DESIGN OPTIONS FOR THE MATERNAL, INFANT, AND EARLY CHILDHOOD HOME VISITING EVALUATION

Draft Final Report

APRIL 2011

DESIGN OPTIONS FOR THE MATERNAL, INFANT, AND EARLY CHILDHOOD HOME VISITING EVALUATION

Disclaimer

Acknowledgments

Chapter 1

Background on Home Visiting and Goals for the Evaluation

Home visiting programs in the U.S. grew from three major approaches that first became prominent in the 1960s: visits by public health nurses to promote infant and child health in disadvantaged families, Head Start home visiting to promote school readiness in hard-to-reach families, and home-based family support to promote positive parenting and prevent child abuse in high risk families. All of these approaches sought to foster early childhood health and development by intervening in the home to support and improve socialization, health, and education practices.[1] Today, home visiting is seen as a particularly important strategy for high risk families who may be difficult to engage in other services. A study by the PEW Center on the States found that in fiscal year 2009-2010, states spent more than $500 million to fund home visiting programs, with additional programs funded by local governments and private foundations.[2]

On March 23, 2010, the President signed into law the Patient Protection and Affordable Care Act (ACA) of 2010. Through a provision that authorizes the creation of the Maternal, Infant, and Early Childhood Home Visiting Program (MIECHV), the Act greatly expands federal funding of evidence-based home visiting programs. According to a Supplemental Information Request (SIR) released by the Health Resources and Services Administration (HRSA) in February 2011, "this program is designed: (1) to strengthen and improve the programs and activities carried out under Title V; (2) to improve coordination of services for at-risk communities; and (3) to identify and provide comprehensive services to improve outcomes for families who reside in at-risk communities." The legislation defines at-risk communities as communities with concentrations of:

- premature birth, low-birth weight infants, and infant mortality, including infant death due to neglect, or other indicators of at-risk prenatal, maternal, newborn, or child health;

- poverty;

- crime;

- domestic violence;

- high rates of high school dropouts;

- substance abuse;

- unemployment; or

- child maltreatment.

The ACA provides $100 million in Federal funds for the Program in fiscal year 2010 with substantial annual growth, and will total $1.5 billion over five years. Seventy-five percent

---

[1]Weiss (1993).

[2]The PEW Center on the States (2010).

of program funding must be used for home visiting models with evidence of effectiveness. A recent review funded by the U.S. Department of Health and Human Services (HHS) and conducted by Mathematica Policy Research – the Home Visiting Evidence of Effectiveness (HomVEE) review – found seven models that met the HHS criteria for evidence of effectiveness. These seven models will be referred to as evidence-based models in this report. Each state can use up to 25 percent of its funds for promising approaches that do not yet qualify as evidence-based. The ACA reserves 3 percent of the total $1.5 billion in MIECHV funding for research and evaluation, as well as for providing technical assistance to grantees after they submit their three-year benchmarks.

Needs Assessments and State Plans

To receive MIECHV funding for fiscal year 2010, states must complete a three-step application process. In the first step, states described their plans for carrying out a needs assessment, their intentions for developing their program, and their anticipated need for technical assistance while carrying out the needs assessment. In the second step, states submitted preliminary needs assessments in response to the SIR for the Submission of the Statewide Needs Assessment issued in August 2010 by HRSA. This document requested that states provide three main categories of information:

- Identify at-risk communities. States were asked to provide several specific indicators meant to help identify at-risk communities (such as number of live births before 37 weeks as a percentage of the total number of live births).

- Identify the quality and capacity of existing programs. The August 2010 SIR also required states to report several indicators of the quality and capacity of existing home visiting programs (such as the home visiting models or approaches in use and the specific services provided). As part of the description of their capacity, states were to discuss the extent to which such programs and initiatives are meeting the needs of eligible families and to identify gaps or duplications in early childhood home visiting services.

- Discuss the capacity for providing substance abuse treatment and counseling services. Finally, the August 2010 SIR required states to provide descriptions of state capacity for providing substance abuse treatment and counseling services.

The third step in the application process for MIECHV funding requires states to submit a more detailed needs assessment for their identified at-risk communities and a state home visiting plan. The SIR released in February 2011 instructed states on completing their plans. This document requested that states provide several types of information, including the at-risk community or communities the state intends to target, the states' primary goals and objectives for the home visiting program, the home visiting models the states are planning to fund with MIECHV funding, the states' plans for implementing and administering their home visiting programs, and state plans for meeting legislatively-mandated benchmarks and for using continuous quality improvement strategies.

The National Evaluation of Evidence-Based Home Visiting Programs

The legislation specifies that there should be an ongoing program of research to increase knowledge about home visiting implementation and effectiveness. Specifically, the legislation requires a national evaluation of MIECHV to report findings to Congress in 2015. The Design Options for Home Visiting Evaluation (DOHVE) project is designing this national evaluation, and is presenting a preliminary design in this report.

The ACA specifies four main components of the national evaluation:

- <u>Analysis of the needs assessments</u>. An analysis, on a state-by-state basis, of the results of assessments of state needs that are required by the legislation and state actions in response to the assessments.

- <u>Effectiveness study</u>. An assessment of the effects of early childhood home visiting programs on child and parent outcomes, with respect to each of the benchmark areas and participant outcomes specified in the legislation.

- <u>Subgroup analysis</u>. An assessment of the effectiveness of such programs on different populations, including the extent to which effects on participant outcomes vary across programs and populations.

- <u>Study of effects on the health care system</u>. An assessment of the potential for the activities conducted under such programs, if scaled broadly, to improve health care practices, eliminate health disparities, improve health care system quality, and reduce costs.

This report addresses each of these components, and also extends the evaluation design to answer additional questions of interest to HHS. In addition to specifying components of the national evaluation listed above, the legislation delineates several outcome domains that must be measured as part of the evaluation:

- Prenatal, maternal, and newborn health;

- Child health and development;

- Parenting skills;

- School readiness and academic achievement;

- Crime and domestic violence;

- Family economic self-sufficiency; and

- Referrals and service coordination.

Chapter 4 discusses the DOHVE team's recommendations on specific constructs to measure within these domains.

This report presents a design for an evaluation that would study the evidence-based home visiting programs, but not the promising approaches that states can fund with up to 25 percent of

their MIECHV funds. The primary reason for excluding promising programs is that each state that chooses to fund a promising program may pick a different approach. This could lead to a great deal of diversity in the promising programs, and a small number of families served using each program model. Tremendously diverse program models with small numbers of participants could make it particularly difficult and expensive to do a rigorous evaluation of all of these models in the national evaluation. In addition, promising programs may vary in their evaluation needs: some may benefit from a study of their effectiveness of family outcomes but others might benefit first from formative research to understand whether the new approach is addressing family needs as intended. For this reason, the legislation requires states to evaluate promising approaches but also allows the states to choose the type of evaluation that is most appropriate for the approach being used.

Key Features of Evidence-Based Home Visiting Programs

The national evaluation will focus on the seven home visiting models that HomVEE determined meet the criteria for evidence-based models: Early Head Start – Home-Based Option, Family Check-Up, Healthy Families America (HFA), Healthy Steps, Home Instruction Program for Preschool Youngsters (HIPPY), Nurse-Family Partnership (NFP), and Parents as Teachers (PAT). While these models all include home visiting services, they differ in many respects. Tables 1.1 and 1.2 summarize some important features of the evidence-based models and their implementation systems, which include the following:

- Program goals. While all of the models try to improve child health and development broadly conceived, specific goals differ among the models: some focus more directly on preventing child maltreatment, others focus on improving maternal and child health, and still others prioritize the promotion of positive parenting or increasing school readiness. Some models also explicitly aim to improve parental self-sufficiency and well-being.

- Target population and age at enrollment. Most of these models serve families whom they identify as being at-risk of poor child outcomes based on one or more family characteristics. Although the definition of at-risk differs by model, most models target low-income families. They may also specifically target young, first-time mothers, parents with past negative school experiences, or families with maternal depression or substance abuse problems. Healthy Steps and PAT have historically served a broad array of families with children in their target age ranges, but programs funded by MIECHV will be required to focus on families in communities defined as at-risk by the legislation.[3] Many of the models begin to work with women when they are pregnant or when they have newborns. Early Head Start and PAT accept pregnant mothers as well as families whose youngest child is up to 3 or 5 years of age, respectively. HIPPY and Family Check-Up target families with children older than 2 years of age.

- Program intensity and duration. Many models require some combination of weekly and biweekly meetings for an extended period of time. Early Head Start, HFA, NFP, PAT,

---

[3]While the MIECHV Program will only fund programs providing services in at-risk communities, funded programs would be allowed to serve families that do not fit in any of the specific at-risk categories as long as those families are in the chosen at-risk communities.

and HIPPY all offer services to clients for several years, provided that their children are not enrolled in the programs near the end of the eligible age range. Healthy Steps and Family Check-Up are a bit different. For Healthy Steps, home visitors conduct six home visits with the family over the course of three years. In Family Check-Up, only three home visits are made to clients, and all are to be completed soon after clients enroll in the program. There is also variation in how long models expect each home visit to last, with the expectations generally ranging from 50 to 90 minutes per home visit.

- Home visitor qualifications. The evidence-based models have a wide range of standards for home visitor qualifications. For example, NFP requires that home visitors are registered nurses, and Family Check-Up recommends hiring home visitors with doctoral or master's degrees in psychology or a related field. Some models focus on finding home visitors that they think will connect well with families based on personal characteristics. Other models allow local programs to set the criteria. Many of the models require home visitors to have relevant experience or knowledge.

- Requirements for data systems. HFA, NFP, and HIPPY have specific data systems that they require implementing agencies to use. Other models do not currently have explicit requirements for the data systems used by agencies implementing their models.

- Home visitor and supervisor training requirements. Most of the evidence-based models have training requirements for home visitors and supervisors, although the requirements differ in terms of timing and intensity. Many of the models require three to five days of pre-service training. While HIPPY requires coordinators to complete a one-week HIPPY pre-service training, it only requires home visitors to receive training during an initial site visit conducted by a national HIPPY trainer. Many of the programs also have ongoing training requirements. For example, HIPPY home visitors receive weekly training from their coordinator, and NFP requires that nurse home visitors and supervisors complete three core education sessions that take place over a nine-month period. Nursing supervisors must also complete additional education sessions.

The Design Options for Home Visiting Evaluation Project

Table 1.1[4]

Key Components of Service Models for the Evidence-based Home Visiting Programs

Home Visiting Model	Program Goals	Target Population/ Age at Enrollment	Program Intensity/Duration	Home Visitor Qualifications
Early Head Start – Home Visiting	Promote healthy prenatal outcomes for pregnant women Enhance the development of very young children Promote healthy family functioning	Low-income pregnant women and families with children birth to age 3 years Families at or below the federal poverty level Children with disabilities who are eligible for Part C services under the Individuals with Disabilities Education Act in their state	Home visits last for a minimum of 90 minutes Minimum of weekly home visits and bi-monthly group socialization sessions Offer a minimum of 32 home visits and 16 group socialization activities per year Services can begin prenatally and are offered until the child is 3 years old	Require that home visitors have knowledge and experience in child development and early childhood education; child health, safety, and nutrition; adult learning; and family dynamics Effectively communicate with children and families with no or limited English proficiency directly or through an interpreter; be familiar with the ethnic background of these families Demonstrated skills and abilities in human services program management
Family Check-Up	Reduce children's conduct, academic, and internalizing problems Reduce maternal depression Improve parental involvement and positive parenting	Families at socioeconomic risk Families with children who have conduct problems, academic failure, or depression and are at risk for early substance use Families with children age 2 to 17 years old	First session, an initial interview, lasts 30 to 60 minutes Second session, an in-home assessment, lasts 60 minutes Third session, a family feedback session, lasts 60 minutes After three sessions, parent consultant recommends a variety of family-based interventions varying in intensity and duration	Recommend that parent consultants have a doctoral or master's degree in psychology or a related field and previous experience carrying out family-based interventions If given additional support, allow programs to use parent consultants who have a bachelor's or associate's degree

(Continued)

---

[4]Information for this exhibit taken from the US Department of Health and Human Services HomVEE website: http://homvee.acf.hhs.gov/programs.aspx and program model websites.

Table 1.1 (Continued)

Home Visiting Model	Program Goals	Target Population/ Age at Enrollment	Program Intensity/Duration	Home Visitor Qualifications
Healthy Families America (HFA)	Increase use of prenatal care, access to primary care medical services, and immunization rates Ensure healthy child development and improve parent-child interactions and school readiness Promote positive parenting and reduce child maltreatment Promote family self-sufficiency and decrease dependency on welfare and other social services	Parents facing challenges such as single parenthood, low income, childhood history of substance abuse, mental health issues, domestic violence Individual programs select the specific characteristics of the target population they plan to serve Require that families are enrolled prenatally or within the first three months after a child's birth	Home visits typically last a minimum of 60 minutes Offer at least one home visit per week for the first six months after the child's birth After the first six months, local programs determine the frequency of the visits Services are provided until the child enters kindergarten	Do not provide specific educational requirements for direct-service staff Recommend selecting staff based on personal characteristics; willingness to work in, or experience working with, culturally diverse communities; experience working with families with multiple needs; and ability to maintain boundaries between personal and professional life
Healthy Steps	Enhance the information and support services available to parents Emphasize a close relationship between health care professionals and parents in addressing physical, emotional, and intellectual growth and development of young children	Implemented by medical practices Open to all families served by the participating practice or organization Parents with children from birth to age 3 Target first visit to occur when the child is 3-6 days old	Home visits should last between 60-90 minutes Can be implemented at three different levels of intensity which offer between two and five home visits that occur from the child's birth until they are 36 months old Serve families from birth until the child is 36 months of age	Recommend bachelor's degree with advanced training or education in child development, family studies, nursing, psychology, or a related field Prefer previous experience and knowledge about early child growth and development and parent-child relationships; demonstrated flexibility to work on an interdisciplinary team; experience working in a medical setting or with health professionals; and relevant life experiences

10

Table 1.1 (Continued)

Home Visiting Model	Program Goals	Target Population/ Age at Enrollment	Program Intensity/Duration	Home Visitor Qualifications
Home Instruction for Parents of Preschool Youngsters (HIPPY)	Increase vulnerable children's success in school and, ultimately, in life Improve language development, problem solving, logical thinking, and perceptual skills Enhance social, emotional, and physical (fine and gross motor skills) development	Families with children age 3 to 5 years Parents who lack confidence in their ability to prepare their children for school, including parents with past negative school experiences, limited financial resources, or other risk factors Encouraged to offer a three-year program serving 3 to 5 year olds, but sites may offer the two-year program for 4 to 5 year olds	Recommend that group meetings last 120 minutes, but no information is available about the length of the home visits Offer weekly activities for 30 weeks of the year, alternating between home visits and group meetings Require that all sites serve children until they are 5 years old	Prefer that home visitors be recruited from the targeted community and have a child of HIPPY-appropriate age, or have access to a practice child with whom they can engage in the HIPPY curriculum Other qualifications may be specified by the local implementing agency
Nurse Family Partnership (NFP)	Improve prenatal health and outcomes Improve child health and development Improve families' economic self-sufficiency and maternal life course development	First-time, low-income mothers and their children Children up to 2 years of age Require first home visit for occurrence no later than the end of week 28 of pregnancy, recommend that programs begin conducting visits in the second trimester (14 to 16 weeks gestation)	Home visits typically last 75 minutes Require that programs offer home visits on a weekly, biweekly, or monthly basis depending on the stage of pregnancy and the age of the child; more intensive services are to be provided earlier in the pregnancy and following the birth of the child Services are provided until the child is 2 years old	Require that nurse home visitors are registered professional nurses with a minimum of a bachelor's degree in nursing

(Continued)

Table 1.1 (Continued)

Home Visiting Model	Program Goals	Target Population/ Age at Enrollment	Program Intensity/Duration	Home Visitor Qualifications
Parents as Teachers (PAT)	Increase parent knowledge of early childhood development and improve parenting practices Provide early detection of developmental delays and health issues Prevent child abuse and neglect Increase children's school readiness and school success	Do not have eligibility requirements for participants, individual programs select the specific characteristics of the target population they plan to serve Many programs are equipped to serve children and their families from pregnancy through kindergarten entry Target enrollment prenatally or soon after birth	Recommend that home visits last between 50 and 60 minutes Local programs can determine the intensity of the services they provide, but recommend that families with high risk factors receive more frequent visits Offer home visits on a monthly, biweekly, or weekly basis Local programs determine the child's age at which families stop receiving services	Do not provide specific requirements for staff background and education

The Design Options for Home Visiting Evaluation Project

Table 1.2[5]

Key Components of Implementation Systems for the Evidence-based Home Visiting Programs

Home Visiting Model	Model Requirements for Data Systems	Home Visitor and Supervisor Training Requirements
Early Head Start – Home Visiting	No specific infrastructure or data system requirements Recommend programs use record-keeping systems	Require that programs provide orientation to all new staff focused on program goals and implementation. Pre-service training opportunities help staff and volunteers acquire or increase knowledge and skills needed to fulfill job responsibilities. Training aimed at improving ability of staff and volunteers to deliver services required by program regulations and policies. Require programs to provide ongoing opportunities for training and professional development. Programs implement structured staff training and development system, offering academic credit where possible.
Family Check-Up	Prefer implementing agencies have high-speed Internet to upload digital images of sessions for supervision	Provide a three-month certification process that employs didactic instruction and role playing, followed by ongoing videotaped supervision of intervention activity. Certification is established by reviewing videotapes of feedback and follow-up intervention sessions to evaluate whether the parent consultants are competent in all components of the intervention. Certification process for parent consultants is repeated yearly to reduce drift from the intervention model.
Healthy Families America (HFA)	Require implementing agencies to use the Program Information Management System (PIMS)	Provide mandatory five-day workshop delivered by certified trainers. Workshop prepares program staff to provide services and describes best practices of family-centered and strength-based service provision. Staff members receive this training before working with clients or within the first six months of employment. Sites can also offer prenatal training, lasting between three and four days, that teaches specific strategies for supporting families prenatally. Encourage sites to provide in-service training about specific challenges faced by community's families and local resources available. Recommend that staff devote one-third of their time (about 80 hours) to in-service training in the first six months of employment.

(Continued)

---

[5]Information for this exhibit taken from the US Department of Health and Human Services HomVEE website: http://homvee.acf.hhs.gov/programs.aspx

Table 1.2 (Continued)

Home Visiting Model	Model Requirements for Data Systems	Home Visitor and Supervisor Training Requirements
Healthy Steps	No information available	Require Healthy Steps Specialists to complete a three-day training institute provided by the training team in Boston or on-site. Recommend that the lead physician and others who will work directly with Healthy Steps families also attend training. Offer a one-day workshop designed for pediatric clinicians. Alternatively, offer a multimedia training kit that practices can use.
Home Instruction for Parents of Preschool Youngsters (HIPPY)	Require that sites implement the HIPPY management information system (MIS) a computer program	Require home visitors to receive training in the use of the curriculum during an initial site visit conducted by a HIPPY trainer. Require coordinators to complete a one-week pre-service training successfully before starting a new program or coordinating an existing one. Pre-service training is five days in length and covers all aspects of administering a program. Home visitors receive weekly trainings from their coordinator, which periodically include professional skill development. National trainers also perform annual site visits during which training may be provided to home visitors.
Nurse Family Partnership (NFP)	Require implementing agencies to use web-based data system, Clinical Information System (CIS)	Require nurse home visitors to complete three core education sessions, in both distance and face-to-face training formats. Sessions take place over a nine-month time frame. Supervisors must complete the same core education sessions as home visitors. They must also complete four supervisor core education sessions, including two face-to-face sessions.
Parents as Teachers (PAT)	No information available	Require that all parent educators attend a five-day training. Offer separate trainings for programs that work with families prenatally until three years and for those that work with preschool-aged children. Require supervisors to complete a separate training. Require additional training for parent educators who will administer developmental, hearing, or vision screenings. Parent educators receive a sixth day of training, held three to six months after their pre-service training. Recommend that supervisors participate in a day-long, advanced training after they been a supervisor for six months. Offer two- or four-day professional development sessions for those who work with special populations. Parent educators complete a minimum of 20 hours professional development during their first year, 15 hours during their second, and 10 hours per year thereafter. Supervisors complete 10 hours of professional development annually.

