if either of them are not comfortable being video-recorded.

4. Home Visiting Staff Survey

Another design option that would more systematically measure the perspectives and experiences of Welcome Baby program staff could involve the administration of a staff survey. The survey could be self-administered; staff could either complete it on-line or fill out a mail-in paper version. Given that the target population would be staff, the expectation is that it would be possible to convince a high percentage of staff to complete an on-line or paper survey without having to resort to a much more labor intensive and costly telephone survey. More importantly, a self-administered survey would elicit more candid responses as employees may be less likely to reveal concerns and personal dissatisfaction with the program to a person interviewing them. Another advantage of an on-line survey is that the survey could be sent to all staff members or at least to a large sample of staff members without seeing big increases in costs.

While response rates should be fairly high, a sample of the staff that did not complete the survey could be contacted by phone and encouraged to complete the survey. The sample of staff that completed the survey only after a phone call reminder would help in assessing the representativeness of the full sample. In addition, a phone call reminder should increase the overall response rate.

The development of the survey questionnaire would include a review of similar staff surveys, cognitive interview testing of questions, and piloting or pretesting. A goal would be set to keep the questionnaire length reasonably short (about 15 minutes) as respondents tend to provide less cognitive effort when surveys are too long. The questionnaire would need to include some questions about the amount of training received and how many client interactions (and type of interactions) a staff member has experienced to better understand the staff members’ responses.

Staff surveys completed during the implementation stage of the Welcome Baby program would provide a useful benchmark measure and could be used to help make timely
adjustments to the program. A second round of staff surveys—or even annual surveys—could be administered after the program becomes fairly stable and would therefore provide a useful comparison to the survey done during early implementation of the program to see if any areas of concern have arisen or been resolved.

5. **Client Satisfaction Surveys**

In addition to (or instead of) conducting focus group with clients, a client satisfaction survey could be administered to collect information from a larger sample of Welcome Baby clients. Similar to the staff survey, clients could complete a self-administered survey either on-line or by filling out a mail-in paper version, though with this population we would expect a lower percentage of on-line completes. Given the target population would be women who enrolled in the program the expectation would be that it should be possible to convince a high percentage of these women to complete the survey without having to resort to a much more labor intensive and costly telephone survey. Again an important reason for choosing a self-administered survey would be to elicit more candid responses as the respondents might generally be less likely to reveal concerns and personal dissatisfaction with the program to a person interviewing them.

For the client survey, a small incentive payment could be offered to obtain reasonably high response rates. Also, for those women who do not complete the survey, phone calls could be made to encourage them to complete the survey. Unlike the staff survey that could be offered to all staff, for this option, drawing a stratified random sample of clients who enrolled during a particular time frame would likely be more appropriate rather than potentially trying to offer the survey to all women. The Stronger Families LA Database would provide a good source for selecting a stratified sample of women. More intensive efforts could then be made to encourage a smaller selected sample to participate. In addition, stratifying the sample by variables such as birth hospital and timing of enrollment (prenatally or postpartum) could support potential subgroup analyses by site and prenatal program exposure. A final sample size of 400 should provide useful results which would in all likelihood require selecting around 650 women (assuming a 60% response rate) to interview.

Another possibility would be piggy-back the client satisfaction survey onto another evaluation component, such as a follow-up survey with Welcome Baby clients (described in the section on Design Options for an Outcomes Study) or site visits for a case study. However, the potential bias should be considered of having the respondent complete the satisfaction survey at the same time that they are being including as part of a larger overall evaluation.

Similar to the development of the staff questionnaire, instrument development would require a review of similar client surveys, cognitive interviewing of test items, and piloting.
or pretesting. A goal should be set to keep the questionnaire length reasonably short (about 15 minutes) as respondents tend to provide less cognitive effort when surveys are too long.

A client satisfaction survey completed during the implementation stage of the Welcome Baby program would again provide a benchmark measure and could help program directors with making some program adjustments. A second round of client survey could be done after the program becomes stable to provide insight into ways the program is maturing.

B. Secondary Data Analysis Options for an Implementation Study

In addition to collecting new data for the evaluation of Welcome Baby implementation, a number of options using available secondary data could further strengthen study findings. Potential data sources include the Stronger Families LA Database, hospital data records, and Welcome Baby cost records. Each of these options is outlined in this section.