Research on Home Visiting Programs

According to the HomVEE review, all of the evidence-based models have "at least one high- or moderate-quality study with at least two favorable, statistically significant impacts in two different domains[6] or two or more high- or moderate-quality studies using non-overlapping analytic study samples with one or more statistically significant, favorable impacts in the same domain."[7] HomVEE also found some evidence that all of the models except Healthy Steps affected primary measures of positive parenting practices, child development, and school readiness.[8] Three models had a study showing that they improved a primary measure of child health. Positive effects on primary measures in other outcome domains were found for fewer than three of the models.

While the evidence-based models have each been found to produce some positive effects, there are many remaining gaps in knowledge about home visiting programs.

Inconsistent effects between different samples for a given model. Even for the evidence-based models, effects have often been modest and inconsistent across different samples. Many times, findings of effects for certain outcomes and subgroups have not been replicated in later studies with different samples. For example, the HomVEE review found that most of the evidence-based models had studies that showed favorable effects on at least one primary outcome measure of child development and school readiness. At the same time, there were a number of other studies that failed to find positive effects on any measures of child development and school readiness. This may have occurred in part because some studies had samples that were too small to detect modest effects. The national evaluation presents an opportunity to do a rigorous evaluation with enough families to detect modest effects so that the field has clearer evidence on the effects of evidence-based home visiting programs on the outcome domains of interest.

Different outcomes tested in different studies. One difficulty in interpreting the home visiting research is that different studies measured different outcomes. In part this is because different program models target different domains, and studies of those models may have only focused on the targeted outcomes. In addition, different evaluators have looked at different measures within a given outcome domain. The national evaluation can add to knowledge about home visiting programs by collecting a consistent set of information across all relevant outcome domains for all of the evidence-based models.

---

[6]HomVEE recategorized the outcome domains specific in the ACA. The review looked at the following eight domains: child health; child development and school readiness; family economic self-sufficiency; linkages and referrals; maternal health; positive parenting practices; reductions in child maltreatment; and reductions in juvenile delinquency, family violence, and crime.

[7]Paulsell, Avellar, Martin, and Grosso (2010).

[8]The HomVEE review defined an outcome measure as primary if "data were collected through direct observation, direct assessment, or administrative records; or if self-reported data were collected using a standardized (normed) instrument."

Insufficient evidence of effectiveness in subgroups. The HomVEE review found that sample sizes were generally too small to conclude whether home visiting models worked for particular types of families and identified this as a gap in the home visiting research. Many studies of home visiting programs involve sample sizes that are too small to allow a precise analysis of subgroup effects, and those studies that have examined how effects differed by subgroup have often focused on different subgroups. This has led to thin evidence on some subgroups. The field would benefit from research that helps identify what works for different types of families.

Lack of information on program implementation. Prior studies of human service programs have found that program effects are associated with a number of factors such as program maturity, clarity of program goals, and the extent to which services target specific outcomes.[9] However, evaluations of home visiting programs have rarely collected detailed information on the services actually delivered. This makes it difficult to know when weak impacts are due to problems of implementation rather than features of the program model. The field could greatly benefit from a systematic examination of how program features are associated with service delivery and impacts.

These gaps in prior research suggest the importance of a national evaluation. To understand how MIECHV affects outcomes for families and children and to inform the field about what works best for whom, and on which outcomes, the national evaluation will need to gather systematic information that is consistent across all of the evidence-based models that states choose to operate. This national evaluation has enormous potential to contribute to the field by collecting common measures across several program models about not only the outcomes of interest, but also the actual services provided to families and the community, organizational, and family characteristics that are associated with service delivery and impacts. Strengthening the field means helping states, communities, program developers, and program operators build programs that produce strong, consistent impacts across the full range of intended outcomes and targeted population subgroups. This report presents an evaluation designed to accomplish these goals.

Overview of the Report

The remainder of the report presents the proposed design for the national evaluation. Each chapter of the report presents an aspect of the evaluation design. Chapter 2 begins by giving a brief overview of research goals and questions, the conceptual framework used to guide the evaluation design, and some of the key challenges that the evaluator will face in conducting this evaluation. This chapter also discusses the basic design of the evaluation and presents the anticipated timeline for the study. Chapter 3 discusses the sampling plan and presents details on the number of families and sites that would be included as well as discussing how sites will be chosen. This chapter also presents a design for conducting the analysis of the needs assessment data. Chapter 4 discusses an implementation study that could be included in the national evaluation, while Chapter 5 presents the measurement and analytic plan for the impact study, including how the evaluation would assess the ability of home visiting programs to affect health

---

[9]Fixsen et al. (2005).

disparities and health care quality. Chapter 6 describes an economic evaluation, which would estimate the cost of achieving key outcomes through home visiting. Appendices A and B provide brief summary of the resources required to conduct the core study and additional research activities that could be conducted as part of either the national evaluation or an ongoing research agenda..

<div align="center">

Chapter 2

Overview of the Proposed Design

</div>

This chapter presents a brief overview of the design of the national evaluation, which will be elaborated on in the remainder of the report. Before presenting the overview, the chapter describes four foundations on which the design is based: the goals that the design is intended to achieve, the conceptual framework of how home visiting programs achieve their effects, the research questions that stem from those goals and framework, and the unique challenges related to a study of home visiting programs.

Goals of the National Evaluation

The national evaluation is designed to meet legislative requirements as well as a number of additional goals set forth by the US Department of Health and Human Services (HHS) that reflect the background on home visiting presented in Chapter 1. Meeting legislative goals will require:

- Using a rigorous design for assessing the effectiveness of home visiting services overall, and variations across programs and populations. The evaluation should seek to obtain credible evidence of the effects of home visiting services, and it should be able to address questions about key subgroups of programs and families. This will require gathering information about the characteristics of families and of programs

- Learning about the effectiveness of home visiting programs across all domains specified in the Affordable Care Act (ACA). As noted in Chapter 1, prior studies of home visiting have varied in the domains they analyzed and the outcomes examined within each domain. The national evaluation will improve what is known about home visiting by measuring outcomes consistently across all sites included in the evaluation.

- Reflecting the national diversity of communities and populations. Home visiting currently takes place in thousands of communities involving many thousands of families. Under Maternal Infant and Early Childhood Home Visiting (MIECHV) program, home visiting may be extended to even more places and serve even more families with particular needs.

In addition, HHS aims to gain information to strengthen future programs by:

- Systematically studying program implementation. Also as noted in Chapter 1, prior studies of home visiting programs have often included little information on the actual services provided to families and on the community, organizational, and family characteristics that influence service delivery. The national evaluation can provide valuable information on these issues.

- Linking information on communities, organizations, and families to program impacts, the national evaluation can deepen our understanding of program features associated with

greater benefits. This understanding can be used to strengthen impacts and to broaden the range of outcomes and populations in which home visiting improves child and family well-being and eliminates health disparities.

A Conceptual Framework of Home Visiting Programs

The evaluation design is based on a conceptual framework for how home visiting programs operate and achieve their effects (Figure 2.1). The framework has three broad aspects – inputs, outputs and outcomes. In discussing this framework and in the remainder of this report, the term *site* will be used to refer to the local implementation of home visiting, and a site will operate a local home visiting program. Each site in the national evaluation will use one of the national evidence-based home visiting *models*. Local sites will be administered by a state agency and may be implemented by a local agency. The *implementation system* refers to the resources used to implement the home visiting model.

Inputs influence how services are provided to families and are shown in the left half of the figure. Starting at the bottom of the figure, community resources provide a foundation from which programs operate. This foundation sets the stage for home visiting by determining the resources available to home visiting program sites and the opportunities available to families.

Multiple organizations influence how a specific home visiting program defines its service model and its implementation system. These organizations may include the state agency that receives MIECHV funds, the local agency that operates the home visiting program, the developer and purveyor of the evidence-based model that has been adopted, and other community organizations with which the implementing agency collaborates. Thus, the service model and implementation system for any two program sites will not be identical, even if they use the same evidence-based model. In many instances, these differences among sites will lead to differences in how they deliver services and the outcomes they achieve.

The service model and implementation system in turn affect the characteristics of home visitors in a given program. As noted in Chapter 1, some national models specify the professional background that home visitors must have, while other national models give local sites considerable discretion in this regard. The implementation system also shapes home visitor attributes including their competence in carrying out their responsibilities. Some programs have more intensive training and supervision than others. Furthermore, within a given program site, adequacy of staff training and supervision might be emphasized more for responsibilities related to some intended program outcomes than for other intended outcomes.

The service model and implementation system also affect the characteristics of families who enroll in a given home visiting program. For example, most of the evidence-based national models specify characteristics of the families that can be served: some models limit enrollment to pregnant women, some limit enrollment to families with children above a certain age, and some serve a broader range of families. State and local agencies may further restrict or expand the eligibility requirements for home visiting in a particular program. This might be accomplished by limiting enrollment to families with a particular need. Alternatively, a program might be required to include families with a particular need outside the model developer's definition of eligible families. Sites will also vary in their procedures for family recruitment. This can include how

staff explain the purpose and intended benefits of the program, which might influence families' understanding of the program and willingness to enroll.

Design Options for Home Visiting Evaluation Project

Figure 2.1

Conceptual Framework for a Home Visiting Program
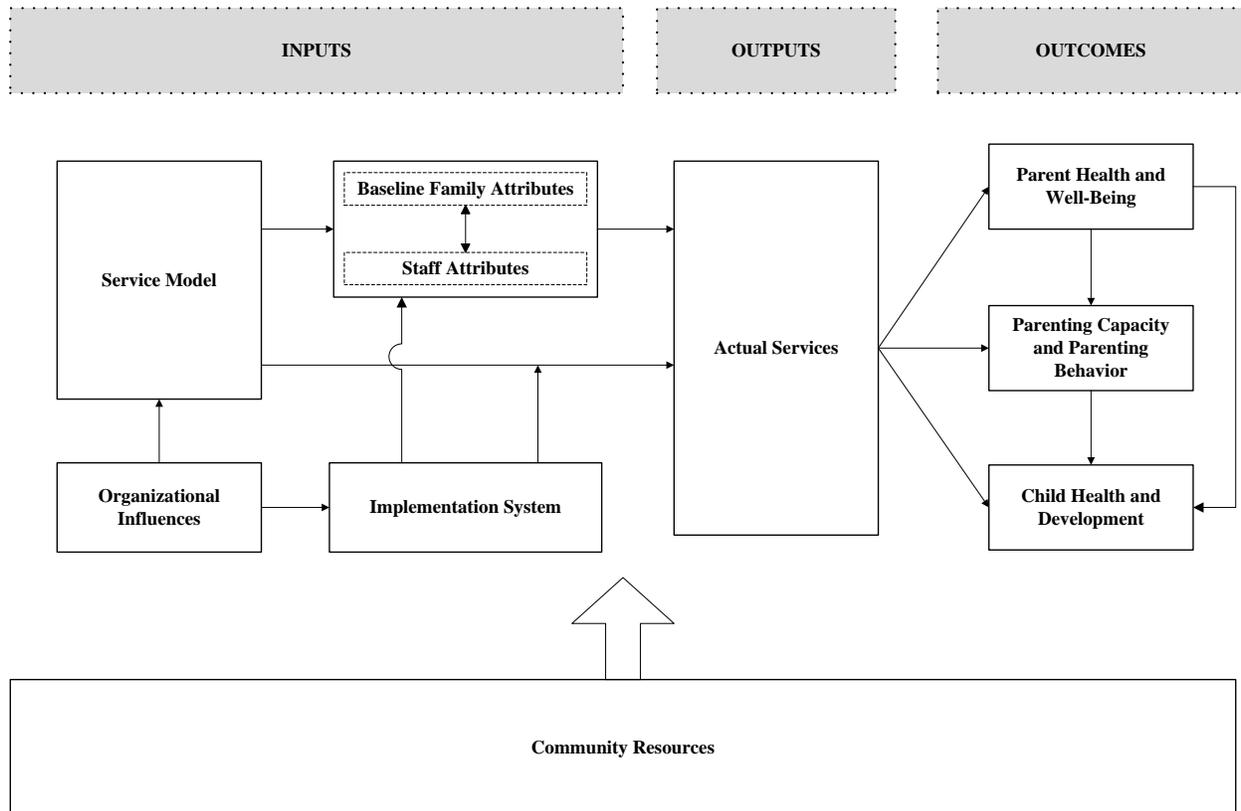
The service model, implementation system, and characteristics of home visitors and families all affect the services that families receive directly from the home visiting program and indirectly as a result of referrals to other services. These services are outputs, as shown in the middle of the figure. Services include program coverage of the target population and the quantity and quality of service delivery.

The service model influences actual service delivery by defining intended outcomes; expected frequency, duration, and content of home visits; and intended linkages with other services. These definitions are sometimes clear and coherent, but may also be underspecified or

contain ambiguous or incongruent elements. As the service model's clarity and congruence increase, so does its fidelity.[10]

The implementation system includes the resources for carrying out the service model. It incorporates policies and procedures for staff recruitment, training, supervision and evaluation; assessment tools, protocols and curricula to guide service delivery; the use of administrative supports such as management information systems to monitor and promote staff adherence to the service model; organizational culture and climate regarding fidelity and the use of evidence-based practices; the availability of consultation to address issues beyond the home visitor's skills and expertise; and the home visiting program's relationships with other community-based organizations to facilitate referral and service coordination. As the adequacy of the implementation system increases, so does its fidelity. [11]

Baseline family attributes also influence actual services in several ways. First, evidence-based home visiting models encourage the tailoring of services based on family strengths, needs, and concerns. Second, families can vary in their understanding of a program, the benefits they are likely to derive from it, and what enrollment entails. Third, parents vary in their cognitive and emotional capacity to engage with services offered by the home visitor.

A range of home visitor characteristics can also influence actual services delivered. The home visitor's understanding of the program model and their roles and responsibilities will inform the services they choose to deliver. Their beliefs about the relative importance of specific tasks and parenting risks such as intimate partner violence will also influence how they provide services. Their ability to ascertain family strengths, needs and concerns will shape their relationship with the family and decision making about which services to provide. Their own psychological well-being, including whether they are depressed or experiencing burnout, will also influence how they approach their work with families.

Lastly, family and staff characteristics interact to influence service delivery. In short, the same home visitor might provide services differently for one family than another, not only because they tailor those services, but also because they deliver services in ways that are not intended by the model developers. Consider, for example, a home visitor's screening for and discussion of psychosocial risks for poor parenting. Home visitors might vary in their self-efficacy in carrying out this role. One might feel comfortable discussing these risks across all families. Another might discuss risks with families she perceives are comfortable with self-disclosure, but shy away from discussion with families that are reluctant to disclose.

The right side of the figure shows outcomes that home visiting is designed to achieve for families. Programs aim to improve family health and functioning (including ACA-noted domains of prenatal and maternal health, crime, domestic violence, and economic self-sufficiency), parenting knowledge and practices, and child health and development (including the domains of child health, school readiness, and academic achievement).

---

[10]Carroll et al. (2007).
[11]Carroll et al. (2007).

Research Questions

The goals and conceptual framework together suggest the following research questions to be addressed by the national evaluation. They are presented in two groups, those that are required in statute and so must be included in the national evaluation, and those that, while not required, would be highly beneficial for HHS to understand. Required research questions include the following:

- What are the average effects of home visiting programs on the range of outcomes specified in the legislation (including maternal and child health, school readiness, child maltreatment, parenting practices, economic status, and crime and domestic violence)?

- What are the effects for key subgroups of families, defined according to characteristics of policy interest?

- What are the effects for groups of programs, defined according to characteristics of policy interest?

- How do home visiting programs affect health disparities and health care quality?

- What are the needs and capacities of the at-risk communities that states decide to target and how do states plan to use MIECHV funding to address these needs?

- What are the costs for achieving key outcomes for families and children?

Questions that, while not required, would be extremely useful in for future policy, programmatic, and implementation decision-making include the following:

- How do the funded home visiting programs actually operate? What organizations are involved as stakeholders, how are service models defined, how adequate are implementation systems, who provides services, what families are enrolled, and what services are provided?

- How are the different types of inputs of home visiting programs related to one another? How do community context and organizations influence service model clarity and congruence? How do they influence the adequacy of the implementation system? How do the service model and implementation system influence the characteristics of the staff who provide services and the families who receive them?

- How is the full set of inputs related to the services provided to families through home visiting and through referrals to other services? Analyses can address several key questions across program sites: How do service model clarity and implementation system adequacy influence program coverage of the targeted population and service dosage, content, and quality? How do staff and family characteristics mediate these influences?

- Beyond basic characteristics of programs required to establish subgroups, which features of program service models and implementation systems are associated with larger effects on key family outcomes?

Challenges Faced by the National Evaluation

The design of the national evaluation has been shaped by a number of challenges that must be addressed by any national evaluation of home visiting programs.

The first challenge stems from the fact that states can choose from seven different evidence-based models. This is further complicated by the fact that some local sites may blend features of different models, for example, following the basic framework of one national model but using a curriculum from a different model. Key differences across the models may complicate the interpretation of findings. For example, impacts on birth outcomes would be relevant only for a subset of sites serving pregnant women. Likewise, measures of child development will differ for infants and preschool children, so results from sites that include very young children could not be combined with sites that include older children. Although the diversity of program models raises issues about how to interpret results, it also provides an opportunity for the evaluation to systematically investigate the link between program features and program effects.

Another set of challenges are related to the fact that states can use funds provided under ACA either to expand existing home visiting programs or to begin operating new programs. This raises two sets of issues. One is the importance of understanding the home visiting service environment for control group members in every community in which the study is conducted. The existence of home visiting programs outside of MIECHV must be considered in selecting sites and in collecting data about services for control members. The second is whether the evaluation should seek to understand the effects of both new programs and programs that are expanded through MIECHV. While an evaluation that includes all programs would provide estimates of the improvements in maternal and child outcomes purchased with MIECHV funds, it is likely that many new programs will become better implemented over time.[12] Thus, including new programs would lead the national evaluation to understate the long-term effects of home visiting. An evaluation that includes only fully operational programs would have a narrower focus, but it might provide better information on the potential of home visiting to benefit families and children. Even more mature programs might continue to evolve throughout the evaluation, as more is learned about how families and children benefit from their interventions. In many places, families enrolling in a program toward the end of recruitment for the evaluation will experience a different program than those who enrolled earlier. The study must monitor these varying conditions, as they might affect data analysis and interpretation of findings.

---

[12]Rubin (2010).

The thousands of currently operating home visiting programs are highly decentralized. This raises several challenges for the evaluation. First, eligibility criteria and how families are referred to home visiting differ from place to place, even across sites using the same national model. This may create relevant differences in the characteristics of families served at different sites. It will be important for the evaluation to include the full range of at-risk families. Likewise, the evaluation will have to take into account local processes in interpreting results from both the implementation and impact studies. Differences in eligibility criteria and referrals also mean that the evaluation may have to design study enrollment procedures site by site. The decentralized nature of services must also be considered in data collection. For example, much of the information on home visiting services will probably need to be collected from each local program rather than from centralized agencies such as state grantees, although this may vary with the national model being used.

The broad set of domains specified in the legislation may require a variety of different types of data. The easiest option for the evaluation would be to collect data through surveys of parents. However, parents might not accurately report some outcomes such as health care use and health outcomes over the course of a year, or they may have incentives to misreport some information, such as instances of child abuse and neglect or their parenting practices. Thus, collecting information across all domains may require the use of surveys of parents, observations of parents interacting with their children, direct assessments of child development, and administrative data on child abuse, health care use, and birth outcomes.

The final two challenges relate to the timing of particular aspects of the evaluation. States have not yet submitted their updated plans for using MIECHV funds. As a result, the national evaluation is designed without knowing which evidence-based program models will be chosen by states, which states will implement or expand home visiting programs with the new funds, which groups of families will be targeted for home visiting services, and how many families states plan to serve. The design therefore makes assumptions about the number of local programs that will be operated by individual states, the frequency with which the various evidence-based models will be used, and the number of families that will be served by local programs. The design may have to be updated when state plans are reviewed and perhaps again when the evaluator learns more about community context and organizational influences on programs during the process of choosing sites for the evaluation.

A last challenge is to provide relevant information for the required report to Congress in 2015. As discussed below, the time it will take to choose sites, enroll families into the study, and collect, process, and analyze data means that the report to Congress in 2015 might include information only on the state needs assessments, and the characteristics of families, sites, and programs in the evaluation, but might not include any information on the effects of home visiting programs. Full follow-up information needed for the effectiveness study and the cost-effectiveness analysis could be analyzed after the report to Congress has been submitted, as would optional long-term follow-up if the federal government chooses to exercise options for long-term follow-up.

The Proposed Design

To provide unbiased estimates of the effects of home visiting programs, the design assumes families recruited into the study will be randomly assigned to either a MIECHV program or to a control group that could use any other services available in the community. A random assignment design would meet the HHS goal of conducting a rigorous evaluation of home visiting programs. Although the feasibility of carrying out random assignment must be assessed community-by-community, the need for home visiting services is likely to far exceed the capacity of local programs, allowing for the ethical creation of a control group. When a program cannot serve all eligible families, a lottery can be a fair way to allocate scarce slots rather than, for example, accepting all families only until slots are full and then creating a waiting list. Other than MIECHV services, control group members would be eligible for all services available in the community for which they would normally be eligible. The evaluation will adhere to all ethical standards for program evaluation and will undergo human subjects review by an Institutional Review Board.

As discussed further in Chapter 3, the study would include 7,200 families spread across 120 sites (that is, local programs). A typical site would include 30 families assigned to a MIECHV program and 30 control group families, although the exact number of families may vary from site to site. This number of families would provide enough statistical power to examine differences in impacts of home visiting across key subgroups of families. The large number of sites is in part due to the small capacity of most local home visiting programs, but also creates an opportunity to learn about the relationship between local program features and impacts of home visiting.

Within each site, the evaluation would seek to randomly assign families that would be served by a particular home visitor into the program and control groups. This would allow the evaluation to compare the estimated effects of home visiting by characteristics of home visitors, such as prior experience or educational credential. Since this would be feasible only in sites that identify the home visitor before a family is randomized, the evaluation would investigate the feasibility of conducting random assignment by home visitor in each site. This method would be used only if it can be done in enough sites to provide precise effects of the relationship between home visitor characteristics and program impacts.

The 120 sites will be selected to include a diversity of program models, families, and geographic locations across the country. For example, the evaluation would seek to include a broad representation of each of the evidence-based program models to maximize what can be learned about how impacts vary by program features and to ensure that the results do not primarily reflect one or two program models chosen most frequently by MIECHV grantees. However, this might not be feasible if some national models are used by few states. Likewise, the evaluation would seek to include a diverse set of families to provide fairly precise estimates of the effects for subgroups of families. For example, the evaluation would seek to include enough families in underserved groups such as racial and ethnic minorities in order to investigate the effects of home visiting on health disparities.

Although the criteria for choosing sites cannot be finalized until after state plans have been submitted, it is expected that sites chosen for the evaluation would have to meet several other criteria. Since new programs might take time to evolve to their full level of effectiveness,

the evaluation would choose only sites operating mature programs that had existed for at least two years. There would need to be sufficient excess demand for services at the site to allow for the ethical creation of a control group. Again, the design assumes that each site would include 30 families for the control group, although this number could be reduced to some extent without a substantial effect on the design's statistical power. Finally, a site would be eligible for the evaluation only if the local environment was such that control group families would be unlikely to receive similar home visiting services on their own.

Sites would be concentrated in a few states to reduce evaluation costs. For planning purposes, the design assumes that the 120 sites would be spread across 12 states. In addition, sites could be concentrated within certain geographic areas in some states to reduce travel costs associated with various aspects of the research.

Five of the evidence-based models work with pregnant mothers or mothers of infants, but two models enroll families with children two and older.[13] Since it can be difficult to compare many outcomes across a broad range of children's age, the evaluation would estimate effects separately for two groups of families: those who enter while the mother is pregnant or in the first few months after the child is born, and those who enter when the child is between two and four years old. Estimated effects of home visiting would be presented separately for the two groups of families.

Follow-up data would be collected at either six months or twelve months. A decision about the timing of follow-up will be made in consultation with the Secretary's Advisory Committee (SAC). The length of follow-up will determine which families are recruited and when data would be collected. If data are collected at six months, the study would be limited to pregnant women and families with children who are, at the time of enrollment, under three months old or between two and four years old. Infants would be required to be less than three months old to be included in the study so that there would be less variation in child's age when six-month follow-up data are collected. In this case, follow-up data would be collected when the child is six months old for families that enter the study when the mother is pregnant but at six months following random assignment for other families. Data would be collected at the same age for the youngest children because child outcomes may vary substantially by age for very young children. Collecting information at the same age for older children is both less important than for very young children, but also more difficult since children might be anywhere from two to four years old when they enter the study.

If data are collected at twelve months, enrollment into the study would be limited to pregnant women and families with children under six months old or between two and four years old. In this case, follow-up data would be collected around the time of the child's first birthday for families that enter the study when the mother is pregnant or when the child is an infant, but data would be collected one year following random assignment for families with older children.

---

[13]These five models that work with pregnant mothers or families with infants are NFP, HFA, PAT, Early Head Start—Home-Based Option, and Healthy Steps. The two models that work only with older children are HIPPY, which works with parents with preschool children, and Family Check-Up, which works with children two and older.

An impact analysis would estimate the effects of home visiting programs across the range of domains specified in the ACA and for key subgroups of families. Results from the impact analysis would also be used to assess the potential of home visiting programs to reduce health disparities and improve health care quality. As discussed in Chapter 5, an assessment of the effects of home visiting programs on health care practices would be limited to sites where home visiting is embedded in those practices (as would be the case for Healthy Steps) or where the concentration of home visiting is so high that health care providers in the community might be affected by the mere existence of home visiting programs.

An implementation study would collect information on community context, influential organizations, the service model, the implementation system, home visitors, families, and actual service delivery. The implementation study is designed to complement the impact study. It has three main goals: (1) to describe home visiting program inputs and outputs; (2) to determine the associations among inputs; and (3) to investigate how inputs are related to outputs. In addition, the implementation study and the effectiveness study will jointly investigate which features of service models and implementation systems are associated with more positive effects for families.

Data for the implementation and impact studies will be collected from a variety of sources to provide the most reliable evidence possible about home visiting services and their effects on families and children. Data sources would include interviews with parents; observations of the home environment; direct assessments of older children; observations of home visitors in their work with families during home visits; logs, observations, and interviews with home visitors, supervisors, and program administrators; program model documentation from program developers, grantees and local sites; administrative data from program management information systems; and administrative data on child abuse and neglect and birth outcomes.

A cost-effectiveness analysis would estimate the cost of achieving key benefits to families. Although the ACA requires the evaluation to assess the ability of home visiting to reduce health care costs, the evaluation could extend the cost-effectiveness analysis to other key outcomes, such as child development and family economic self-sufficiency. This component of the evaluation will build on results from the impact analysis as well as data on program costs.

Appendix B describes some additional research activities that could be undertaken as part of either the national evaluation or an ongoing research agenda.

Timeline for the Evaluation

The timeline for the study (see Figure 2.1) is assumed to have the following key dates:

- <u>September 2011</u>. It is assumed that a request for proposals will be issued this summer and that the evaluator will be chosen by the end of the current fiscal year.

- October-December 2011. The first few months would include a number of start-up activities, such as finalizing the evaluation design with the most current information on state plans, developing data collection instruments, and disseminating information about the evaluation to states.

- January–December 2012. The design assumes that it will take one year to recruit 120 sites into the evaluation. Other activities that would take place would include submitting a package to the Office of Management and Budget to obtain approval for baseline data collection and obtaining approval from Institutional Review Boards as needed.

- October 2012-June 2014. Sites are assumed to enter the study on a rolling basis. The first families would enter the study in October 2012. Since some sites may take as long as 18 months to enroll 30 families into the program group, the last families would enter the study in June 2014.

- April 2013-December 2015. The main outcome data will be based on a six or twelve-month follow-up survey, observations, and child assessments. If families begin to enter the study in October 2012, data collection would begin in April 2013 for a six-month follow-up and October 2013 for a twelve-month follow-up. Since it may take some time to find and interview families, and because it will take more than one year for children whose parents enroll prenatally to reach their first birthday, data for the last families that enter the study in June 2014 would not be collected until June 2015 for a six-month follow-up or December 2015 for a twelve-month follow-up (six months after their anniversary of entering the study in the latter case).

- March 2015. The ACA requires a report to Congress in March 2015. Because little data from the six-month or twelve-month follow-up would be available at the time, the design assumes that the report will include a description of the evaluation, the families in the study, and the local programs.

- 2016-March 2017. Once all six-month or twelve-month follow-up data are collected by the end of 2015, they would be analyzed and a final report would be issued in March 2017.

Design Options for Home Visiting Evaluation Project

Figure 2.1

Timeline for National Home Visiting Design

	Year 1				Year 2				Year 3				Year 4				Year 5				Year 6		
	2011	2012				2013				2014				2015				2016				2017	
Task	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2
Submit OMB application and obtain OMB approval	X		X																				
Select sites (assumes generic OMB clearance)		X	X	X	X																		
Recruit sample and conduct baseline interviews					X	X	X	X	X	X	X												
Collect implementation data					X	X	X	X	X	X	X	X	X	X	X								
Collect 12-month follow-up data									X	X	X	X	X	X	X	X	X						
Prepare and submit 2015 report to Congress									X	X	X	X		X									
Analyze, prepare, and submit 12-month impact report																		X	X	X	X	X	X

NOTE:

For the six-month follow-up option, data collection would begin in Q3 2013. Data for the last families that enter the study at the end of Q3 2014 would not be collected until the end of Q3 2015.

29

Chapter 3

Sampling Plan

This section describes the number of families and sites that would be included in the evaluation, principles underlying how sites would be chosen, and the statistical power of the proposed sampling plan.

A number of considerations affected the choices described in this chapter. Perhaps foremost was designing a study that would be likely to produce statistically significant estimated effects. In particular, the sampling plan was developed to provide enough statistical power to draw inferences about differences for subgroups of families and to investigate the relationship between program features and program impacts. This is an important consideration because the important unanswered questions about home visiting include the effects for families with particular needs that are served with new program funds and the association between program effects and program implementation. The statistical power of the designed evaluation is discussed later in this chapter.

One particular constraint is imposed by the possibility that home visiting programs could enroll families beginning before the child is born through age 5. Few measures of child development are relevant to both infants and preschool children, and it is unclear whether reliable measures of child development exist for very young children. The sampling plan was consequently developed as if there are two distinct evaluations being conducted, one for sites that serve pregnant women and those with very young children, and one for sites that serve families with toddlers and preschool aged children. The design therefore includes enough families to draw inferences concerning both types of families.

The sampling plan was also affected by the resources expected to be available for the evaluation. Increases in data collection costs associated with additional families, additional sites and additional states imposed some constraints on the total size of the evaluation. For example, the costs of collecting administrative data from state agencies increases with the number of states in the study. Likewise, it would take more time and resources to explain the study to more grantees and obtain their agreement to be involved in the evaluation. The design therefore assumes that sites will be concentrated in a relatively small number of states rather than being spread nationwide.

Number of Sites and Families

The proposed study would be conducted in 120 sites. The large number of sites was chosen for several reasons. First, it is likely that many sites will serve a small number of families, so a greater number of sites is needed to obtain a sample large enough to detect program effects. Second, the relatively large number of sites will make it easier for the study to reflect the diversity of communities and families involved in Maternal Infant and Early Childhood Home Visiting (MIECHV) nationally. Third, many sites are included to enhance the ability of the evaluation to identify the features of local programs that are associated with stronger program effects.

The design assumes that the average site would enroll 30 program group and 30 control group families for a total sample of 7,200 families across the 120 sites. For a site with four home visitors, 30 program group families could be enrolled in a year if each home visitor serves between seven and eight new families each year. Prior studies suggest such enrollment levels could be achieved in about a year for programs adding families to the caseloads of existing home visitors. Where MIECHV funds are used to expand programs, more slots may be available for new families to receive home visiting services, which would reduce the time needed to recruit families into the study.

Minimum Detectable Effects

The statistical power of the proposed sampling plan was assessed using a concept called *minimum detectable effect.* A minimum detectable effect is the smallest true effect that is likely to generate statistically significant estimated effects. For purposes of the design, calculations were performed to find the smallest effects that would generate statistically significant findings in 80 percent of studies with a similar design, using two-tailed t-tests with a 10 percent significance level.[14] As noted earlier, families are assumed to be assigned in equal proportions to the program and control groups because this results in the greatest statistical power of the study. However, relatively small deviations from this allocation would have only slight effects on statistical power. For example, assigning 60 percent of families to the program group would increase minimum detectable effects by only 1 percent (for example, from 0.10 standard deviations to 0.101 standard deviations).

Impact Estimates Pooled across Sites

As noted elsewhere, the evaluation would estimate separate effects for families enrolled before or near the child's birth and for those enrolled when the child is a toddler or preschool age. For purposes of determining the statistical power of the evaluation, it is assumed that 85 sites would enroll families before or near the child's birth while 35 would serve families with older children.[15] Although it is expected that many states will include programs for infants and pregnant mothers in their plans, it is less certain that many states will include programs for older children in their plans. If few states plan to use MIECHV funds for children two to four years old, the evaluation might exclude this group and focus on infants and prenatal families.

Table 3.1 shows minimum detectable effects for results pooled separately for the two groups of sites. All results are presented as effect sizes, that is, in terms of number of standard deviations of the outcome being examined. Results are presented both for administrative data,

---

[14]Although many disciplines assess statistical significance at the 5 percent level, the design used the 10 percent level for two reasons. First, conventions about statistical significance are not universal and many prior studies have assessed significance at the 10 percent level. More important, for making policy decisions, it can be useful to know that a result is significant at a level between 5 and 10 percent. The evaluation should report the exact significance of results using p-values or standard errors to minimize the importance of deeming one specific level as being "significant."

[15]Several of the evidence-based models serve families in both groups. It is therefore likely that some sites would include families in both groups. To simplify the discussion, this is assumed not to be the case, but the essential features of the design are not affected by this assumption.

which would be available for all families, and for data such as surveys and observational interactions between parents and children, which are assumed to be available for 80 percent of families.[16] In addition, results are presented for the case in which baseline family characteristics do not increase the precision of estimated effects and for the case where those characteristics explain 30 percent of the variation in outcomes across families. This provides a range of estimates since the ability of baseline characteristics to increase the precision of estimated effects may vary from outcome to outcome.

For the 85 sites assumed to serve families near the time of the child's birth, the minimum detectable effect would be 0.070 standard deviations for the pooled sample for administrative records and 0.078 for data such as surveys or other data types provided for 80 percent of families. For example, if a site had a rate of child abuse and neglect of 20 percent in the control group, this design would have an 80 percent chance of finding a statistically significant impact if the true impact is a reduction of 2.8 percentage points (from 20 percent of the control group to 17.2 percent of the program group).[17] This calculation assumes that information on baseline family characteristics would have no effect on the precision of the estimated effect, which is a very conservative assumption. If baseline family characteristics explained 30 percent of the variation in outcomes across families, the minimum detectable effects would decrease by about 16 percent, to 0.058 for outcomes measuring using administrative data and 0.065 for outcomes measured using survey data. Using this less conservative assumption, the study would have an 80 percent chance of finding a statistically significant effect on substantiated cases of child abuse and neglect if the true impact is a reduction of 2.3 percentage points (to 17.7 percent of the program group).

For the 35 sites serving families with two to four year old children, the minimum detectable effect would be 0.109 standard deviations for results from administrative records and 0.121 for survey-based results with an 80 percent response rate, assuming baseline family characteristics do not improve the statistical power of the estimates. For a child abuse and neglect rate of 20 percent, this would give the evaluation an 80 percent chance of finding statistically significant effects if the true effect is 4.3 percentage points (from 20 percent of the control group to 15.7 percent of the program group). If baseline family characteristics explained 30 percent of the variation in outcomes across families, the minimum detectable effects would decrease by about 16 percent, to 0.091 for outcomes measured using administrative data and 0.101 for outcomes measured using survey data. In other terms, the study would have an 80 percent chance of finding a statistically significant effect on substantiated cases of child abuse and neglect if the true impact is a reduction of 3.6 percentage points (to 16.4 percent of the program group).

---

[16]A response rate of 80 percent is assumed for surveys because this is the standard set by the federal Office of Management and Budget. Such response rates have been achieved in numerous studies of home visiting and other evaluations with similar target populations.

[17]A rate of 20 percent was chosen because home visiting is unlikely to find statistically significant impacts if substantiated cases of child abuse and neglect are rare. In addition, 20 percent is a reasonable rate given studies such as Duggan et al. (2007), which found substantiated cases of child abuse and neglect for 17 percent of control group families.

Design Options for Home Visiting Evaluation Project

Table 3.1
Minimum Detectable Effects of Proposed Home Visiting Evaluation

Scenario	Number of families		Administrative data	Survey or observational data
	Program Group	Control Group		
85 sites serving pregnant women and families with infants				
No baseline covariates	2,550	2,550	0.070	0.078
Covariates explain 30 percent of variation in outcomes			0.058	0.065
35 sites serving families with older children				
No baseline covariates	1,050	1,050	0.109	0.121
Covariates explain 30 percent of variation in outcomes			0.091	0.102

These pooled minimum detectable effects provide reasonable statistical power for the evaluation given prior evidence of effectiveness from Home Visiting Evidence of Effectiveness (HomVEE). Table 3.2 shows the range of effects across studies and outcome measure, the average effect in each domain weighted by sample size, and the number of effects included in the calculation. The range and average are presented as effect sizes, or number of standard deviations for the given outcome.

Results summarized in Table 3.2 were restricted to those that were considered primary by HomVEE. This means that the results are limited to direct observations, direct assessments, administrative records, and self-reported data using standardized instruments. In addition, results were restricted to those for which an effect size was available, either from the original study or calculated by HomVEE.[18] Because effect sizes were not available for most Nurse Family Partnership (NFP) studies and many Healthy Families America (HFA) studies, the results presented here may understate the true average effects of studies included in the HomVEE review.