1. Analysis of Client Characteristics and Program Fidelity Using the Stronger Families LA Database

To describe the clients being enrolled and served through Welcome Baby, and to identify providers’ adherence to key elements of model fidelity (and areas in need of improvement), one design option for an implementation study would be to analyze administrative data from the Stronger Families LA Database (which is currently in final stages of development). The Database will track the history of each case from initial intake, to assignment to a nurse and parent coach, to program exit, and includes all data collected on Welcome Baby clients (e.g., demographics, risk profiles, engagement points completed, referrals received, and scores on assessments). These data are entered by home visiting staff at all levels (e.g., outreach specialists, parent coaches and nurses, clinical supervisors, etc.), with each user having access to different fields based on their position. Under this design option, markers of program fidelity—based on the Welcome Baby Fidelity Framework—could be identified in the database and tracked over time. Examples include caseload size, staff training, completion of engagement points, follow-up on referrals, and use of a family-centered approach (according to supervisor performance ratings on the LABBN-designed fidelity monitoring tool). Specifically, these data could be aggregated by site at periodic assessment points to examine trends in program performance during the course of the evaluation. The fidelity markers could be identified early in the evaluation so that de-identified data by local sites could be extracted for ongoing analysis.

The Stronger Families data are also useful for describing the Welcome Baby clients themselves. Potential fields that would be useful for describing the client population and estimating differences among subgroups of clients include:

- Demographic characteristics (mother’s age, race/ethnicity, language, marital status, education level, employment);
- Residential information (if the client lives alone, with family, with non-relatives); and
• Family risk profile (Modified Bridges for Newborns score and other risk items).

2. **Analysis of Hospital Data Records to Determine Program Reach**

Getting aggregate data on women delivering babies in hospitals in the Best Start communities would provide a strong indicator of whether or not the Welcome Baby program is serving the intended population. These data could be compared to descriptive information about home visiting participants’ demographic characteristics obtained from the Stronger Families LA Database. The database includes descriptive information on all women who are screened with the Bridges tool—those who are referred to Select Home Visiting, those referred to Welcome Baby, and those referred to Welcome Baby Lite (who are high risk, give birth in a Best Start hospital but do not live in a Best Start community). Hospital records would likely include fields such as mother’s age, father’s age, marital status, education level, employment status, whether the birth was the mother’s first, whether the mother had prenatal care, and whether the mother and child were insured. These fields are also available in the Stronger Families LA Database.

Getting data on hospital clients—particularly those who do not enroll in any home visiting programs—would likely require building strong rapport with the hospitals and providing in return some useful feedback to the hospitals. If implemented, this option may require intensive initial coordination with hospitals so it is expected that data would not be available for analysis until after a minimum of several months; therefore, this task should be initiated early on during the implementation study, even if analysis does not occur until later in the first year of the evaluation.

In addition to looking at the hospital data, Census tract data for the communities would provide some supplemental social and economic information about the communities that the Welcome Baby program is serving and how the women being served fit into the community. These data could be accessed at any time and similarly compared to the demographic characteristics of the population of Welcome Baby clients.

3. **Analysis of Welcome Baby Costs**

As part of an implementation study, another option would be to examine the costs associated with Welcome Baby, including start-up costs for a new site and ongoing program operations. Start-up costs may include hiring and training staff, purchasing computer hardware and materials, and staff time to coordinate procedures for recruitment and screening of clients. Ongoing operational costs may include personnel salaries, staff development and training, supplies, gas and mileage reimbursement, and costs for maintaining the program’s Stronger Families LA Database.

Assuming good cost records are available (potentially from hospital invoices), the analyses could focus on determining what levels of intervention are most cost-effective. Then, depending on differences in dosage received by clients a detailed cost analysis could
provide some insight as to the optimal level of services that should be provided. These data could also be helpful in making staffing decisions, how many families can be effectively served, and the ideal level of services per family. The cost analysis could potentially uncover unexpected costs, which in turn could lead to making more immediate adjustments to the Welcome Baby program.

**IV. Design Options for an Outcomes Study**

An outcomes evaluation could also include both primary data collection and analyses of administrative data. Potential design options are listed in Table 2. These research tasks could begin at different timeframes during the course of the evaluation. Certain administrative data may take time to access while primary data collection would require initial planning, instrument development and testing, and participant recruitment. Each of these options is described in detail in this section.