As noted in Chapter 1, prior results vary substantially from study to study and sample to sample. For each domain, the results range from roughly -0.5 standard deviations to 0.5 standard deviations. In part, the wide range stems from the small samples used to calculate many of the

---

[18]For parenting practices, results from PAT on days attended were excluded. Results for child development and school readiness were limited to measures that are likely to be included in the national evaluation: the Bayley, the Peabody Picture Vocabulary Test, the MacArthur CDI, Woodcock-Johnson, SSRS, and Bracken. Most of the outcomes removed by this restriction measured cognitive outcomes of school-age children, who would not be represented in the national evaluation as it is currently designed.

effects. For example, the effect of -.36 standard deviations on positive parenting practices is from a study with 246 families, while the effect of .49 is from a study with 92 children.

When results are averaged across studies and outcomes and weighted so that larger samples have more influence on the results, the average effect is largest for maternal health, where the average effect size is 0.17 standard deviations, and for referrals and coordination, where the average effect size is 0.14. By contrast, the average effects for positive parenting practices, child maltreatment, and child development and school readiness are all about .03 to .04 standard deviations. While the proposed design would easily detect the larger effects, it probably would not generate statistically significant effects if the true effect of home visiting is 0.03 standard deviations.

Design Options for Home Visiting Evaluation Project

Table 3.2

Summary of Results from the Home Visiting Evidence of Effectiveness Review

Domain	Range	Average	Number of effects
Positive parenting practices	-0.36 to 0.49	0.03	40
Child maltreatment	-0.45 to 0.30	-0.03	14
Child health	-0.50 to 0.43	0.08	24
Child development and school readiness	-0.14 to 0.34	0.06	26
Maternal health	-0.34 to 0.80	0.17	13
Referrals and coordination	-0.62 to 0.67	0.14	18

Differences in Estimated Effects across Subgroups

In addition to looking at the average effect across sites, the evaluation would assess whether home visiting had larger effects for some subgroups. For purposes of investigating the statistical power of subgroup estimates, it is assumed that the evaluation would be interested in detecting significant differences across subgroups.[19] For example, if the estimated effect on parenting practices were 0.20 standard deviations for women who entered the study while

---

[19]Bloom and Michalopoulos (2010).

pregnant and 0.10 standard deviations for women who entered the study soon after childbirth, the evaluation would ask whether those two estimates were significantly different from one another. If they were not significantly different from one another, the evaluation would not have strong evidence of a differential effect for one group compared with the other.

Table 3.3 presents minimum detectable differences between subgroups for cases where families in the 85 sites serving parents near the time of childbirth are divided into two groups (for example, mothers who are pregnant when they enter the study compared with those who enter the study when their child is an infant). Since statistical power depends on the number of families in each subgroup, minimum detectable differences are presented for four cases: (1) half the sample is in each subgroup, (2) 60 percent of the sample is in one subgroup, (3) 70 percent of the sample is in one subgroup, and (4) 80 percent of the sample is in one subgroup. As in Table 3.1, results are presented once using the assumption that baseline information does not improve the precision of estimated effects and once assuming that baseline information explains 30 percent of the variation in outcomes across families.

Consider a subgroup that divides the sample in half. The minimum detectable differences range from 0.117 using administrative data when baseline information is useful to 0.156 standard deviations for outcomes measured using survey data when baseline information does not increase statistical precision. If 20 percent rate of control group families had a substantiated case of child abuse and neglect, the study would have an 80 percent chance of finding significantly larger effects for one subgroup than for another if the difference in true effects was 4.7 percentage points (for example, reducing child abuse and neglect by 4.7 percentage points for one subgroup but having no effect for the other subgroup).

These minimum detectable differences increase gradually as the proportion of families in one subgroup increases. They are quite similar if 60 percent of families are in one subgroup, but increase by 25 percent if 80 percent of families are in one subgroup.

Exploring the Relationship between Program Features and Program Impacts

In addition to estimating the average effect of home visiting programs and effects by subgroup, the evaluation would include 120 sites so that it could explore the relationship between program features and program impacts. Program features could include any aspect of the community context, implementation system, service models, organizational influences, or home visitor characteristics that will be described in Chapter 4. For example, this analysis could explore how program impacts vary with the duration of home visits, the background and training of home visitors, the support provided by supervisors for home visitors, the clarity of the goals of the local program, the intended targets of the national model being used, and so on.

Design Options for Home Visiting Evaluation Project

Table 3.3
Minimum Detectable Differences Between Subgroups of Families
For 85 Sites Serving Families with Infants

Scenario	Percent of sample in one subgroup	Administrative data	Survey or observational data
Baseline family characteristics do not improve statistical precision			
	50	0.139	0.156
	60	0.142	0.159
	70	0.152	0.170
	80	0.174	0.195
Baseline family characteristics explain 30 percent of variation in outcomes across families			
	50	0.117	0.130
	60	0.119	0.133
	70	0.127	0.142
	80	0.146	0.163

Results are the smallest true impact that would generate statistically significant impact estimates in 80 percent of studies with a similar design using two-tailed t-tests with a 10 percent significance level. No adjustment for multiple comparisons is assumed. Results are based on fixed effects estimates. Adminsitrative data are assumed available for all families, while survey or observational data would be available for 80 percent of families.

A framework for exploring the link between program features and program impacts is described in Greenberg, Meyer, Michalopoulos and Wiseman (2003). Within this framework, the precision of the estimated relationship between program features and program impacts depends on a number of factors, including (1) the number of sites in the evaluation, (2) the precision of impact estimates within each site (which will increase with the number of families in the site), (3) the variation in characteristics across sites, (4) the number of program features to be investigated, and (5) how related the various program features are to each other. It is easier to detect differences by program feature if there are more sites, if there are more families in each site, if different sites vary more across the program feature being examined, if fewer program features are being examined at any one time, and if the program features are not closely related to one another. As an example of the last point, it may be very difficult to distinguish the effect of planned duration of home visits from the effect of actual duration since the two are likely to be closely related in a particular site.

Table 3.4 shows the minimum detectable effects of program features for several scenarios. The top half of the table shows results for a program feature that is binary and takes on one value in half of the sites and a different value in half of the sites. For example, half of the sites might plan to visit families weekly while half would visit only every other week. The bottom half of the table shows results for a continuous program feature, such as how many weeks home visits would take place. In each panel, results are presented depending on whether

10, 20, or 30 program features would be examined at one time. As noted above, the ability to detect the effects of program features will worsen as more features are examined. Finally, results for each scenario are presented for three assumptions about how highly correlated various program features are with one another. As noted above, the ability to detect the effects of program features worsens as features become more highly correlated with one another.

Design Options for Home Visiting Evaluation Project

Table 3.4
Minimum Detectable Effects of Program Features
For 85 Sites Serving Families with Infants

Type of variable	No. of variables representing program features	Correlation across program features	Administrative data	Survey or observational data
Binary, half of sites have the feature				
	10	Low	0.204	0.232
		Medium	0.215	0.244
		High	0.227	0.259
	20	Low	0.232	0.264
		Medium	0.265	0.302
		High	0.318	0.363
	30	Low	0.269	0.307
		Medium	0.350	0.399
		High	0.630	0.717
Continuous				
	10	Low	0.102	0.116
		Medium	0.107	0.122
		High	0.114	0.130
	20	Low	0.116	0.132
		Medium	0.133	0.151
		High	0.159	0.181
	30	Low	0.135	0.153
		Medium	0.175	0.199
		High	0.315	0.358

Consider the first row of the table, which shows the case where 10 program features are being examined simultaneously and there is a low correlation across them. For outcomes measured using administrative data, the model would be able to detect differences of 0.204 standard deviations between sites of one type and sites of another type. If the overall effect on an

outcome were 0.15 standard deviations, for example, the study would have an 80 percent change of finding a statistically significant relationship between the program feature and impacts if the true impact were 0.252 in one set of sites and .048 in the other set of sites.

The ability to detect an effect of a program feature is only slightly worse if the features are more highly correlated, or if 20 program features are being examined. The statistical power gets considerably worse, however, if more features are being examined and the correlation across features is high. For example, the minimum detectable difference is 0.318 (for example, an effect of 0.309 standard deviations in one set of sites compared to -0.009 in the second set of sites) if 20 program features are being examined and the correlation across them is high, and 0.350 if 30 features are being examined and the correlation across them is medium.

Although it is not shown in the table, minimum detectable effects increase fairly modestly if sites are not equally divided by the program feature. For example, they would increase by about 2 percent if 60 percent of the sites fall into one category, by about 9 percent if 70 percent of the programs fall into one category, and by about 25 percent if 80 percent of the programs fall into one category.

The second half of the table shows minimum detectable effects if the program feature is continuous and normalized to have a variance of 1.0 across sites. Because there can be greater variability in continuous variables than in binary ones, the design would have a greater ability to detect differences for such measures. For example, for a study examining 10 program features that are not highly correlated, the minimum detectable effect size of the program feature would be 0.102 using administrative data and 0.116 using survey data. Even for the most extreme case shown in the table – 30 highly correlated program features – the design could detect differences in impacts of 0.315 using administrative data and 0.358 using survey data.

Choosing Sites

The process for choosing sites cannot be finalized until after state plans for using MIECHV funds have been finalized Those plans will provide information on where home visiting programs will operate, the national program models that will be used, and the number and types of families that will be served. Even so, some principles of site selection can be presented at this time.

Principles of Site Selection

Sites would meet four criteria to be eligible for inclusion in the evaluation. First, they would have to be operating local programs that were expanded rather than created using MIECHV funds. In particular, the evaluation would be limited to sites that have been operating local programs for at least two years when decisions are made about which sites will be in the evaluation. This restriction is made because many new programs will become better implemented over time, and including them in the evaluation would understate the long-term effects of home visiting.[20] Congress will consequently learn more about the potential for home visiting to achieve the goals described in the legislation if the national evaluation measures the

---

[20]Rubin (2010).

impacts of programs that are in a relatively "steady state" of operations rather than in an early start up period. Because some national models are currently used in only a few places, the evaluation may want to oversample some national models in order to increase the diversity of models in the study.

A site would be eligible for the evaluation only if it could recruit enough families to fill program slots and to allow for a randomly chosen control group. For example, the evaluation might focus on sites that receive more referrals for home visiting services than they can meet, such as those with waiting lists for home visiting services. This could exclude very small sites as well as those serving all families identified as eligible for services. To ensure representation of small or rural communities in the national evaluation, several small communities in one state might be combined to form one site for purposes of the evaluation..

The evaluation will also prefer locations where fewer control group members would be expected to receive home visiting services. This would allow for a clear differential between the home visiting services received by the program group and any services received in the community by the control group. Because some control group members are likely to receive home visiting in many places, the evaluation would collect information on the extent to which control group members are receiving services and some features of those services. Sites that are chosen for the study might later be dropped from the study if little difference found in service receipt between program group and control group families. Data to be collected on control group members will be described in Chapters 4 and 5.

Finally, sites would be chosen based on their contribution to the diversity of program models and family characteristics. To ensure that estimated effects are not dominated by one or two program models and to maximize what the field learns about how variation in program inputs is associated with variation in program impacts, the evaluation should seek to include a similar number of sites for each of the seven evidence-based program models. This assumes that enough states are choosing to use each of the evidence-based models. If fewer than, say, ten states choose to use MIECHV funds for a particular evidence-based model, that model may be excluded from the evaluation. That is, states would not be selected solely for purposes of including some sites operating rarely chosen models.

Sites would also be chosen to ensure diversity of families and communities included in the evaluation. Since one of the goals of the evaluation is to understand the effects of home visiting on health disparities, the evaluation would include sites that ensure large enough samples of underserved groups to investigate health disparities. For example, the evaluation should be sure to include enough Latino and African-American families to obtain precise estimates of home visiting's effects on health care use and health outcomes.

The proposed design assumes that sites will be concentrated in a relatively small number of states to reduce administrative costs of the evaluation and data collection. In particular, the design assumes that the 120 sites would be spread across 12 states. The first step in choosing sites, therefore, will be to choose which states to include in the evaluation.

Preference would be given to states with sites running different program models. This would avoid the possibility that a particular program model appears to be effective because it is adopted by a state with a well-run administering agency. Having multiple program models being used in a state would allow the evaluation to look for differences across program models, holding constant the characteristics of the state administering agency. According to a review of program models being used by various states conducted by Pew, a number of states currently include more than one program model, so this assumption may be a reasonable one.

States will be chosen to build geographic and demographic diversity into the study. For example, the evaluation could choose at least one state from each of the ten Administration for Children and Families (ACF) and Health Resources and Services Administration (HRSA) regions, assuming each region contains at least one state that would meet the other criteria for site selection. In addition, sites would be chosen within states to provide diversity in race, ethnicity, and other key demographic characteristics, such as teen parents, to provide enough families of various types for precise subgroup estimates.

To achieve the desired diversity across family characteristics and program models, sites and states would be chosen using stratified sampling. Although the exact criteria for doing this would be determined after state plans have been updated, here is an example of how it might work. First, to promote geographic diversity, states would be divided into the ten ACF and HRSA regions. States would be characterized in terms of the number of home visiting sites that would meet the criteria described above (such as having a program that has been in operation for at least two years and can contribute 60 families to a random assignment evaluation) and the number of families in those sites. Since the average state would contribute 10 sites and 600 families, a state would be excluded from the evaluation if it could not be expected to contribute at least 5 sites and 300 families to the evaluation. States would then be characterized in terms of key characteristics of families that would be served with home visiting funds. Since health disparities are especially problematic for low-income families and families in certain racial and ethnic groups, preference would be given to states that are likely to contribute a substantial number of low-income and minority group families. States without a sufficiently diverse set of families or with too few low income families would be excluded from the evaluation. Finally, priority would be given to states that are using MIECHV funds for more than one evidence-based model. If more than one state in each region met the criteria for size, characteristics of families that would be served, and number of evidence-based models being used, one state from that region would be chosen randomly to be included in the evaluation.

Stratified sampling can also be done within a state to ensure enough families with key characteristics are included in the sample. Once a state has been chosen for the evaluation, potential sites can be characterized in terms of the characteristics of the families they serve, especially race, ethnicity, and income. Sites would receive higher priority for being in the study if they served primarily low-income families. In addition, sites would be grouped according to the main racial or ethnic group being served, or they would be classified as serving a diverse group of families. If the state has more sites meeting those criteria than are needed for the evaluation, the evaluation would randomly choose one or more site from each grouping based on race and ethnicity of families that are expected to be served.

In principle, stratified sampling can also be done within each site to enhance the diversity of the families in the study. If a site has more families available than are needed for the evaluation, families that contribute to the diversity of the sample could be oversampled and random assignment could be conducted within each demographic strata. However, it might be difficult to obtain program administrators' cooperation with the evaluation if they believe certain subgroups of families are more likely to be denied services than others.

Analysis of Needs Assessment Data

As described in Chapter 1, states must submit home visiting plans – including a detailed needs assessment for their identified at-risk communities – in order to receive MIECHV funding for fiscal year 2010. In addition to fulfilling one of the requirements of the Affordable Care Act (ACA), the analysis of state needs assessments and state plans may provide information for choosing sites for the evaluation.[21]

The state plans will be critical to the site selection process because they will provide information on where home visiting programs will operate, the national program models that will be used, and the number and types of families that each state expects to be served. In addition to informing the site selection process, the state plans and detailed needs assessments will be used to fulfill the ACA requirement for an analysis of the needs assessment data. The legislation requires "an analysis, on a State-by-State basis, of the results of such assessments, including indicators of maternal and prenatal health and infant health and mortality, and State actions in response to the assessments." To fulfill this requirement, the evaluation would include the following:

(1) A set of state-by-state charts that summarize community needs reported in the needs assessments; existing services in those communities; and plans to fill the gap between needs and services. The charts should have one row for each state with columns for 1) each of the 16 indicators of at-risk communities that states were required to report on in their needs assessments; 2) each of the 8 types of information requested on the quality and capacity of existing programs/initiatives for early childhood home visiting in the states; and 3) the major components of the state plans (summarized at a high level). While the chart would be organized at the state-level, the information presented in the chart would focus on the at-risk communities that each state chooses to target.

(2) A narrative that will provide a) a description of community needs, services and plans across states and any patterns that emerge, and b) a description of how community needs compare to existing services and to the Grantees' future plans. This narrative will not focus on identifying states that did a particularly good or bad job at developing the needs assessments and state plans.

In addition to informing the sampling plan for other components of the national evaluation, the design described above would organize and summarize the needs assessment and state plan data, creating a user-friendly summary that would help policy makers understand the

---

[21]In addition to the state plans for FY2010 MIECHV funding, the evaluator will have state plans for FY2011 funding to help with site selection.

state of home visiting programs in the U.S., and compare the home visiting landscape across states. In addition, this design could reveal common issues and barriers to effective service expansion or provision that face a number of different states - information that could be used to inform later federal funding for home visiting programs.

<div align="center">

Chapter 4

Implementation Study

</div>

As discussed in Chapter 2, the goals of the national evaluation include the systematic study of program implementation to describe what services are delivered and to provide information to strengthen future home visiting programs. This chapter describes the design of implementation research for the national evaluation. It begins by providing an overview of the relevance of implementation research to home visiting. It then introduces the overall implementation study design and broad research questions. The chapter then provides detail on each of the specific research questions, describing the conceptualization of constructs, the measurement plan, and the analytic plan for each of the broad research questions.

The legislation requires the evaluation to examine "the extent to which the ability of programs to improve participant outcomes varies across programs…" Addressing this requirement requires collecting information on characteristics of programs. A further goal that could be met by a more in-depth implementation study is to inform future policy, programmatic, and implementation decisions.

Relevance of Implementation Research for Home Visiting

Prior studies of home visiting programs have found modest and variable impacts, but it is not clear why. A strong implementation study can supplement the impact analysis study and begin to fill this gap in the research.

For example, there is strong evidence that program implementation is related to program effects. A review of over 500 studies of prevention and health promotion programs for children and adolescents found that mean effect sizes were at least two to three times higher when programs were carefully implemented and free from serious implementation problems.[22] However, very little evidence exists about the degree to which home visiting programs are well implemented. The national evaluation can begin to build a body of evidence by systematically studying the quality of implementation of home visiting programs across the 120 sites.

Although little evidence exists about the quality of implementation of home visiting programs, the existing evidence suggests there are high levels of unintended variability.[23] The national evaluation can advance the field by more thoroughly studying the degree to which programs deviate from intended services and the factors that are associated with unintended variability. Doing so may suggest ways of limiting unintended variability and thus achieving more consistent effects across programs.

In addition to the quality of implementation, the effects of home visiting are presumably linked to the services that families receive. However, most published reports of home visiting programs do not describe the actual services that families receive. Instead, researchers such as

---

[22]Durlak and DuPre (2008).
[23]Stavrakos, Summerville, and Johnson (2009); Duggan et al. (2007); Duggan et al. (2004); Duggan et al. (2004); Duggan et al., (1999).

Sweet and Appelbaum (2004) have been restricted to examining the relative influence of factors such as "planned" duration and activities. Given empirical evidence of the gap between planned and actual services,[24] an important goal of the study will be to understand how services are actually provided and which program features are associated with greater impacts.

If the national evaluation is able to provide a better understanding of which program services are associated with larger program effects, it will also be important to provide information to future program operators on how home visiting service models and implementation systems are linked to how services are provided. For example, the effects of parent training programs on parenting behavior and child's externalizing behavior have been linked to specific program components and service delivery strategies.[25] In addition, the implementation science literature supports the importance of implementation system components such as training, supervision, and technical assistance.[26] What is lacking is a systematic approach to measuring and testing the associations of service models and implementation systems with actual service delivery and program impact on outcomes across the nationally disseminated evidence-based home visiting models.