<table>
<thead>
<tr>
<th><strong>Table 2. Design Options for an Outcomes Study of Welcome Baby</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Collection Methods</strong></td>
</tr>
<tr>
<td><strong>Primary Data Collection</strong></td>
</tr>
<tr>
<td>Follow-up survey of Welcome Baby clients</td>
</tr>
<tr>
<td><strong>Secondary Data Analysis</strong></td>
</tr>
<tr>
<td>Analysis of Welcome Baby clients’ initial outcomes using Stronger Families LA Database</td>
</tr>
<tr>
<td>Analysis of Welcome Baby Lite clients using Stronger Families LA Database</td>
</tr>
<tr>
<td>Analysis of linked administrative data from WIC, Medi-Cal, and DCFS records</td>
</tr>
<tr>
<td>Secondary data sources for local comparison estimates</td>
</tr>
</tbody>
</table>

**A. Primary Data Collection Options for an Outcomes Study**

A principal option for an outcomes study that would require primary data collection would be a longitudinal, follow-survey of Welcome Baby clients. However, such a survey could be executed in a number of ways, as described in this section.
1. Follow-up Survey of Welcome Baby Clients

To measure long-term program outcomes past the 9-month home visit, one option would be to conduct a longitudinal, in-home survey of a sample of Welcome Baby clients. The purpose of the survey would be to examine: 1) the health and well-being of mothers and their children after leaving the program; and 2) the associations between the variation in family’s Welcome Baby experiences and their outcomes. The specific program experiences of interest could be the service model used (hospital-run or hospital plus CBO), level of program fidelity, timing of enrollment (prenatally or postpartum), dosage (number and timing of completed engagement points), and quality of experiences (based on parental report and supervisors’ performance ratings of parent coaches). These program-level data (besides the reports of quality) would be accessible through the Stronger Families LA Database and could be linked to survey data for the purposes of analysis.

A follow-up survey could be administered at one or more time points, but to estimate long-term outcomes on children and their families, a survey would follow participants at least 2 years, and potentially even 3 years, post-intervention (i.e., roughly 24 and 36 months post-partum). Surveying families at multiple time points during early childhood development would allow for the examination of both sustained and emerging, “sleeper” outcomes. The Metro LA Pilot Community Evaluation surveyed families at 12, 24, and 36 months postpartum to capture short-term outcomes within three months of exiting the program (e.g., breastfeeding duration) and long-term outcomes of child health and development (e.g., BMI; social competence; communications skills), that were precursors of school readiness, an ultimate goal of Welcome Baby. A similar design with even assessment intervals could be implemented to capture outcomes in late infancy, toddlerhood, and early preschool-age. If funding prohibits the inclusion of three survey waves, then the 12- and 24-month survey (thus following up at 3 months and 15 months post-intervention) would offer some estimates of associations between clients’ experiences and short- and intermediate-term outcomes.

Survey sample. To estimate associations between service model (i.e., hospital-run versus hospital plus CBO), the sample should include women from multiple sites with varying service models. A minimum number of hospitals should be included to recruit a sufficient sample size during recruitment period. Given current enrollment numbers across sites, a minimum of six sites would likely be sufficient in recruiting the targeted sample—three sites that provide direct services to clients and three that partner with a community-based organization to deliver services.

To estimate differences in client outcomes based on program exposure (specifically, whether women had the opportunity for prenatal home visits in addition to postpartum visits), the sample could be drawn to include a minimum number of women who enrolled in their first or second trimester, their third trimester, and postpartum. Since fewer women enroll prenatally than postpartum, and even fewer enroll early in their pregnancy, a simple
random sample of women would make it difficult to estimate differences in outcomes by timing of enrollment. Therefore, one option would be to draw a stratified random sample of women to essentially oversample prenatal enrollees. Stratification variables could include birth hospital (to sample women within the hospitals where they deliver to have relatively equal representation) and timing of enrollment in Welcome Baby (to sample women who enroll in the first or second trimester, third trimester, or postpartum).

For illustrative purposes, Table 3 provides the margin of errors for various measures under different sample size assumptions. Specifically the table shows what the margin of errors would be for a measure that had an estimate of about 50%, 10% or 90%, and for a mean estimate based on a 5-point scale with a standard deviation of about 1.

The number of enrolled women is the number of women that is expected to be in enrolled into the Welcome Baby program in the Best Start LA communities over a one-year period. For illustrative purposes an estimate of 2,800 is used, which would need to be revised once a more accurate estimate of projected enrollment is obtained. The survey sample size column is the number of completed outcome surveys. The lower half of the table is useful in that it shows the margin of error for a sub-group that makes up about 10% of the overall population (n=280). The survey sample size needed for each subgroup is indicated in the second column, ranging from 25 to 150 per group depending on the measure and margin of error.