Lastly, the effects of home visiting reported in prior studies may be related to the control group's receipt of similar services. Thus, implementation research can be used to measure home-based and related services received by control group families in national evaluation sites.

Overall Implementation Study Design and Broad Research Questions

The implementation study addresses three broad research questions. Table 4.1 lists the broad research questions and identifies the components of the conceptual framework and study design used to answer each broad question. For each broad research question, this chapter: 1) identifies specific questions and, where applicable, states hypotheses and their rationale; 2) conceptually defines constructs in the components used; 3) describes the measurement plan for each construct; and 4) describes the analytic plan. In addition to the three broad research questions, the implementation study and impact analysis will jointly investigate which program features are associated with program impacts. That analysis is described in Chapter 5.

Design Options for Home Visiting Evaluation Project

Table 4.1

Broad Research Questions, Components of Framework Included, and Study Design for the Implementation Study

Broad Research Question	Components of Framework	Study Design
1. How do sites actually operate?	Inputs and outputs	Descriptive
2. How do inputs relate to one another?	Inputs	Analytic, cross-sectional
3. How do inputs influence outputs?	Inputs and outputs	Analytic, longitudinal

---

[24]Elliot and Mihalic, (2004); Ennett et al., (2003); Hallfors and Cho, (2007).
[25]Kaminski, Valle, Filene, & Boyle (2008).
[26]Fixsen et al. (2005); Durlak and DuPre (2008).

Broad Research Question 1: How do programs actually operate?

The first broad research question will seek to address the following questions:

- What is the community context in which home visiting programs operate?
- What stakeholder organizations are involved?
- How are service models and implementation systems defined?
- What are the characteristics of home visitors and supervisors?
- What are the characteristics of families enrolling in home visiting?
- What services are actually provided to families?
- How do actual services differ from intended services?

Conceptualization of Constructs

Community Context: The community context in which home visiting programs are implemented can play a significant role in both the implementation of the program and outcomes achieved by the program. For example, the extent to which a community has a broad array of supportive services and resources available to meet needs of clients outside the scope of the home visiting program has implications for the success of referrals and comprehensiveness of wraparound care. In addition, research has established a link between community and neighborhood factors and a variety of family development and well-being outcomes.[27] For example, community buy-in and support for a particular program has been associated with program implementation and positive program outcomes.[28] For this reason, community-level factors should be measured and incorporated in hypothesis testing as covariates and potential moderators of home visiting service delivery and impact.

Influential Organizations: An implementing agency provides home visiting at one or more sites. Typically, a site serves a specific geographically-defined community. An implementing agency works with other organizations to provide these services. The implementing agency works with the national model purveyor to develop the infrastructure to provide services. It works with other community-based organizations to arrange referrals and coordinate services. Finally, it works with public agencies and funders to secure funding, promote coordination, and conform to regulations.

Service Model and Implementation System: Together, the implementing agency and the other organizations with which it works adopt and adapt an evidence-based home visiting model for a specific site. For any given site, a home visiting model has two primary aspects: the service model and the implementation system. Thus, a given site aims to implement the national home visiting model it has adopted as it is modified and adapted by other influences.

The defining features of the service model are its specifications for: 1) goals and intended outcomes; 2) eligible families; 3) intended service dosage; 4) intended service content; 5) intended approach to service delivery; and 6) staffing. Service dosage includes aspects such as

---

[27]Molnar, Buka, Brennan, Holton, and Earls, (2003); Xue, Leventhal, Brooks-Gunn, and Earls, (2005); and Brooks-Gunn, Duncan, Klebanov, and Sealand, (1993).
[28]Durlak and DuPrees (2008); Bond et al. (2009).

intended visit frequency and length and intended duration of family enrollment. Service content includes aspects such as information that is intended to be covered during home visits, specific activities that should be carried out, intended referrals to and from other services, and intended coordination of home visiting with other services. Home visiting models specify how services are to be delivered. Examples include the use of a strengths-based, family empowerment, and shared decision-making approaches. Staffing includes qualifications for hiring staff, staff roles and responsibilities, required competencies, and caseload limits.

Sites vary not only in how these service model attributes are defined, but also in how clearly and coherently they are defined with respect to each outcome specified in the ACA national home visiting program legislation. Insofar as influential organizations define the service model fully and clearly, the model has clarity. As these organizations expand the service model's intended outcomes, activities, staffing configurations and other factors, the model increases in complexity. Insofar as service model features are logically related to intended outcomes, the model has coherence.

The implementation system refers to the policies, procedures and resources to implement the service model. Like the service model, the implementation system is influenced by organizations. These organizations can include the model developer and purveyor, who might provide the implementing agency with training, technical assistance, and a management information system. Still, influential organizations will not be limited to the model developer and purveyor. Thus, implementation systems will vary from one site to another, even among sites adopting the same national evidence-based model. The defining features of the implementation system can be categorized as: 1) policies and procedures for staff selection, training, supervision and evaluation; 2) facilitative clinical supports; 3) facilitative administrative supports; and 4) systems interventions. Facilitative clinical supports include screening and assessment tools, protocols, and curricula; the availability of peer support; and the availability of professional consultation to home visitors for situations that require expertise beyond that of the home visitor. Facilitative administrative supports include the availability and use of a management information system and continuous quality improvement procedures to monitor and promote adherence to the service model. The implementing agency's organizational culture and climate, including supportive policies for implementation and acceptance of the program, also influence the level of adherence to the service model.[29] Systems interventions include formal agreements and shared information systems that make it easier for staff to link families with needed services and coordinate services.

Home Visitor Characteristics: Home visiting providers also influence actual service delivery and, hence, fidelity. Home visitors vary in their understanding and acceptance of each of the responsibilities of their role. They also vary in their actual and perceived capacity to carry out each responsibility, both in general and in the context of challenging situations. Home visitors vary in their responsiveness to training, supervision and evaluation activities.

Family Characteristics: Families themselves influence how services are delivered. Family responsiveness to home visiting refers to how family members (especially parents) react to or engage in program activities, particularly those occurring during home visits, but also including

---

[29]Bond et al. (2009).

other activities such as recommended referrals or suggested parenting behaviors. Aspects include level of interest and willingness to engage in discussion and to follow through on suggested behaviors.

Actual Services: There are three main aspects to actual service delivery – dosage, content and quality.

- Dosage is the frequency, intensity, and duration of services to which a family is exposed. In home visiting, dosage is commonly measured by indicators such as the number of visits, the length of each visit, and the duration of family enrollment in a program.

- Content refers to the information conveyed in home visits as well as the methods used to convey that information, and the set of activities carried out in home visits. For a specific intended outcome, content can be measured as the information provided and the activities carried out to achieve the outcome. For promoting child development, for example, content might include parenting education on child developmental milestones, periodic developmental screening, and modeling, role playing, and reinforcement of positive parenting techniques. For a specific service model, some of these activities might be specified as core components.

- Quality of delivery is the manner in which a home visitor provides services. This construct can include the home visitor's interaction style, responsiveness to family members' questions and concerns, adherence to program protocols in challenging situations, ability to tailor services and to motivate behavior change, and cultural appropriateness.

Fidelity is the extent to which actual services conform to what is specified in a service model.[30] Fidelity of dosage is the amount of service actually received relative to what is called for in a service model. For example, fidelity of dosage can be measured as the ratio of the actual number of visits to the number specified in the service model. Fidelity of content is adherence in carrying out activities that are specified in a service model. For example, adherence in conducting developmental screening can be measured as the proportion of families where required screenings are carried out as specified in the service model. Fidelity of service quality is the extent to which activities are delivered in the intended manner. For example, quality of delivery to a strengths-based service model can be measured as the extent to which services reflect the defining attributes of that approach.

Measurement Plan

Community Context: Although measurement of community context is important, researchers have noted the challenge of doing so in a precise, robust, and unbiased manner.[31] The measurement of community context in the national evaluation can and should draw from lessons learned and the recommendations for improving measurement identified in the growing body of research in this area. The evaluator will collect data related to community and neighborhood

---

[30]Mowbray, Holter, Teague, and Bybee (2002); Dane and Schneider, 1998.
[31]Duncan and Raudenbush (2001).

characteristics which are known to affect health and development more broadly. These data should not only include quantitative data about communities, such as census data, but also include emerging data on features such as the built environment. Furthermore, this study can take advantage of emerging analytic techniques, such as geographic information systems, to create community-level profiles of resources and opportunities. In addition, primary data can be collected on the availability of resources and support for the program in the community through interviews and web-based surveys with key administrators, staff, and community stakeholders. Data regarding perceptions of community norms and capacity can be collected through maternal interviews.

Service Model and Implementation System: Table 4.2 lists candidate constructs to be measured regarding service model and implementation systems, as well as potential data sources and the likely timing of data collection. As shown, the evaluator will collect baseline information on community context, influential organizations, and the service model and implementation system as each site enters the evaluation. This could take place during visits to recruit sites into the evaluation, to explain the evaluation, or to train staff in random assignment and other evaluation activities. It could also happen through web-based surveys with key administrators and staff, using follow-up by phone as needed. It is likely that service models and implementation systems will evolve over time. Thus, the national evaluation should update this information annually as part of the implementation study.

Design Options for Home Visiting Evaluation Project

Table 4.2

Data Collection to Measure the Service Model and Implementation System

	At Site's Entry into Study	Annually Thereafter
Influential Organizations		
Identification of organizations	X	X
Working relationships	X	X
Service Model[1]		
Goals and Intended Outcomes	X	X
Eligible Families	X	X
Intended Overall Dosage	X	X
Intended Direct Services for Each Outcome Domain	X	X
Intended Referrals/Coordination for Each Outcome Domain	X	X
Staff Roles and Competencies; Caseload Limits	X	X
Implementation System[1]		
Staff Recruitment and Hiring	X	X
For Each Outcome Domain:		
Staff training, supervision, evaluation, feedback	X	X
Facilitative clinical supports	X	X
Facilitative administrative supports	X	X
Systems interventions	X	X

NOTE: [1]Measurement includes source for each, for example if defined or provided by program developer/purveyor

Supervisor and Home Visitor Attributes. Table 4.3 lists candidate constructs and the likely timing of data collection to measure characteristics of both home visitors and supervisors. The home visitor influences the activities that take place during home visits. Through supervision and feedback on performance, the supervisor influences home visitor behavior.

As shown in Table 4.2, the evaluator will collect information on home visitors and supervisors at multiple points in time. Close to the site's entry into the evaluation – or as soon as new home visitors and supervisors are hired after a site has entered the national evaluation – staff would be asked to complete a web-based questionnaire to provide background information on their education, prior experience, and demographics. The survey would also assess the staff member's psychological well-being. For each outcome domain, the survey should elicit the staff member's understanding of the service model, self-efficacy in carrying out role functions, knowledge and skills required to carry out roles and responsibilities, and ratings of the implementation system for each outcome domain. Around the same time, supervisors would be asked to assess how well home visitors respond to training, supervision, and evaluation.

The evolution of these staff member characteristics will be tracked through annual follow-up web-based surveys with home visitors and their supervisors.

Design Options for Home Visiting Evaluation Project

Table 4.3

Data Collection[1] to Measure Home Visiting Program Staff Attributes

	At Site's Entry into Study[2]	Annually Thereafter
Demographics	X	
Experience as Home Visiting Recipient	X	
Education, discipline, employment background	X	X
Psychological Well-Being		
Depressive Symptoms	X	X
Relationship Security	X	X
Burnout	X	X
For each outcome domain		
Perceived roles and responsibilities	X	X
Self-efficacy to fulfill roles in challenging situations	X	X
Knowledge and skills to carry out roles	X	X
Ratings of implementation system components	X	X
Responsiveness to training, supervision, feedback[3]	X	X

NOTES: [1]Measured through home visitor survey unless otherwise specified
[2] Or upon assumption of role for staff hired after site entry into the study
[3]Supervisor report of home visitor's responsiveness

Family Attributes: Baseline family attributes also influence service delivery and outcomes. A more detailed discussion of baseline family attributes is provided in Chapter 5. Table 4.4 lists some candidate constructs that are relevant to investigating how services are delivered. These include family risks and strengths (relationship security, depression, substance use, cognitive capacity, social support, and parenting beliefs) as well as the family's reasons for enrolling in home visiting and expectations of what enrollment entails and the benefits to be derived.

Design Options for Home Visiting Evaluation Project

Table 4.4

Baseline Data Collection to Measure Family Attributes[1]

Demographics
Education, discipline, employment background
Parenting Risks / Strengths
Depressive Symptoms
Substance Use
Cognitive Capacity
Social Support
Parenting beliefs / perceived norms
Reasons for enrolling in program
For each outcome domain
Perceived need for / value of program services
Perceived roles of home visitor and family in services

NOTES: [1]Family baseline attributes are conceptualized as factors for service delivery but many of these attributes are also malleable risk / protective factors that the program aims to impact.

Actual Services: Table 4.5 shows proposed methods for collecting information on program outputs. These include information on the family's responsiveness to home visiting (for example, the mother's relationship with the home visitor), the dosage of services provided (for example, the duration and frequency of services), the content of services and techniques used by home visitors, and the quality of home visiting as measured by the process of care (for example, interaction style). Information would be collected directly from the home visitor through management information systems, web-based questionnaires, and, because of concerns about the accuracy of home visitors' reports on their own activities, videotapes of home visitors interacting with the families during home visits. Videotapes will also allow the evaluator to examine aspects of home visits that are more difficult to measure through home visitors' self-report logs. Home visitors would be videotaped at least once with each family they visit. This will provide vital information on what actually takes place in the home, something that many prior evaluations of home visiting programs lack.

Design Options for Home Visiting Evaluation Project

Table 4.5

Measurement of Family Responsiveness and Services Actually Provided

	Program Records	Home Visitor Report	Maternal Report	Observation
Family Responsiveness				
Relationship with home visitor		X	X	
Engagement in activities		X	X	
Actual Services – Dosage				
Duration of enrollment	X			
Reasons for disenrollment	X	X	X	
For each Visit:				
Date; Length; Distribution of Time	X	X		
Participants	X	X		
Actual Services – Content, Techniques				
For each Visit:				
Content/activities specified in theories of change for each outcome domain	X	X		
Techniques used	X	X		
For selected Visits:				
Content/activities specified in theories of change for each outcome domain				X
Techniques used				X
Actual Services – Quality				
Interaction Style				X

Analytic Plan

The first descriptive analytic task is to characterize inputs: community context, stakeholder organizations, service model, implementation system, home visitors and supervisors, and families. This will address five specific questions:

- What is the community context in which home visiting programs operate?
- What organizations play a role in defining the service model and contributing to the implementation system?
- How are service models and implementation systems defined?
- What are the characteristics of home visitors and supervisors?
- What are the characteristics of families enrolling in home visiting?

This area of the implementation study will yield detailed quantitative and qualitative data on each input for each site. For many inputs, the data will be specific to each outcome domain. Thus, the analysis will not only describe overall site-level inputs but also describe and compare inputs for each outcome domain.

The analyses are primarily descriptive. The distribution of quantitative variables can be characterized in terms of basic descriptive statistics such as proportions and confidence intervals, means and standard deviations, and medians and inter-quartile ranges. For qualitative data, the first step will be to use qualitative software to extract, compile, and examine the occurrence of key information obtained across data sources. The evaluator will need to systematically review the data and generate specific indicators and a hierarchy of codes. The preceding analytical steps will involve the development of constructed variables that reduce and simplify quantitative and qualitative data collected from multiple sources over the course of the evaluation. These mixed categorical and interval variables will represent a range of implementation characteristics.

Beyond this, for each site, summary measures of model clarity and implementation system adequacy can be derived from the qualitative and quantitative data for each outcome domain. The summary measures for each site could be formatted in matrices to facilitate pattern identification across sites. These patterns, in turn, could be used to create a codebook to assign ordinal ratings such as poor, moderate, or good to each aspect of a program's definition of its model and each aspect of its implementation system pertinent to each outcome domain. Building on this, a codebook could be created to arrive at global ratings of model clarity and implementation system adequacy for achieving each outcome domain. Such a categorization is consistent with the emphasis in qualitative research on use of analytical categories to describe and explain social phenomena.

In descriptive analyses, the distribution of sites in terms of model clarity and implementation system adequacy could be characterized across each outcome domain.

The second descriptive analytic task is to characterize outputs to address the question:

- What services are actually provided?

As with inputs, this part of the implementation study will yield detailed quantitative and qualitative data on each output for each site. For many outputs, the data will relate to outputs relevant to specific outcome domains. Thus, analysis will describe not only overall site-level outputs but will also be able to describe and compare outputs pertinent to specific outcome domains. As with inputs, quantitative analyses would involve basic descriptive statistics such as proportions and confidence intervals, means and standard deviations, and medians and inter-quartile ranges. As with inputs, conventional approaches to qualitative analysis could be applied.

The third descriptive analytic task is to characterize outputs in relation to service model definitions of intended service dosage, content and quality to answer the question:

- How do actual services differ from intended services?

To assess outputs, it is important not only to measure services delivered, but also to compare these to what is expected. Service models will vary. One site might provide fewer visits or carry out specific activities less often than another site not because its services are poorer but because its service model calls for less intensive services. It would be misleading to compare sites on the basis of actual service delivery alone. Thus, the third analytic task is to compare actual service delivery to the service model to arrive at measures of fidelity.

Broad Research Question 2: How are inputs related to one another?

This aspect of the implementation study involves analyses to determine the associations among inputs. Across sites, the study seeks to address five questions.

- How is community context associated with a site's definition of its service model and the adequacy of its implementation system?
- How are the characteristics of stakeholder organizations related to a site's definition of its service model and the adequacy of its implementation system?
- How are features of the service model associated with features of the implementation system?
- How are a site's service model and implementation system related to the characteristics of its staff?
- How are a site's service model and implementation system related to the baseline characteristics of the families it enrolls in home visiting?

The knowledge gained in addressing these specific questions is important for policy and program decision-making to enhance the effectiveness of evidence-based home visiting models as they are taken to scale. It is important to examine the influence of the community context and stakeholder influence to identify possible reasons for observed variations across sites in the clarity and coherence of their service models and the adequacy of their implementation systems. This knowledge can be used to establish policies and procedures that promote sound adaptation of evidence-based models and the resources needed to achieve implementation fidelity. It is important to understand how features of the service model are associated with features of the implementation system because this can help explain why home visitors differ on characteristics believed to influence actual service delivery and impact. Understanding this will guide policy and practice to improve home visitors' capacity to fulfill their roles. How a site's service model and implementation system are related to the baseline characteristics of the families it enrolls is important to address because baseline family characteristics are likely to influence family engagement in home visiting. Furthermore, baseline characteristics are likely predictive of family outcomes without home visiting. If we learn which features of service models and implementation systems lead to engagement of families that are most likely to benefit from home visiting, this can improve the targeting of services.

Conceptualization of Constructs and Measurement Plan

The constructs were described in the section on Broad Research Question 1.

Analytic Plan

- How does community context influence a site's definition of its service model and the adequacy of its implementation system?
- How do stakeholder organizations influence a site's definition of its service model and the adequacy of its implementation system?

The first two sets of analyses are conducted only for the program group, so they will provide primarily descriptive evidence of how different aspects of program implementation are associated with one another under different conditions, rather than the rigorous causal analysis that is possible in the impact evaluation.