The design effect adjustment shown in Table 3 is 5%, though this is something that varies by measure and could end up a little higher or lower depending on the final sample design. However, unless a simple stratified random sample of all enrolled women could be drawn, some design effect is expected as a result of dividing women into sampling strata (i.e., birth hospital and timing of enrollment). If there is a strong intra-class correlation of outcome survey estimates by hospital or timing of program entry, higher design effects are expected. Once the final stratification and sample design procedures are set, it should be possible to get a reasonable estimate of the expected design effect by looking at demographic and other characteristics of the enrolled women by sampling strata.

The final three columns show the margin of error at the 90, 95, or 99 percent confidence level. For example using the 95 confidence level, if the estimate was 50% and the margin of error was .04, than one could be 95% confident that the estimate is no less than 46% and no greater than 54%.

Based on this table, an overall sample size of 400 woman would yield a margin of no worse than plus or minus 5 percent. Also, if a targeted sub-group is included that equals 10% of all enrolled women (e.g., first/second trimester enrollees), then a sub-group sample size of 100 would yield a margin of error no worse than plus or minus 7 percent.
Table 3. Margin of Error for Various Measures and Sample Sizes

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Number of Enrolled Women</th>
<th>Survey Sample Size</th>
<th>Design Effect</th>
<th>Margin of Error (90%)</th>
<th>Margin of Error (95%)</th>
<th>Margin of Error (99%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion 50/50</td>
<td>2800</td>
<td>600</td>
<td>0.05</td>
<td>0.031</td>
<td>0.037</td>
<td>0.049</td>
</tr>
<tr>
<td>Proportion 90/10</td>
<td>2800</td>
<td>600</td>
<td>0.05</td>
<td>0.019</td>
<td>0.022</td>
<td>0.029</td>
</tr>
<tr>
<td>5-Point scale (SD=1)</td>
<td>2800</td>
<td>600</td>
<td>0.05</td>
<td>0.035</td>
<td>0.042</td>
<td>0.055</td>
</tr>
<tr>
<td>Proportion 50/50</td>
<td>2800</td>
<td>500</td>
<td>0.05</td>
<td>0.083</td>
<td>0.095</td>
<td>0.109</td>
</tr>
<tr>
<td>Proportion 90/10</td>
<td>2800</td>
<td>500</td>
<td>0.05</td>
<td>0.021</td>
<td>0.025</td>
<td>0.033</td>
</tr>
<tr>
<td>5-Point scale (SD=1)</td>
<td>2800</td>
<td>500</td>
<td>0.05</td>
<td>0.070</td>
<td>0.083</td>
<td>0.110</td>
</tr>
<tr>
<td>Proportion 50/50</td>
<td>2800</td>
<td>400</td>
<td>0.05</td>
<td>0.040</td>
<td>0.048</td>
<td>0.063</td>
</tr>
<tr>
<td>Proportion 90/10</td>
<td>2800</td>
<td>400</td>
<td>0.05</td>
<td>0.024</td>
<td>0.029</td>
<td>0.038</td>
</tr>
<tr>
<td>5-Point scale (SD=1)</td>
<td>2800</td>
<td>400</td>
<td>0.05</td>
<td>0.080</td>
<td>0.095</td>
<td>0.125</td>
</tr>
<tr>
<td>Proportion 50/50</td>
<td>2800</td>
<td>300</td>
<td>0.05</td>
<td>0.047</td>
<td>0.056</td>