The first analytic task is to assess how community context and stakeholder organizations decide which evidence-based home visiting models to adopt and whether and how to adapt their selected models. These analyses will draw on both quantitative and qualitative indicators of community context and organizational influence. Consider, for example, the outcome of reducing maternal depression. A given site might or might not explicitly specify this as one of its intended outcomes. It should be possible to test several community and organizational features as predictors of whether reducing maternal depression is adopted as an explicit outcome. These include the availability of community-level statistics on the prevalence of maternal depression, the actual prevalence, influential organizations' perceptions of the importance of addressing this outcome through home visiting, and both the actual and perceived availability of community resources to reinforce and coordinate with home visiting services in addressing the outcome.

- How are features of the service model and implementation system related to one another?

The second analytic task is to estimate the strength of associations among features within the service model, within the implementation system, and between the service model and implementation system. For many features, strong correlations would be expected. For example, whether a program explicitly aims to promote maternal mental health (service model feature) should be linked with whether its implementation system includes staff training in identifying and responding to maternal depression, the availability of tools to screen or assess for depressive symptoms, and access to clinical and administrative supports for responding to identified mental health problems. Other features of the service model and implementation system might not be related. For example, the service model's definition of eligible families is unlikely to be related to the adequacy of staff training or the use of management information systems to monitor and promote service quality.

Exploring these relationships would primarily involve uncontrolled bivariate tests of association. Building on this, the analysis could proceed to explore patterns of association among features and the identification of profiles of features for specific outcomes. Analyses could determine whether and how specific combinations of features co-occur and could describe the prevalence of specific combinations across sites.

- How do service models and implementation systems influence staff characteristics?
- How do service models and implementation systems influence the baseline characteristics of the families who enroll in home visiting?

The third analytic task is to determine the associations of organizational-level factors with individual-level factors for home visiting implementation and impact. Exploring these relationships would require several steps. To illustrate, consider analyses to assess how service model and implementation system features influence home visitor self-efficacy in addressing maternal depression. First, using uncontrolled bivariate analyses, one would estimate the associations of home visitor self-efficacy with individual features of the service model (for example, whether promoting maternal mental health is an explicit program outcome) and the implementation system (for example, whether home visitors received training in how to screen for depression, whether specific instruments are used, whether the home visitor has access to staff with expertise in this area, and whether the program facilitates home visitors in referring families to treatment services). Building on this, the analysis could proceed to build regression models incorporating multiple service model and implementation system features to test their independent and combined influence on home visitor self-efficacy. The results of analyses carried out to estimate the associations between service model and implementation system features (the second broad implementation study research question) would inform model building.

A parallel set of analyses could be carried out using family rather than home visitor characteristics as the dependent variable.

Broad Research Question 3: How are inputs related to outputs?

Specific Questions

This aspect of the implementation study involves longitudinal, hierarchical analyses of data for families assigned to the program group. Families assigned to the control group would be excluded. Across sites, this portion of the study seeks to address four questions:

- How do service model and implementation system features influence the dosage, content and quality of actual services?
- How do staff and family characteristics influence actual service delivery?
- How do staff and family characteristics interact as influences on actual service delivery?
- How do staff and family characteristics mediate the influence of the service model and implementation system on actual service delivery?

The answers to these questions would provide crucial information for home visiting policy and practice by describing how organizational and individual-level factors for service delivery can be shaped to assure high quality services that are faithful to evidence-based models. Addressing the first question will help explain how the characteristics of service models and implementation systems are related to service delivery in ways that are likely to influence program impact. Addressing the second question will identify home visitor characteristics that should be considered in staff recruitment or that could be modified through training and supervision to promote fidelity to evidence-based home visiting models. It will also identify family baseline attributes that should be considered in targeting families and in introducing them to program services in ways that promote their engagement in services. Addressing the third

question will determine which home visitors and families are likely to be more successful working together. This information can be used to shape decisions around program targeting, and staff recruitment, training, and supervision.

Conceptualization of Constructs and Measurement Plan

The constructs were described in the section on Broad Research Question 1.

Analytic Plan

- How do service model and implementation system features influence the dosage, content, and quality of actual services?

Exploring the relationship between program features and actual service delivery would proceed in steps. To illustrate, the first step could be to estimate the associations of features of the service model and the implementation system with actual service delivery, first in uncontrolled bivariate analyses and then controlling for the associations of the service model and implementation system with family characteristics. Regression models could be developed for service model features and then for implementation system features. These models could then be expanded by including both service model and implementation system features. This model building will be informed by results of analyses carried out to test associations between service model and implementation system features (the second broad implementation study research question).

A parallel set of analyses could be carried out using fidelity rather than actual service delivery as the dependent variable. For example, instead of using the number of visits to measure dosage, one would use the actual number of visits as a proportion of the number of expected visits. This approach could be applied to service model definitions as put forth by national model developers and also as adapted by sites.

- How do staff and family characteristics influence actual service delivery?
- How do staff and family characteristics interact as influences on actual service delivery?
- How do staff and family characteristics mediate the influence of the service model and implementation system on actual service delivery?

Exploring the relationship between individual-level characteristics and actual service delivery would also proceed in steps. For example, the first step could be to estimate the independent associations of home visitor and family characteristics with service delivery. The next step would be to test the interactive effects of home visitor and family characteristics with service delivery. For example, consider program services to promote positive mother-child interaction. One could test first whether visit content to move this outcome is different for home visitors who are themselves depressed as compared to those who are not depressed, adjusting for whether the mother is depressed. The next step would be to test the interactive effects of home visitor and maternal characteristics. For example, the model could be expanded to test whether the effects of home visitor and maternal depression on service content to promote positive parenting is different if both are depressed. That is, one could determine if service content is

diminished if only the home visitor is depressed, only the mother is depressed, or either is depressed, and if the effects when both are depressed are additive or multiplicative.

- How do staff and family characteristics mediate the influence of the service model and implementation system on actual service delivery?

The first set of analyses in this implementation substudy will have estimated the influence of organizational level factors on actual service delivery. The second set of analyses will have estimated the influence of individual-level factors. This specific research question is addressed by the third set of analyses, which will estimate the extent to which the influence of individual-level factors explains the observed influence of organizational level factors on service delivery. This can be achieved by expanding the models developed in the first set of analyses by introducing home visitor and family characteristics as covariates and testing for changes in the estimates of service model and implementation system features as factors for impact.

Chapter 5

Impact Analysis and Analysis of Health Systems Outcomes

As discussed in Chapter 1, the Affordable Care Act (ACA) requires the national evaluation of home visiting programs to assess the effectiveness of home visiting to improve outcomes for families and children, both overall, and for subgroups of families. Because of the potential of home visiting programs to affect health care use and health outcomes, the ACA also calls for the national evaluation to assess the potential for home visiting to reduce health disparities and improve health care quality and practices.

Here are the research questions addressed by this area of the evaluation:

- What are the effects of home visiting programs across the domains of outcomes mentioned in the ACA?
- Do the effects of home visiting programs vary across subgroups of families?
- What is the relationship between the features of home visiting programs and their effects on family outcomes?
- What are the effects of home visiting program on health disparities, health care quality, and health care practices?

This chapter describes these analyses, along with a plan for collecting information to inform the impact analyses.

Measurement Plan for the Impact Analysis

The impact analysis will use baseline information collected through surveys of parents and data collected by the programs at the time of recruitment. It will also include outcomes from a range of data sources collected either six months or a year after parents enter the study (depending on what follow-up data collection option is chosen). These include surveys of parents, administrative data on child abuse and neglect, observations of parents interacting with their children, and, for older children, direct assessments of child development.

Overview of data collection

The legislation indicates that the evaluation should assess the effects of home visiting on a number of domains, including prenatal, maternal, and newborn health; child health and development; parenting; school readiness and academic achievement; crime and domestic violence; family economic self-sufficiency; and referrals and service coordination.

As described in Chapter 2, two follow-up periods are being considered: six months and twelve months. For each of these follow-up periods, outcomes would be collected separately for an older and a younger cohort of children. For a six-month follow-up period, the younger cohort would include pregnant mothers and families that enter the study before the child is three months of age, and the older cohort would include families that enroll when their children are between two and four years old. For a twelve-month follow-up period, the younger cohort would include

those who enter the study before the child is six months old, and the older cohort would be the same as in the first design option.

For both possible follow-up periods, the first wave of data collection will occur at baseline, or when study participants enter the evaluation. Baseline data collection provides information needed to describe the population of families and children targeted by home visiting programs. Baseline data are also used to identify subgroups of the population for the impact analysis and to provide covariates to improve the precision of estimated effects. Some components of the baseline data may vary depending on whether the family is enrolled prenatally or postnatally.

The second wave of data collection will be used to assess program impacts on parents and children. For a six-month follow-up period, the timing of the second wave of data collection would depend on when families enter the study. For families enrolled prenatally, the second wave would occur when children are six months of age. For families of older children, the second wave would occur six months after random assignment. For a twelve-month follow-up period, the second wave of data collection would occur for all children in the younger cohort when children are 12 months of age. For the older age group, the second wave of data collection would occur one year after enrollment, when children are three to five years of age.

There are potential advantages and disadvantages for each possible follow-up period. A six month follow-up would reduce program attrition, would provide initial impact estimates six months earlier, and could be important if a primary goal of the study is to understand impacts on parenting or child development in early infancy. For example, if a primary goal were to understand home visiting impacts on maternal depression at about six months post-partum, rather than at 12 months post-partum, the shorter follow-up period would be more desirable. Finally, to the extent that the study relies on maternal report of birth outcomes and infant health (rather than administrative data), maternal accuracy of reporting will arguably be more reliable the closer in time mothers are to the birth of the child.

The potential advantages of a one-year follow-up period relate to study enrollment, dosage, and measurement. This longer follow-up period features an enrollment window for infants up to six months of age, rather than just three months of age, which would increase the speed with which the study can reach full enrollment. In addition, it would allow for a longer dose of intervention for most children, which is especially relevant since most evidence-based models are intended to provide home visits for more than a year. Another advantage to the twelve-month follow-up is that there are more high-quality options for measuring child development and parenting when children are one year of age. These include increased options for parent-child interaction measures, parent self-report measures on child development, and – if a follow-up child development assessment for younger children, such as the one presented in Appendix B, were to be included = – assessment of cognitive, language, and social-emotional and behavioral development.

Whichever follow-up data collection strategy is chosen, it will include assessment of the domains described in the legislation as benchmarks and outcomes for families and children: a) maternal and newborn health; b) prevention of child injuries, child abuse, neglect, or

maltreatment, and reduction of Emergency Department visits; c) improvement in early child development, school readiness, and achievement; d) reductions in crime and domestic violence; e) improved parenting; f) improved family self-sufficiency; and g) greater coordination of referrals to community resources. Although for many outcome domains, the impact measures would be identical for the two different age groups, we note that the precursors to school readiness and school readiness battery of measures would vary for these two different age cohorts.

The Process of Identifying Constructs for the Impact Study

Crosswalk between the legislative benchmark areas and the evaluation outcome domains. The impact study for the national evaluation needs to address all of the domains described in the legislative benchmarks domains and participant outcomes. For purposes of conceptualizing the measurement work, the benchmark domains and participant outcomes were grouped into five distinct domains: parent health and well-being, parenting capacity, parenting behavior, child health and development, and actual services (see Table 5.1). As Table 5.1 shows, in some cases, a given domain or outcome may include specific outcomes that fall into more than one domain in this framework. For example, for the benchmark area "prenatal, maternal and newborn health," one outcome domain includes prenatal and newborn health (child health and development) but a different domain includes maternal health (parent health and well-being). In the rest of the chapter, constructs and measures are organized into these five domains.

Design Options for Home Visiting Evaluation Project

Table 5.1

Benchmark Outcomes and Organization in Conceptual Framework

Benchmarks/Participant Outcomes	Location in Framework
1. Prenatal, maternal and newborn health	Parent Health and Well-Being; Child Health and Development
2. Child health and development (including injuries, hospitalizations, ED visits, maltreatment)	Child Health and Development
3. Parenting skills	Parenting Capacity; Parenting Behavior
4. School readiness and academic achievement	Child Health and Development
5. Crime and domestic violence	Parent Health and Well-Being
6. Family economic self-sufficiency	Parent Health and Well-Being
7. Referrals and service coordination	Actual Services

Identifying key constructs within each domain. After identifying the critical domains, the next step in the process was to identify the key baseline and outcome constructs within each domain that should be measured to gauge the direct effects of the home visiting programs, as well as constructs that may be key moderators or mediators of impacts. Important constructs were identified by drawing from the following resources: 1) the conceptual models and theories of change underlying the ways in which evidence-based home visiting models are hypothesized to affect maternal and child well-being and development; 2) the results of prior evaluations of these programs, particularly the constructs and measures that were included in impact analyses; and 3) early input from the HHS staff and other stakeholders about high priority constructs to measure in light of the goals of the initiative and the range of evidence-based programs to be included in the evaluation.

Baseline Constructs for the Impact Study

As with any large, random assignment study, this evaluation should include measurement of key baseline constructs. These constructs serve two purposes in the impact analysis: they serve as covariates to increase the statistical precision of impact estimates, and they are used to identify subgroups. In addition, as described in Chapter 4, they will provide information to help the implementation study link program services to particular family needs. Table 5.2 presents a list of possible baseline and follow-up constructs.

Baseline parent health and well-being constructs. Baseline covariates should include demographic information such as maternal and paternal age, race and ethnicity, number of other living children, the respondent's dominant language, and a measure of acculturation. Other useful demographic information would include other household characteristics such as age and relations of other household members, whether the respondent was homeless, a measure of housing mobility, and the dominant language spoken in the household. The baseline survey may also include measures of family self-sufficiency, such as maternal and paternal earned income, total household income, maternal and paternal employment, maternal and paternal highest grade completed and current schooling, maternal educational aspirations, current public assistance receipt, and maternal monetary assistance from the father (material support and amount contributed to household income). All of these are key variables for understanding the characteristics of the population served by home visiting programs and to identify subgroups of interest.

The baseline interview should also assess key aspects of the mother's health, including physical health (global health, illness, nutrition, and quality of life), mental health (depression, other mental illness, and social support), and substance use (tobacco use, substance use, and problem alcohol use). Respondents would also be asked to report on the physical health and substance use characteristics of the father of the child. Baseline data should also include an assessment of the mother's desired timing of future subsequent births. Finally, measures of maternal and paternal prior arrests and convictions should be collected at baseline. Again, these maternal and paternal risk factors may be used to identify important subgroups of interest.

Baseline child health and development constructs for children born prior to random assignment. Birth outcomes for children who have already been born at the time of random assignment could include the following: birth weight and length, gestational age, special health

care needs at birth, and length of stay in the hospital. These newborn health indicators are associated with long-term health and development and therefore serve as key baseline covariates and variables to identify at-risk subgroups of children. Baseline data should also include measures of children's current health status, height and weight, special health care needs, any prior substantiated or unsubstantiated claims of abuse and neglect, prior injuries, and the child's temperament in infancy.

    Baseline parenting constructs for families with children born prior to random assignment. For those families that enroll postnatally, various parenting constructs could be measured at baseline. This could include parenting behaviors such as cognitive stimulation, social-emotional responsivity, harsh parenting, discipline strategies, breastfeeding and nutrition, sleep routines, prior focal child maltreatment, safety of the home environment, parenting knowledge and attitudes, attachment style and aspects of the mother-father relationship (establishment of paternity, father involvement, and parents' relationship quality).

Outcome Constructs for the Impact Study

    Child Health and Development. Outcomes related to child health and development can be further classified as related to birth outcomes, postnatal health outcomes, and school readiness.

    Birth outcomes for those families enrolled prenatally are key outcomes of interest for the evaluation. Prior research suggests that birth weight, gestational age, size for gestational age, and health status at birth are linked to children's short- and long-term health and development, as well as family well-being and health system costs. A low birth weight infant can be born too small, too early, or both. Compared to infants of normal weight, low birth weight infants may have an increased risk for many negative outcomes. The negative outcomes can be immediate, such as infection or perinatal morbidity (illness through the first week of life) or mortality. Longer-term consequences of impaired development can be delayed motor and social development or learning disabilities. These constructs - birth weight, gestational age, size for gestational age, and health status at birth - would be critical to measure as important pregnancy-related program outcomes ultimately linked to other aspects of children's development, as well as for the cost-effectiveness analysis.

    Children's physical health and development are key outcome domains for most home visiting programs. Physical health problems can be harmful to other domains of a child's development, as well as financial stressors on families and broader health systems. Because of this, the design includes measures of children's height and weight, physical delays and motor development, special health care needs, child abuse and neglect, injuries, and whether the child has been determined to be eligible for early intervention services.

    Children who enter kindergarten behind in pre-academic and social-emotional and behavioral development are at increased risk for struggling in school. Thus, improving at-risk children's school readiness remains a national priority and is increasingly an important element of many home visiting programs. Accordingly, the national evaluation will estimate the effects of programs on school readiness of the older cohort of children, and precursors to school readiness for the younger cohort of children. Prior studies suggest that the evaluation include children's outcomes in the following domains: cognitive, language, social-emotional and

behavioral development, and approaches to learning and executive function skills. For the younger sample of children, who will be about 12 months old at follow-up, the evaluation includes parent report assessments of children's cognitive, language, and social-emotional, and behavioral development. For the older sample of children, who will range in age from three to five years, the evaluation includes a battery of parent report as well as direct assessments of children's cognitive, language, social-emotional and behavioral development, as well as children's approaches to learning and executive function.

Parenting Capacity. Sufficient parenting capacity is a crucial precursor to improvements in the quality of parental responsivity and home environment. Parental motivation, knowledge, reflective capacity, and resources to engage in the home visiting program, to improve parenting skills, and to improve the quality of the home environment are critical precursors to change in parenting behaviors and decisions involving children, and ultimately, child development. That is, unless parents are motivated and have the knowledge and resources to attend to those factors that influence child development, they will be less likely to actually do the things that matter for supporting their child's development. Thus, the study should assess these aspects of parenting capacity as part of the impact evaluation.

Parenting Behavior. Researchers have identified many different ways to describe parenting behaviors that are important to young children's development. The measurement plan classifies parenting behaviors that influence development into two broad types: social-emotional responsivity and cognitive responsivity, and recommends that both are measured in the evaluation as key parenting outcomes. Social-emotional responsivity refers to the parent's ability to quickly, appropriately, and sensitively read and respond to an infant or young child's needs and cues and to provide a secure source of attachment, supportiveness, and warmth. This type of responsivity is crucial in infancy for healthy development, but remains important through preschool and beyond. Cognitive responsivity refers to a parent's quantity and quality of verbal and cognitive attention, stimulation, and interaction with his or her child. It has been linked to cognitive development in infancy and early childhood. The evaluation should also include a measure of harsh parenting, as it has been linked to poor child outcomes. In addition, although it is sometimes difficult to capture treatment effects on substantiated or unsubstantiated claims of child abuse or neglect due to low incidence, prior evaluations have found treatment effects on harsh parenting. The evaluation should also include a measure of the discipline strategies the parent commonly uses.

Many home visiting programs promote health-related parenting practices, such as nutrition and healthy sleep habits. Both of these have been linked to children's short and long-term health and physical development, and even to school readiness. Therefore, the evaluation should measure nutrition practices, such as breast-feeding and other child nutrition practices, as well as sleep habits including sleeping routines and sleeping arrangements.

Reduction of child maltreatment is a major targeted outcome of home visiting programs. The evaluation should therefore measure substantiated and unsubstantiated reports of neglect and abuse. In addition, the study should measure whether the mother or father of the child has relinquished his or her role as parent.

Finally, the evaluation should include several measures of the quality of the home environment. Two aspects of the home environment strongly linked to children's development are the quality of the home learning environment and the physical safety of the home. The quality of the home learning environment is an important target of many home visiting programs and has also been linked to children's long-term outcomes, particularly cognitive outcomes. The physical safety of the home is also a primary target for many home visiting programs, and has been linked to children's health outcomes, including risk for injury and long-term health outcomes, while other safety hazards (for example, lead paint) are also associated with long-term cognitive and behavioral outcomes.

Parent Health and Well-Being. Improving parent health and well-being is also a goal of many home visiting programs and, therefore, a significant focus of the evaluation. These outcomes can be classified as related to maternal or paternal health, domestic violence and crime, and family self-sufficiency.

Maternal health is a key outcome domain of interest. For mothers enrolled prenatally, it may be appropriate to assess key pregnancy-related health constructs, such as prenatal health problems (for example, gestational diabetes and high blood pressure), pregnancy weight gain, and pregnancy-related nutritional practices. It will also be critical to measure actual services received (discussed below), including services, coordination, and referrals during and immediately after pregnancy for this subgroup.

Maternal mental health and substance use are other key constructs of interest and serve as major risk factors for reduced family well-being and child development outcomes. Parents with mental and behavioral health problems are at increased risk for poor parenting quality, committing child maltreatment, and providing lower quality home environments. Although major depressive disorder is particularly harmful for child outcomes, even mild depression or elevated depressive symptoms have been linked to poorer parenting quality. Other major mental health issues, including bipolar disorder, anxiety, and schizophrenia, are also associated with marked reductions in the quality of parenting and the home environment. We note that maternal substance abuse and mental health conditions are often co-morbid, further increasing the risk for diminished child outcomes. Thus, the evaluation would measure maternal depression, anxiety, presence of other mental illness, and social support, as well as maternal substance use, including tobacco, other substances, and problem alcohol use.

Prior studies of home visiting programs have found treatment impacts on maternal reproductive health, with reductions in the number of subsequent pregnancies as well as the distance between subsequent pregnancies, with long-term implications for family self-sufficiency and health care system expenditures. Therefore the evaluation should measure the desired and actual timing of subsequent pregnancies in the follow-up study.

Although mothers are usually the enrolled adults in home visiting programs, the physical and mental health of the father is a key aspect of healthy family well-being. Therefore, the evaluation should also measure father's overall health and any major illnesses, as well as substance use, tobacco use, and problem alcohol use at follow-up.

Domestic violence and the risk for domestic violence are primary targets of many home visiting programs. For this reason, the evaluation should measure parental conflict, battery, and restraining orders at follow-up. The evaluation should also measure maternal and paternal arrests and convictions since baseline (when prior arrests and convictions are to be measured).

Key outcome constructs in the domain of family self-sufficiency include maternal and paternal income, total household income for the household in which the child resides, maternal and paternal employment, maternal and paternal highest grade completed since baseline, maternal and paternal current schooling, maternal educational aspirations, current public assistance receipt, and maternal monetary assistance from the father (material support and amount contributed to household income). The evaluation should also measure constructs related to housing, including household composition, homeless status, and housing mobility since baseline. All of these are key constructs for understanding program impacts on family self-sufficiency.

Actual Services: Referral and Coordination with Other Services. Many home visiting programs aim to improve access, efficiency, and quality of services within the family and child health and welfare system. Measuring the adequacy of screenings, referrals, and receipt of referred services as a result of the home visiting service will be critical for the evaluation.

The evaluation should assess impacts on screenings, referrals, service coordination, and service usage of various medical and non-medical services for children enrolled in the study. These include medical insurance status, usual source of care, well-child visits, immunizations, developmental screenings, subspecialist health care, hospitalizations, injury care, visits to the Emergency Department, other outpatient services, and use of prescription drugs. Finally, we suggest measures assessing whether the child has been screened, received referrals from the home visiting services, and received help with accessing and coordinating supplemental nutrition programs, child care services, and early education or preschool services.

The evaluation should also assess impacts on screenings, referrals, service coordination, and service usage for mothers enrolled in the study. For those mothers enrolled prenatally, mothers' screenings, referrals, and coordination and use of standard prenatal and postpartum care services should be assessed. For all mothers, the evaluation should assess the following: health insurance; a regular source of care; primary, reproductive, mental, and substance use services; hospitalizations; injuries requiring health care; visits to the Emergency Department; domestic violence services; education and workforce services; and public benefits.

Design Options for Home Visiting Evaluation Project

Table 5.2

Key Baseline and Outcome Domains: Proposed Timing and Constructs

	Prenatal/Infant Sample		Toddler/Preschool Sample	
	Baseline	6 mos/1 year FUP	Baseline	6 mos/1 year FUP
CHILD HEALTH AND DEVELOPMENT				
Health				
Newborn health [a]				
Birth weight	X		X	
Gestational age	X		X	
Size for gestational age	X		X	
Special health care needs	X		X	
Length of hospital stay	X		X	
Infant and Child Health and Physical Development				
Overall health	X	X	X	X
Height/weight[b]	X	X	X	X
Physical delays/motor development		X	X	X
Special health care needs		X	X	X
Child abuse or neglect		X		X
Injuries		X		X
Infant temperament[c]	X		X	
Child is receiving EI services		X		X
School Readiness and Precursors to School Readiness				
Cognitive Development		X		X
Language Development		X		X
Social-Emotional Development and Behavior		X		X
Approaches to Learning and Executive Function				X

(continued)

NOTES: [a]Newborn health will be collected as an impact measure for families enrolled prenatally. For all other children it will be collected as a baseline covariate.

[b]Height and weight and selected other health measures will measured at baseline for families enrolled postnatally.

[c]Infant temperament will be measured at FUP for families enrolled prenatally and baseline for families enrolled postnatally.

67

Table 5.2 (continued)

	Prenatal/Infant Sample		Toddler/Preschool Sample	
	Baseline	**6 mos/1 year FUP**	**Baseline**	**6 mos/1 year FUP**
PARENTING[d]				
Parenting Behavior				
Cognitive stimulation		X	X	X
Social-emotional responsivity		X	X	X
Harsh parenting		X	X	X
Discipline		X	X	X
Breastfeeding, nutrition		X	X	X
Sleep routines, arrangement		X	X	X
Child Maltreatment				
Neglect		X	X	X
Substantiated reports		X	X	X
Unsubstantiated reports		X	X	X
Relinquishment of role		X	X	X
Home Environment				
Support for learning		X	X	X
Home safety environment		X	X	X
Parenting Knowledge		X	X	X
Parenting Attitudes/Beliefs		X	X	X
Parenting Stress		X	X	X
Attachment Style	X		X	
Parents' Relationship and Father Involvement				
Paternity established	X		X	
Father involvement		X	X	X
Parents' relationship quality	X	X	X	X
PARENT HEALTH AND WELL-BEING				
Maternal health (including prenatal)				
Physical Health				
Global Health	X	X	X	X
Illness	X	X	X	X
Nutrition	X	X	X	X
Quality of Life	X	X	X	X
Mental Health				
Depression	X	X	X	X
Other mental illness	X	X	X	X
Social support	X	X	X	X
Substance Use				
Tobacco use	X	X	X	X
Substance use	X	X	X	X
Problem alcohol use	X	X	X	X

(continued)

NOTES (continued): [d]Selected parenting measures may be collected at baseline.

Table 5.2 (continued)

	Prenatal/Infant Sample		Toddler/Preschool Sample	
	Baseline	6 mos/1 year FUP	Baseline	6 mos/1 year FUP
PARENT HEALTH AND WELL-BEING				
Maternal reproductive health				
Subsequent pregnancies		X		X
Subsequent births (date of birth, gestational age, birthweight, overall health)		X		X
Desired timing of subsequent births	X	X	X	X
Paternal health				
Physical Health				
Global health	X	X	X	X
Illness	X	X	X	X
Substance Use				
Tobacco use	X	X	X	X
Substance use	X	X	X	X
Problem alcohol use	X	X	X	X
Intimate Partner Violence				
Conflict		X		X
Battery		X		X
Restraining orders		X		X
Crime				
Maternal crime				
Prior arrests, convictions	X		X	
Sub. arrests, convictions		X		X
Paternal crime				
Prior arrests, convictions	X		X	
Sub. arrests, convictions		X		X
Family Self-Sufficiency				
Income				
Maternal earned income	X	X	X	X
Paternal earned income	X	X	X	X
Household income	X	X	X	X
Employment				
Maternal employment	X	X	X	X
Paternal employment	X	X	X	X
Education				
Maternal highest grade completed	X	X	X	X
Maternal current schooling	X	X	X	X
Maternal educational aspirations	X	X	X	X
Paternal highest grade completed	X			
Paternal current schooling	X	X	X	X

(continued)

Table 5.2 (continued)

	Prenatal/Infant Sample		Toddler/Preschool Sample	
	Baseline	6 mos/1 year FUP	Baseline	6 mos/1 year FUP
PARENT HEALTH AND WELL-BEING				
Current Public Assistance (TANF, SNAP, WIC, Health, UI)	X	X	X	X
Assistance from father				
Material support	X	X	X	X
Income	X	X	X	X
Housing and household composition				
Age and relations of other members	X	X	X	X
Homeless status	X	X	X	X
Mobility	X	X	X	X
Demographics of Index Child's Parents				
Parents' ages	X		X	
Parents' race /ethnicity	X		X	
Parents' relationship	X	X	X	X
Parents' other living children	X	X	X	X
Language spoken at home	X	X	X	X
Acculturation	X	X	X	X

(continued)

Table 5.2 (continued)

	Prenatal/Infant Sample		Toddler/Preschool Sample	
	Baseline	6 mos/1 year FUP	Baseline	6 mos/1 year FUP
ACTUAL SERVICES				
Child-related services: screenings, referral, coordination, and use				
Insurance	X	X	X	X
Preventive/Primary Care[e]	X	X	X	X
Usual source of care	X	X	X	X
Immunizations	X	X	X	X
Developmental screening	X	X	X	X
Early intervention services	X	X	X	X
Receipt of well child care	X	X	X	X
Subspecialist health care	X	X	X	X
Hospitalizations	X	X	X	X
Injuries requiring health care	X	X	X	X
ED Visits	X	X	X	X
Outpatient service use	X	X	X	X
Rx drug use	X	X	X	X
SNAP	X	X	X	X
Child care	X	X	X	X
Mother-related services: screenings, referral, coordination, and use				
Insurance coverage	X	X	X	X
Use of services				
Prenatal/Postpartum care	X	X	X	X
Preventive/Primary Care	X	X	X	X
Usual source of care	X	X	X	X
Receipt of primary care	X	X	X	X
Reproductive health care	X	X	X	X
Mental/Substance use care	X	X	X	X
Hospitalizations	X	X	X	X
Injuries requiring health care	X	X	X	X
ED Visits	X	X	X	X
IPV services	X	X	X	X

NOTES (continued): [e]Measured at baseline for families enrolled postnatally only.

Impact Analysis

As described earlier in the report, the impact analysis will assess the effectiveness of home visiting programs in improving outcomes of families and children, both overall and across key subgroups of families and groups of parents. In addition, the impact analysis and implementation research will be linked to explore the program features that are associated with larger impacts. This section describes some principles for conducting the impact analysis.

Intent-to-Treat Impact Estimates

The proposed starting point for the impact analysis is to estimate intent-to-treat effects in which all program group members – regardless of whether they actually received home visiting services – are compared to all control group members, some of whom may have received home visiting outside of MIECHV. Random assignment ensures that these estimates are unbiased effects of allowing program group families to be eligible for home visiting services.

Intent-to-treat impact estimates could be calculated for a number of comparisons. First, there may be some outcomes that can be pooled across all sites to get most precise estimates of effects across the range of domains. Examples might include the degree of economic self-sufficiency, evidence of child abuse and neglect, and whether the children are receiving appropriate preventive health care.

With the sampling plan described in Chapter 3, most outcomes would be analyzed separately for programs serving pregnant women and mothers of infants on the one hand, and programs serving mothers with toddlers and preschool aged children on the other. Intent-to-treat impacts would be calculated for each group of children. Because the outcomes are expected to be quite different, the analysis would not directly compare the magnitude of the two sets of estimates. In addition, impacts on birth outcomes such as low birth weight would be estimated using only families that were enrolled prenatally.

Impact estimates would be regression adjusted, controlling for baseline characteristics of families and home visitors. In notation, regression-adjusted impacts would be calculated according to equation (1).[32]

$$(1)$$

In equation (1), y_{ij} indicates an outcome for family i in site j, E_{ij} is an indicator of whether the family was assigned to the program (home visiting) group or the control group, and X_{ij} are baseline characteristics of the family. A separate intercept would be estimated for each site and is represented by α_j, while the program effect is captured by the parameter β. The term ε_{ij} captures all parts of the outcome that are not explained by the baseline characteristics, the program group assignment, or the site the family comes from. If random assignment of families is done at the home visitor level, as discussed in Chapter 2, equation (1) would also include measures of baseline home visitor characteristics. Regression adjustment is intended to increase the precision

---

[32] Equation (1) assumes a linear regression, but the approach also can be used for other methods, such as logistic regression for binary outcomes or Poisson regression for count data.

of estimated impacts by reducing the unexplained variation in outcomes across families. Covariates would consequently be chosen because they are expected to be correlated with key outcomes, for example, maternal depression and maternal age at the time of enrollment may be predictive of a range of subsequent maternal and child outcomes.

Intent-to-treat estimates would also be calculated for key subgroups of families to address the question of whether home visiting programs have larger effects for some groups of families. The ACA places priority on serving a number of subgroups of families through MIECHV funding. These include pregnant women under 21 years old, those with a history of child abuse and neglect, parents with a history of substance abuse, and families with a child with development delay or disabilities. Other subgroups of policy interest include families enrolled before the child is born, first-time mothers, and families with depressed mothers.

The main question for these subgroup calculations is whether impacts differ across subgroups. For example, in estimating the effects for mothers who were pregnant when they entered the study to those whose children were infants, the impact analysis would investigate whether estimated effects were larger for one group than for another. If there are not statistically significant differences across subgroups and the pooled effects are significantly different from zero, the presumption would be that home visiting is effective for all subgroups. This approach is recommended because estimated effects for subgroups are less precise than estimated effects for the full sample. Consequently, it is likely that estimated effects for some subgroups would not be statistically significant even if the program was modestly effective for that subgroup.

For drawing conclusions about the effectiveness of home visiting by subgroup, the evaluator would specify which subgroups would be examined before the analysis begins. Subgroups would be chosen based on prior evidence of differential effects across subgroups, theory that suggests effects should be larger for one group than another, or policy interest in understanding the effects across subgroups. The evaluation would conduct such analyses across a limited number of subgroups to reduce the possibility that a chance result leads to a conclusion that impacts are different for one particular subgroup.

Impacts would also be estimated for groups of programs. To learn as much as possible about the specific features of programs that affect the direct experiences of families, this analysis would be designed to highlight individual measureable features that are hypothesized to affect program impacts (either directly or through the content, dosage, or quality of home visits). Examples of groups are maturity of the program, whether the program is highly networked with other community programs, and the clarity and complexity of program goals.

The evaluation should monitor participation rates in home visiting services by site and for both the program and control groups. In most random assignment studies, some families assigned to the program group receive no program services. Often this is because families volunteer for services, consent to be in the study and are randomized but later decide they do not want to receive services. Families also move and cannot be located by the programs. In addition, in an evaluation of home visiting, it is likely that some control group members will receive similar home visiting services. The information gained by monitoring participation rates will be

used in a second set of analyses to adjust impacts estimates for differences in the proportion of program and control group members who receive home visiting services.[33]

Exploring the Relationship between Program Features and Impacts

The legislation requires examining how impacts vary across programs. An in-depth interpretation of this requirement would also meet HHS's goal of informing policy, programmatic, and implementation decision-making through examining how the features of communities, service models, implementation systems, and home visitors are associated with program impacts. The next stage of the impact analysis would explore how various inputs into home visiting programs are related to impacts of those programs. Because sites would not be randomized to have different program features, a finding that sites with certain program features had larger effects would not necessarily mean that those features are responsible for the larger effects. Instead, those program features might be related to aspects of the program that were not measured or not included in the analysis. Unbiased estimates generated through random assignment of the effects of home visiting at each site would be linked to program features of that site, but the associations uncovered through the analysis might not be causal.

The idea behind this analysis is expressed in notation in equation (2):

$$(2)$$

In equation (2), impacts for home visitor k in site j are related to home visitor characteristics, as represented by $X_{jk,}$ and site characteristics, as represented by Z_j. Site characteristics could include any of the implementation factors described in Chapter 4, including features of the service models, implementation systems, and community context. Because sites would not be assigned to have different features of their home visiting models or communities, results of this analysis would be less rigorous than the intent to treat analysis. In addition, to the extent that the analysis does not include important measures of program implementation, the results might suggest the importance of one aspect of program implementation that is really representing an unmeasured aspect of implementation. Thus, results would be suggestive of the features that lead to larger effects but would not provide definitive evidence that certain program features cause larger effects.

To explore the relationship between program features and program impacts the evaluator would likely proceed in steps. For example, the first step could be to estimate the relationship between impacts and features of service models, controlling for the relationship between program features and family characteristics. Features of service models include the frequency of planned visits (for example, weekly or biweekly), whether the program directly targets maternal and child health or economic outcomes (for analyses that include those outcomes), and so on. This step is likely to provide the most rigorous causal conclusions because the features being examined are typically defined by the program model before the sample family has entered the study, similar to the family's baseline characteristics.

---

[33]Bloom (1984), Gennetian et al. (2005); Angrist, Imbens, and Rubin (1996); Duncan, Ludwig, and Hirschfield (2001); Peck (2003).

The regression model could then be expanded by adding in features of the implementation system, such as ratings of the training used for home visitors, the quality of the supervision of home visitors, what supports are available for facilitating program administration, and so on. These features are likely to be somewhat independent of features of the service model, so that both groups of features could easily be included in one regression. The results of this step would need to be interpreted somewhat more cautiously because these features of the system can theoretically be influenced by characteristics of the home visitors and families in the study site and by their responses to the program as it is implemented.

The regression model could then be expanded by adding information about what actually happens in the home. Although this set of program features is likely to be most closely linked to program effectiveness, it is mentioned last for several reasons. First, what a home visitor does for a specific family will depend on that family's needs. For that reason, estimates of the relationship between what happens in the home and impacts are less likely to represent causal relationships, compared with the black box related to program model, implementation systems, and home visitor characteristics. Second, what happens in the home might be closely related to the program model being used in a site. For example, home visits will presumably happen more frequently in sites that use program models with weekly visits than in sites that use program models with biweekly visits. From a statistical point of view, this would make it difficult to distinguish the independent effects of what happens in the home visits from what is intended to happen in the home visits. Because there is likely to be widespread variation in program implementation across sites, this might not be a problem in practice. Despite these potential problems, understanding the role of what happens during home visits is important, and the evaluation would certainly explore this question.

In investigating the link between program features and program impacts, the evaluator would need to be parsimonious about inclusion of features in order to preserve statistical power. Both theory and prior evidence would be used to choose which features would be included in the analysis. For example, theory and prior evidence might suggest that one set of features is important for studying how maternal health is improved, but a different set of features is important for studying how child development is improved. Likewise, theory might suggest that the duration and intensity of home visits may affect the full range of outcomes.

In addition, priority would be given to program features that are determined at the site level rather than on a family-by-family basis. For example, the planned intensity and duration of home visiting services would be preferred over the actual intensity and duration for a specific family because the latter would be influenced by unobserved family attributes and may consequently provide biased estimates of the effects of actual intensity and duration on family outcomes.