<td>0.074</td>
</tr>
<tr>
<td>Proportion 90/10</td>
<td>2800</td>
<td>300</td>
<td>0.05</td>
<td>0.028</td>
<td>0.034</td>
<td>0.044</td>
</tr>
<tr>
<td>5-Point scale (SD=1)</td>
<td>2800</td>
<td>300</td>
<td>0.05</td>
<td>0.095</td>
<td>0.112</td>
<td>0.148</td>
</tr>
<tr>
<td>Proportion 50/50</td>
<td>2800</td>
<td>200</td>
<td>0.05</td>
<td>0.059</td>
<td>0.070</td>
<td>0.092</td>
</tr>
<tr>
<td>Proportion 90/10</td>
<td>2800</td>
<td>200</td>
<td>0.05</td>
<td>0.035</td>
<td>0.042</td>
<td>0.055</td>
</tr>
<tr>
<td>5-Point scale (SD=1)</td>
<td>2800</td>
<td>200</td>
<td>0.05</td>
<td>0.118</td>
<td>0.140</td>
<td>0.184</td>
</tr>
<tr>
<td>Proportion 50/50</td>
<td>280</td>
<td>150</td>
<td>0.05</td>
<td>0.048</td>
<td>0.057</td>
<td>0.075</td>
</tr>
<tr>
<td>Proportion 90/10</td>
<td>280</td>
<td>150</td>
<td>0.05</td>
<td>0.029</td>
<td>0.034</td>
<td>0.045</td>
</tr>
<tr>
<td>5-Point scale (SD=1)</td>
<td>280</td>
<td>150</td>
<td>0.05</td>
<td>0.097</td>
<td>0.115</td>
<td>0.151</td>
</tr>
<tr>
<td>Proportion 50/50</td>
<td>280</td>
<td>100</td>
<td>0.05</td>
<td>0.109</td>
<td>0.115</td>
<td>0.132</td>
</tr>
<tr>
<td>Proportion 90/10</td>
<td>280</td>
<td>100</td>
<td>0.05</td>
<td>0.106</td>
<td>0.115</td>
<td>0.136</td>
</tr>
<tr>
<td>5-Point scale (SD=1)</td>
<td>280</td>
<td>100</td>
<td>0.05</td>
<td>0.124</td>
<td>0.140</td>
<td>0.179</td>
</tr>
<tr>
<td>Proportion 50/50</td>
<td>280</td>
<td>75</td>
<td>0.05</td>
<td>0.105</td>
<td>0.120</td>
<td>0.134</td>
</tr>
<tr>
<td>Proportion 90/10</td>
<td>280</td>
<td>75</td>
<td>0.05</td>
<td>0.061</td>
<td>0.072</td>
<td>0.080</td>
</tr>
<tr>
<td>5-Point scale (SD=1)</td>
<td>280</td>
<td>75</td>
<td>0.05</td>
<td>0.204</td>
<td>0.268</td>
<td>0.326</td>
</tr>
<tr>
<td>Proportion 50/50</td>
<td>280</td>
<td>50</td>
<td>0.05</td>
<td>0.132</td>
<td>0.152</td>
<td>0.181</td>
</tr>
<tr>
<td>Proportion 90/10</td>
<td>280</td>
<td>50</td>
<td>0.05</td>
<td>0.104</td>
<td>0.120</td>
<td>0.144</td>
</tr>
<tr>
<td>5-Point scale (SD=1)</td>
<td>280</td>
<td>50</td>
<td>0.05</td>
<td>0.347</td>
<td>0.417</td>
<td>0.517</td>
</tr>
</tbody>
</table>
Recruitment strategy. For this Outcome Study Design, since a follow-up survey with Welcome Baby clients would not include a comparison group or random assignment to treatment groups, participant sampling and recruitment could occur as women are exiting the Welcome Baby program at nine months postpartum—as opposed to when women enroll in the program. Given the roll-out of the program across sites and potential implementation issues (e.g., limited staff capacity, lack of staff training and experience with the protocols, lack of staff familiarity with community services), the evaluation could target women who enroll in Welcome Baby no earlier than six months after sites began providing services. This would ensure that women in the sample interacted with experienced home visiting staff and were not exposed to initial implementation challenges.

For illustrative purposes, the timetable displayed in Table 4 provides estimated time frames for each task involved in a follow-up survey. The initial steps of selecting and coordinating with partnering hospitals would likely extend through the first quarter of the project. Subsequently, data collection instruments would be drafted, piloted, revised, and programmed into a computer-assisted personal interviewing (CAPI) system. Staff would be trained and then begin recruitment of sampled participants 2-3 months prior to the targeted assessment point. Administration of a 12-month survey could begin approximately one year into the project, at which point women would be graduating from Welcome Baby who were not part of the initial cohort of enrollees (i.e., exposed to the program during the first six months of implementation). Data analysis and report writing could occur following each wave of survey administration. Alternatively, a descriptive report could be produced early in the process and a culminating outcomes report based on multivariate analyses could be released after completing all data collection activities.

Table 4. Timetable for Follow-up Client Survey

<table>
<thead>
<tr>
<th>Survey Task</th>
<th>Estimated Time Frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>Select hospitals; gain hospital consent to participate in the study; develop MOU and hospital IRB packages</td>
<td>Year 1 Quarter 1</td>
</tr>
<tr>
<td>Design draft instrument; pilot and revise instrument; program CAPI</td>
<td>Year 1 Quarters 2-3</td>
</tr>
<tr>
<td>Train field staff</td>
<td>Year 1 Quarter 3</td>
</tr>
<tr>
<td>Sample and recruit participants, about 2-3 months prior to children’s first birthdays</td>
<td>Year 1 Quarter 4</td>
</tr>
<tr>
<td>Begin 12-month survey administration</td>
<td>Year 2 Quarter 1</td>
</tr>
<tr>
<td>Analyze 12-month survey data</td>
<td>Upon completion of data collection</td>
</tr>
<tr>
<td>Submit final report on 12-month survey findings</td>
<td>Within 6 months of completing data collection</td>
</tr>
<tr>
<td>Begin 24-month survey administration¹</td>
<td>Year 3 Quarter 1</td>
</tr>
<tr>
<td>Analyze 24-month survey data</td>
<td>Upon completion of data collection</td>
</tr>
<tr>
<td>Submit final report on 24-month survey findings</td>
<td>Within 6 months of completing data collection</td>
</tr>
<tr>
<td>Begin 36-month survey administration</td>
<td>Year 4 Quarter 1</td>
</tr>
<tr>
<td>Analyze 36-month survey data</td>
<td>Upon completion of data collection</td>
</tr>
<tr>
<td>Submit final report on 36-month survey findings</td>
<td>Within 6 months of completing data collection</td>
</tr>
</tbody>
</table>

¹The estimated end date for survey administration would depend on sample size and number of completes per month.
Potential survey measures. The measures selected for a follow-up survey should align with the Welcome Baby program’s targeted program outcomes, which include:

- Positive parenting behaviors and practices;
- Healthy behaviors during pregnancy and interconception;
- Breastfeeding (initiation, duration, and exclusivity);
- Access to concrete services (e.g., WIC, CalFresh, MediCal); and
- Medical, dental, and psychosocial services (e.g., prenatal care, immunizations, well-baby visits).

Outcome measures could be built into a parent interview protocol and could also include observational assessments of the home environment and parenting behaviors, and direct assessments of children’s cognitive, language, and physical development, such as height and weight. This section includes some examples of potential survey items and scales that could be used to measure these identified outcomes.

Across many national studies (e.g., Early Head Start Baby FACES, the NICHD Study of Early Child Care and Youth Development, the National Longitudinal Survey of Youth–Child Supplement), the quality of home environment is measured with the Home Observation for Measurement of the Environment (HOME Inventory; Caldwell & Bradley, 2003). The HOME Inventory measures several aspects of parenting, such as acceptance, responsiveness, and involvement, as well as the presence of learning materials in the home and the variety of children’s care experiences. Other potential survey items, not drawn from validated instruments, could measure activities such as the frequency of reading to the child and engaging in other early learning activities, frequency of physical exercise, and frequency of television viewing.

Parenting behaviors and practices can be measured in multiple ways, through parent-report and observations. Another observational tool being used in Early Head Start is the Parenting Interactions with Children: Checklist of Observations Linked to Outcomes (PICCOLO; Roggman et al., 2009), which assesses the frequency of positive parenting behaviors, including affection, responsiveness, encouragement, and teaching, during a parent-child play session. Other observational rating schemes that require videotaping parent-child play interactions and later coding for specific parent and child behaviors have been developed for several early childhood studies, such as the Early Head Start Research and Evaluation Project and the NICHD Study of Early Child Care and Youth Development. The “bag task,” as it is referred to in these studies, involves providing parents and children with a bag or multiple bags of toys with the instructions that parents play with their children as they normally would, for a set amount of time (typically 10-12 minutes). These coding schemes require more extensive training as they include a wider range of ratings compared to the PICCOLO which is rated on a 3-point scale and can be used to code live. It is also more expensive to videotape but may lead to greater reliability.
Besides observational measures, some parent-report measures offer evidence of parenting practices and the relationship between parents and their children. One example is the Parent Attitudes towards Childrearing: Infant and Toddlers (PACR; Easterbrooks & Goldberg, 1990). Similarly, a subscale of the PSI measures parent-child dysfunctional interactions, which indicates the quality of the relationship between the parent and child. Additionally, the Knowledge of Infant Development Inventory (KIDI; MacPhee, 1981) is used to measure the skills and behaviors parents expect children to have at certain ages and is appropriate for parents of infants. To assess parenting practices beyond those measured by the identified scales, a parent interview could include questions about safe sleeping arrangements, use of car seats, discipline strategies, and frequency of physical punishment (particularly for toddlers and preschoolers).

Data on breastfeeding could be collected at the first survey wave (such as the 12-month survey) so mothers can reflect on their breastfeeding experiences and report if they ever initiated breastfeeding, how long they breastfed (if not still currently), and how long they breastfed exclusively. The survey could also ask mothers reasons for not breastfeeding or discontinuing breastfeeding to separate mothers who declined due to health reasons or inability to produce sufficient milk from mothers who preferred not to or did not like it.

Although positive maternal mental health is not a specific outcome of Welcome Baby, parent coaches do screen mothers for depression and link them to needed mental health services. Measuring maternal depression and parenting stress at multiple time points would be important to identify whether high-needs mothers who were identified and received services continued to have poor mental health.

Potential depression screeners include the Patient Health Questionnaire for Depression (PHQ-9), which is currently used as a Welcome Baby assessment and would therefore offer the benefits of a repeated measure to examine change in symptoms. Other common depression screeners include the Center for Epidemiologic Studies Depression Scale (CES-D; Radloff, 1977)—used in the Head Start Family and Child Experiences Survey (FACES) and the Early Childhood Longitudinal Study—Birth Cohort (ECLS-B), and the Composite International Diagnostic Interview Short Form (CIDI-SF; Kessler et al. 1998)—also used in the ECLS-B. Parenting stress has been best evaluated using the Parenting Stress Index (PSI; Abidin, 1995)—for example, it was used in the Early Head Start Baby FACES study and was adapted for the ECLS-B.

Perceived social support and family functioning could be assessed at any or all survey waves using several different measures, including the Family Support Scale (Dunst, Trivette, & Deal, 1988), the McMaster Family Assessment Device: General Functioning Scale (Epstein, Baldwin, Bishop, & Keitner, 1983), the Family Adaptability and Cohesion
Evaluation Scale (FACES; Olson, Portner, & Bell, 1982), and the Social Networks Interview (Cochran et al., 1990). The purpose of these measures is to elicit information about strength of participants’ support system and the cohesiveness of family units.

A number of validated child development measures are available to assess child outcomes at different ages. The Ages and Stages Questionnaire (ASQ; Squires, Bricker, & Twombly, 2002) is used by Welcome Baby staff to screen children at multiple time points for developmental delay in communication, problem solving, fine motor, gross motor, and personal-social skills. This tool is appropriate from 1 month to 5½ years and could be repeated during follow-up surveys for a continuous measurement of development. The ASQ-Social-Emotional (ASQ-SE) screener is an additional measure used to screen children for developmental risk in the area of social-emotional development. This tool is appropriate for ages 3 months to 66 months. Both tools rely on parent-report but also directly test children’s skills in particular areas where the parent is unsure of whether the child has achieved the skill.

Although the ASQ-SE screens for risk, the Brief Infant-Toddler Social & Emotional Assessment (BITSEA; Briggs-Gowan & Carter, 2002) is another commonly used evaluation tool for ages 1 to 3 years that identifies both social competence and problem behaviors, according to parent report. Other developmentally appropriate measures of early childhood outcomes typically assess receptive and expressive language skills. A few of the most commonly used tools for research and evaluation include the Peabody Picture Vocabulary Test, 4th Edition (PPVT-4; Dunn & Dunn, 2007; for ages 2 years and older), the Preschool Language Scale, 4th Edition (PLS-4; Zimmerman et al., 2002; for ages birth through 6 years 11 months), and the MacArthur Communicative Development Inventories (CDI; Fenson et al., 2000; for ages 8 through 37 months). The PPVT-4 and PLS-4 are direct assessments while the CDI is parent-report. All three have Spanish versions.

Child nutrition and healthy weight are key goals of Welcome Baby. The parent interview at each wave should assess the types of foods children eat and parental beliefs about a healthy diet to determine whether children are offered healthy and nutritious foods at every meal. Additionally, a measure of household food insecurity could be included, such as the Food and Nutrition Services’ Food Insecurity Scale (Bickel et al., 2000). This scale can be used to identify food insecurity among adults and among children. To assess healthy child weight and cases of overweight and obesity, direct measurements of children’s height and weight would need to be collected. Prior to age 2, body mass index could not be determined, but weight-for-age percentiles (i.e., how the child’s weight compares to other children of the same age and sex) could be.
B. Secondary Data Analysis Options for an Outcomes Study

Multiple data sources are available that can be utilized to examine the characteristics, experiences, and outcomes of Welcome Baby clients. Potential design options include an analysis of data from the Stronger Families LA Database to examine Welcome Baby clients’ initial outcomes; a comparison of characteristics and experiences of clients receiving Welcome Baby and Welcome Baby Lite also using data from the Stronger Families LA Database; analyses of linked administrative data from WIC, Medi-Cal, and the Department of Children and Family Services; and analyses of national survey data available for LA County to draw comparisons to Welcome Baby clients on related measures. These options are detailed in this section.

1. Analysis of Welcome Baby Clients’ Initial Outcomes Using Stronger Families LA Database

The Stronger Families LA Database includes all data collected on Welcome Baby clients, including demographics, risk profiles, engagement points completed, referrals received, and scores on assessments. These data also offer important information on immediate child and family outcomes that could be analyzed to understand the associations between program experiences and potential benefits. One option for an Outcomes Study would select a cohort of women who enroll during a specific period of time, such as over the course of one year. It would be important to wait until several months after all sites are offering services to ensure greater reliability of data entries since it may take staff some time to adjust to the database.

Data specific to clients’ program experiences could include:

- Outreach location;
- Timing of program entry (gestational weeks or postpartum in birth hospital);
- Dosage (date of each completed engagement point);
- Education given during visits;
- Number of referrals made, for which services, and whether services were received;
- Timing of case closures; and
- Reasons for case closures.

Potential fields that could be used to assess clients’ initial outcomes include:

- Receipt of prenatal care (to compare prenatal and postpartum enrollees);
- Health insurance enrollment;
- Breastfeeding initiation, duration, and exclusivity (up to 9 months);
- PHQ-9 scores (maternal depression);
- Life Skills Progression scale score;
- ASQ-3 scores (child development); and
- Home safety checklist score and safe sleeping arrangements.
Additionally, the database can track women who return to the Welcome Baby program after a subsequent birth. These data could be analyzed to estimate birth spacing.

2. **Analysis of Stronger Families LA Database for Welcome Baby Lite Clients**

Another design option is to analyze the Stronger Families LA data, similar to option 1, but for both women enrolled in Welcome Baby and women enrolled in Welcome Baby Lite (high-risk women living outside Best Start communities), to compare their characteristics, program experiences, and initial outcomes. Since Welcome Baby Lite offers only three postpartum engagement points, up until 2 months postpartum, (compared to five points ending at 9 months for full Welcome Baby) the data on initial outcomes would be limited to those collected up until that point in time across groups, such as breastfeeding initiation, prenatal care, health insurance enrollment, home safety, and safe sleeping arrangements. These analyses could offer an opportunity to compare characteristics and initial outcomes for low-risk and high-risk women.

3. **Linking to WIC, Medi-Cal, and Department of Children and Family Services (DCFS) Records**

The purpose of this task would be to develop a more thorough description of women who participate in the Welcome Baby program by linking to administrative records from WIC, Medi-Cal, and DCFS. These data would provide information on Welcome Baby clients’ ongoing use of these programs after exiting Welcome Baby as well as reports of child abuse and neglect from DCFS. The need for additional consent to access and link Welcome Baby data to administrative records is being determined at this time. A major challenge, however, is that linking WIC, Medi-Cal and DCFS data would require obtaining permission from the state to access and link the data.

4. **Secondary Data Sources for Local Comparison Estimates**

The National Survey of Children’s Health is sponsored by the Maternal and Child Health Bureau of the Health Resources and Services Administration. The survey examines the physical and emotional health of children ages 0-17 years of age. This study was last completed in 2012 and with national sample size of 95,677 child-level interviews, which suggests that for many measures estimates could be obtained for Los Angeles County. The survey estimates would provide a unique source of comparison to Welcome Baby clients on measures that relate to child well-being, including medical homes, family interactions, and parental health.

The National Survey of Children with Special Health Care Needs is also sponsored by the Maternal and Child Health Bureau of the Health Resources and Services Administration. The survey is used to assess the prevalence and impact of special health care needs among children 0 to 17 in the US. This study was last completed in 2010 with a
national sample size of 40,242 detailed child-level interviews, which suggests that for many measures estimates could be obtained for the greater Los Angeles region. The estimates from this survey would also provide a unique source of comparison to Welcome Baby clients on measures that relate to adequate health insurance, access to needed services, and adequate care coordination. Other topics from this survey include functional difficulties, transition services, shared decision-making, and satisfaction with care.

V. Conclusion

This design report identifies multiple design options for an evaluation of Welcome Baby implementation across Best Start communities and the outcomes associated with Welcome Baby participation. The options include both primary data collection and secondary data analysis, and employ both qualitative and quantitative research methods. Within each design option, additional options are provided for the methods used with some discussion of the feasibility and pros and cons of each method. For example, key informant interviews, structured observations, and focus groups are each optional, and the frequency and intensity of the approach may yield a more complete picture, but there is a tradeoff in the amount of necessary resources. Similarly, following a cohort of Welcome Baby clients until children are 36 months old versus 24 months and examining national survey data to draw sample comparisons are both options that could offer fuller information on client outcomes relative to national estimates, but each component would carry additional costs, necessitating a review of the associated costs and benefits of each option before study execution.