Because the statistical power of this analysis depends on how related program features are to one another, final decisions about the analysis might not be made until after data are collected. If the data suggest that many program features are unrelated to one another, a more expansive analysis could be conducted. If, as is more likely, program features are highly related within a site, the evaluator would prioritize which small number of features could be included in

the analysis while minimizing the possibility that the results would be biased by the exclusion of important features.

Analysis of Health Care Practices, Health Disparities, and Health Care System Quality

Because of the potential of home visiting programs to affect health care use and health outcomes, the ACA calls for the federal evaluation of home visiting programs to include an assessment of the "potential, if scaled broadly, for improved health care practices, elimination of health disparities, and improved health care system quality, efficiencies, and cost reduction." This section describes proposed methods for defining outcomes related to the health care system and investigating the effects of home visiting programs on that system.

Home visiting could affect health disparities and the health care system in several ways. First, by working directly with families to address issues such as maternal and child health, positive parenting practices, safe home environments, and access to services, home visiting programs may improve health care and health outcomes, and, as a result, directly reduce health disparities. Second, home visiting programs may affect the health care system by interacting with health care providers. For example, a home visitor might seek information from a provider on behalf of the families or advocate for families to ensure that all of their needs are met. By doing so, the home visitor may influence the quality of care provided to the family. Finally, home visiting might indirectly affect health care practices by changing families' information about and use of health care services. For example, if more families with similar needs go to a particular practice, that organization might change the way care it provides care to accommodate the needs of these families.

This section is organized into three subsections. The first section describes health disparities and discusses the home visiting programs' potential to reduce health disparities. The second subsection includes a similar discussion of health care quality and efficiency. The third section examines how home visiting programs may affect the practices of health care providers.

Health disparities

Health disparities are differences in the presence of disease, health outcomes, or access to health care between segments of the population, which may be defined by social, demographic, environmental, and geographic attributes.[34] Socioeconomic disparities have been documented for children and adults in access to and use of health care, and in health outcomes.[35] It is particularly important to examine the issue of disparities as it relates to the home visiting population because studies have shown that children's experiences and environment early in life influence their entire life course. Research by Clyde Hertzman and others shows that early child development— including the physical, social-emotional, and language-cognitive domains—strongly influences basic learning, school success, economic participation, and health.[36] There is a consistent association between socioeconomic status and a variety of development and health outcomes throughout the life course.

---

[34]Carter-Pokras and Banquet (2002); Truman et al. (2011).
[35]Agency for Healthcare Research and Quality (March 2010).
[36]Irwin, Siddiqi, and Herzman (2007).

Although they vary in magnitude by condition and population, disparities are observed in almost all aspects of health care including health care quality, access to care, and treatment of many clinical conditions. Family socioeconomic status has an impact on outcomes as diverse as low birth weight, risk of dental cavities, poorer cognitive test scores, and difficulties with behavior and socialization.[37] Health disparities are greatest for certain diagnoses that are particularly sensitive to delays in medical care, such as asthma and diabetes, as well as others that are consistent with social stress, such as intentional injuries.[38] There is still insufficient evidence on the effectiveness of particular interventions in reducing specific disparities among particular populations.

Despite the gains in health care delivered to children, there are still specific populations of children who have not benefited from these improvements.[39] For example, recent findings show that large disparities in infant mortality rates persist even though child mortality has declined overall. The infant mortality rate is an important indicator of the health of a nation because it is associated with maternal health, quality of and access to medical care, socioeconomic conditions and public health practice. In the U.S., infants born to African American women are 1.5 to 3 times more likely to die than infants born to women of other races and ethnicities.[40] African American and American Indian or Alaskan Native children had death rates about 1.5 times as high as white children in 2003 and, in 2005, African American infants were more than twice as likely as white infants to die during their first year.[41] Examination of child mortality by state, race, and age from 1985 to 2004 found that declines in child mortality were substantial for children of all races and ethnicities.[42] However, the ratio of infant mortality between African American and white children remained unchanged during the study period.

To assess whether home visiting reduces health disparities, the evaluation would follow the steps used to report outcomes of Healthy People objectives.[43] Using this method, the effects of home visiting on health disparities would be estimated by comparing the estimated effects from the national evaluation for an underserved subgroup to the gap in outcome levels between that group and a reference group as suggested by prior research. For example, the estimated effect on fetal death for African American families in the national evaluation would be expressed as a percentage of the difference in fetal death rates between African American and Asian families as indicated in the literature or published statistics. In this example, Asian families would serve as the reference point because they are the racial group with the lowest fetal death rate. Similarly, the effect on disparities can also be measured by gender, education level, income level, and geographic location. For example, estimated effects on health outcomes for the subgroup of families below the poverty level in the national evaluation could be compared to the gap in outcomes between poor families and families with income above 400 percent of the federal poverty level.

---

[37]Irwin, Siddiqi, and Herzman (2007).
[38]Wise (2004).
[39]Berry, Bloom, Foley, and Palfrey (2010).
[40]MacDorman and Mathews (2011).
[41]Agency for Healthcare Research and Quality (March 2010).
[42]Howell et al. (2010).
[43]Centers for Disease Control and Prevention and Health Resources and Services Administration (2007).

The effects on health care disparities would be assessed using the constructs related to health and health care discussed earlier in this chapter. These include measures of child health such as birth weight and overall child health; use of health care services, including the range of preventive and primary care that were shown in Table 5.2; and measures of maternal health including the global health measure, depression and other mental illness, and subsequent pregnancies.

Health care quality

The widely accepted definition of health care quality formulated by the Institute of Medicine is: "Quality of health care is the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge".[44] High quality of care occurs when appropriate and skillfully provided care is available when an individual needs such care, is delivered in a humane manner consistent with the individual's preferences, and the best possible outcomes are achieved.[45] The Agency for Healthcare Research and Quality (AHRQ) defines quality of health care as getting care that is effective, safe, timely, patient centered, equitable, and efficient.[46] Health care quality can be measured by examining how well providers deliver needed services or by outcome measures that may be affected by the quality of health care received. It can also be assessed from the patients' perspective on how well providers meet the patients' health care needs.

The literature suggests that a different framework and method should be used to measure health care quality for parents and children.[47] Childhood is a unique period of life with unique health care needs. Children undergo rapid and continuous cognitive, social, and physical developmental change, which requires different health systems than adults and different approaches to quality measurement.[48] The child health system has a greater reliance on public health, community clinics, and other safety net providers, and this implies a potential for fragmentation of care and discontinuity (for example, not having a medical home). Children also get their care from multiple sites (for example, the health system, schools, juvenile justice, social services, community clinics), implying a greater potential for problems with coordination of care.[49]

Pediatric health care quality should include a focus on primary care activities such as preventive services and anticipatory guidance. Along with immunization, anticipatory guidance and health monitoring are important to well-child care for healthy children. The American Academy of Pediatrics (AAP) and Maternal and Child Health Bureau (MCHB) both provide recommendations for improving the quality of health promotion and preventive services for infants, children, and adolescents.[50] In 2004, the Commonwealth Fund published a report titled "Quality of Health Care for Children and Adolescents: A Chartbook" to raise greater public awareness of the state of health care quality for children and adolescents, incorporating AAP and

---

[44]Institute of Medicine (1990).
[45]McGlynn (1997).
[46]Agency for Healthcare Research and Quality (March 2010).
[47]Leatherman and McCarthy); Seid, Varni, and Kurtin (2000).
[48]Seid, Varni, and Kurtin (2000).
[49]Seid, Varni, and Kurtin (2000).
[50]Hagan, Shaw, and Duncan (2008).

MCHB recommendations.[51] These recommendations call for periodic monitoring, screening, and guidance for all children.[52] Furthermore, preventive care is an essential part of the AAP's Medical Home policy statement. Specifically, the AAP states that primary care services should include "growth and developmental assessments, appropriate screening, health care supervision, and patient and parent counseling about health, nutrition, and safety."[53]

Another approach to thinking about health care quality is through the current AHRQ recommendations to improve quality of care.[54] These items suggest examining a variety of health care practices related prevention and health promotion (for example, prenatal care, screening, immunizations), availability of services (for example, having a usual primary care provider, having access to a dentist), the management of acute conditions (for example, by examining appropriate use of antibiotics and dental care), management of chronic conditions, and family experiences of care (for example, as indicated by communication between health care providers and families).

To analyze the effects of home visiting programs on health care quality, the evaluation would estimate the effects of home visiting on process of care measures described in Chapter 4. These include parent reports on usual source of care, immunizations, developmental screening, early intervention services, receipt of well child care, and so on. In addition, collecting State Medicaid and SCHIP claims data, as described in Appendix B, could provide more detailed information on the receipt of primary and preventive care.

Health care practices

As noted earlier, home visiting might indirectly alter how care is provided by health care providers, especially those working with home visiting enrollees.[55] However, a home visiting program seems likely to influence the way health care organizations operate only under two circumstances: (1) the program is so saturated in the community that it affects how health care is provided to low income families, or (2) the home visiting program works closely with the health care provider to lead it to implement changes in practice at the organizational level. This suggests that the evaluation should prioritize investigating the effects on health care practices in locations that meet one of these two criteria (saturation or close coordination between the home visiting program and health care providers). For this reason, the national evaluation would not collect information on health care practices of a wide range of health care providers. However, the implementation study will collect information on changes in health care practices if the evaluation includes some programs that embed home visitors in health care providers.

Scaling up

---

[51]Leatherman and McCarthy (2004).
[52]Medical Home Initiatives for Children With Special Needs Project Advisory Committee (2002).
[53]Medical Home Initiatives for Children With Special Needs Project Advisory Committee (2002).
[54]Agency for Healthcare Research and Quality (March 2010).
[55]Health care practice is assumed to refer to how health care providers interact with patients and provide care. Changes in health-related behaviors by families are included in the discussion of health disparities and health care quality. In addition, as described in Chapter 4, the evaluation would collect a range of information on such behaviors.

The impact analysis will provide direct information on the effects of home visiting programs included in the national evaluation on health disparities and health care quality. The ACA, however, calls for an assessment of the potential for home visiting, when scaled up, to affect those outcomes. Such scaling up of home visiting programs may extend to the community, regional, or national levels any decreases in disparities found for families and children in the national evaluation.[56]

One relatively straightforward means of inferring the effects of home visiting when scaled up is simply to multiply the estimates from the national evaluation by the number of families who could be served by the program. This would, in a sense, provide an estimate of the potential effect nationally. If, for example, the national evaluation found that home visiting programs increased the percentage of children with recommended well visits by 10 percentage points, that result can be applied to the number of families who could be served by home visiting to estimate the potential of a scaled up program. Going a step further, the evaluation's results for different types of families and programs could be used to project the effects of a fully scaled up program. In this case, the projections would assume different impacts for each subgroup of families or children who represent a proportion of families and children served by the scaled up program. Similarly, sensitivity tests could be used to project the range of possible effects that could be expected depending on which program features are adopted by expanded home visiting programs.

An alternative, more sophistical approach is suggested by system dynamics modeling. System dynamics modeling is a computer simulation method that aims to explain or anticipate potential outcomes of a policy decision.[57] It works by outlining the consequences and feedback loops embedded in a system. It assumes that the complex behavior of organizational and social systems emerge from accumulations of people, material or financial assets, information, and through balancing and reinforcing feedback loops that result from these build ups.[58] The model has had numerous applications in the health area since it was developed nearly 50 years ago. This model can use organizational characteristics, such as size, funding source, culture, adaptability of staff, willingness to collaborate, and standard operating procedures collected from stakeholder interviews and administrative records, to predict how organizations will implement changes. System dynamics modeling can help to identify feedback loops between various factors and simulate the potential impact of expansions in home visiting programs on children and families.

---

[56]Irwin, Siddiqi, and Herzman (2007).
[57]Norman (2009).
[58]Homer and Hirsch (2006).

Chapter 6

Economic Evaluation

As discussed in Chapter 1, the seven evidence-based models included in the national evaluation have some prior evidence of effectiveness. However, prior studies provide little information regarding the costs or the cost-effectiveness of these programs. The ACA requires that the national evaluation assess the potential for the home visiting program, if scaled broadly,to improve health care system quality, efficiencies, and reduce costs." The national evaluation must include an economic evaluation to address that requirement. HHS has additional questions that could also be answered. The potential research questions underlying the economic evaluation are the following:

- What is the cost to deliver home visiting services that use the evidence-based models, and how do these costs vary across groups of families and local programs?

- What is the cost to achieve key outcomes for families and children, and how do these costs vary across groups of families and local programs?

The information from the economic evaluation will play an important role in supporting the implementation of evidence-based programs by helping organizations that implement home visiting programs make decisions regarding the allocation of their resources. This is especially significant because preventive and early intervention services do not typically receive the level of funding received by programs that work with individuals after they have been diagnosed with a disease.[59]

The economic evaluation would combine results from the impact evaluation with an analysis of the costs of delivering home visiting services. Cost would be collected at the site level to provide information to help agencies determine if they have the resources required to deliver home visiting services in their own communities. Collecting information on program costs and program impacts together allows the analysis to compare costs to outcomes in more rigorous way, for example, by investigating the features of local programs that lead to more cost-effective programs.

An Overview of the Economic Evaluation

Cost-effectiveness analysis compares the costs of providing an intervention to the effects it achieves, resulting in a cost per unit of effect.[60] By expressing results in these terms, costs can be compared across programs designed to affect similar outcomes. For example, the cost to reduce child maltreatment through home visiting can be compared to the cost of reducing child maltreatment through other means. The analysis can estimate the unit cost of achieving any of

---

[59]Barnett (1993).
[60]Gold, Siegel, Russell, and Weinstein (1996).

the outcomes being assessed in the impact analysis, but results will be most policy relevant if the most important outcomes are included in the analysis.

The cost-effectiveness analysis would be done at two levels. A *micro-level* cost-effectiveness analysis would compare the net costs of operating programs in the study to the impacts outlined in Chapter 5. For example, the cost-effectiveness analysis could estimate the cost for averting a case of child maltreatment in each site by comparing the estimated effects at that site to that site's programmatic costs. Although such estimates would be fairly imprecise because the underlying estimates of impacts and program costs would be based on a small number of families, such site-level cost-effectiveness measures could be used to investigate the relationship between program features and program cost-effectiveness. The micro-level analysis can also conduct this analysis for different subgroups of families enrolled in the programs.

A *macro-level* analysis would compare the total costs of the MIECHV legislation to the aggregate impacts of the legislation on key outcomes. This would provide policy makers with an overall understanding of the returns on investment in home visiting programs.

Finally, a benefit-cost analysis, if feasible, would place a dollar value on all program effects and compare the monetized benefits of home visiting to the program costs. Because it can take some time for the full benefits and costs of home visiting programs to be realized, the evaluation would undertake a benefit-cost analysis only if results were collected for five years following random assignment. The following sections describe the procedures and data required for these analyses.

Programmatic Cost Analysis

The first step in conducting the analysis is to collect information on program costs for each site. Program costs, in this case, are the value of all resources necessary to provide a home visiting program, before and during implementation. The costs of these programs are expected to be very different as each program provides a different array of services to its clients, not just visits to the home. Costs are also expected to vary depending upon the population receiving the services, the geographic setting within which the program is being delivered, the education and experience level of program providers, and the level of program development already in place at the time of the analysis. Common categories of costs and the sources for data on costs are shown in Table 6.1.

In short, program costs should include a valuation of all of the resources required to provide the overall program, not just the resources specific to actual visits in the home. According to Barnett (1993), a full analysis of costs for home visiting programs should focus on those categories which are expected to create the largest difference. These include the number and duration of visits, home visitor caseload, home visitor credentials and characteristics, supervision and administration, and parent time.

Design Options for Home Visiting Evaluation Project

Table 6.1

Common Categories and Types of Costs
Present in Home Visiting Models

	MIS	Surveys	Interviews
HV Costs			
Personnel			
Home visitor	X		
Educators	X		
Transportation to visits			
Mileage	X		
Program Vehicles	X		
Supplies and Materials			
Family support materials	X		X
Books and brochures	X		X
Screening materials and tools	X		X
Medical supplies	X		X
Participant Costs			
Time at Home Visit	X		
Time at Other Events		X	
Other Costs		X	
Non-HV Administrative Costs			
Personnel			
Program Manager			X
Supervisor			X
Support Personnel			X
Administrative Personnel			X
Data Entry Person			X
Supplies and Materials			
Paper			X
Office supplies			X
Postage			X
Equipment			
Computers/printers			X
Cell phones/service			X
Copiers			X
Training and Education			
Training costs			X
Professional development			X
Conferences			X
Transportation			X
Buildings and Facilities			
Rent/lease			X
Utilities			X
Phones			X
Internet provider fees			X
Miscellaneous			
Data management systems			X
Certification/recertification			X

Following Barnett's approach, the cost-effectiveness analysis would include three primary areas of costs. The first are the costs required to carry out home visits. This is the primary area in which programmatic costs can be connected to individual families and satisfies most of the categories outlined by Barnett above. The second are the costs of the support structure that exists within each program. This is more commonly referred to as administrative or indirect costs and generally cannot be connected to specific families. The third category of costs includes those incurred by families involved in home visiting programs. Each of these three areas is discussed in more detail below.

Costs of Home Visits

The costs of the actual home visits are best collected through a program's management information system and home visitor logs. These systems could detail for each home visit: 1) the time spent at the home visit; 2) the salary and staff level of the provider delivering the home visit; 3) the travel costs and time for travel required for the home visit; 4) the number of attempts and time to schedule the home visit; 5) the time and other resources required to prepare for the home visit; 6) the resources (supplies and materials) required for the home visit; 7) community referrals resulting from the home visit; and 8) any time or resources required to follow-up on the home visit. By collecting this resource use for each home visit, it will be possible to assess average costs of a home visit by program, by provider type, and by family characteristics. If these data are collected for all participating families in the impact analysis, it will also be possible to aggregate total costs of home visiting services by intensity and dosage of participation in the program.

Administrative Costs

The other significant area of costs is the cost of the support or administrative infrastructure that programs have in place that allows the home visits to occur. This includes administrative and supervisory personnel time, supplies, materials, equipment, building space, overhead, and training. These data could be collected at baseline and on a quarterly basis throughout the impact analysis period, from a senior administrative staff person responsible for program budget. A cost collection form could be developed at the beginning of the evaluation with input from senior administrative staff. Webinars and other effective means for training could be administered to all participating administrative staff. Data could be collected via phone interviews or in-person meetings by an evaluation team member explicitly responsible for the economic evaluation.

An example of a cost collection form that has been developed for collecting infrastructure costs is the Drug Abuse Treatment Cost Analysis Program (DATCAP), which has been used to collect costs from drug abuse treatment centers.[61] This instrument provides a solid foundation for how a cost collection form to assess administrative infrastructure costs could be constructed. The DATCAP collects information on administrative personnel time, buildings and facilities, equipment, supplies and materials, and miscellaneous resources and costs used to support a program's infrastructure. To date, this form has been successfully used to collect costs from a

---

[61]French et al. (1997); French and McGeary (1997).

wide variety of substance abuse treatment programs across the country through interviews with program staff. The collection of home visitation infrastructure costs could be collected in a similar manner and is likely to have equally successful results.

Once a collection form has been developed, webinars are an effective and inexpensive method of training for sites included in the evaluation. Sites would have a copy of the form and instructions and a key individual tasked with collecting site information would attend the webinar. This would give the research team a chance to inform sites on the purpose of the cost analysis including important issues such as the perspective of the study and why budgets are not sufficient for the analysis. Following the webinars, telephone interviews should be used to complete the collection forms on a quarterly basis.

As mentioned above, the programmatic costs specific to each home visit can be allocated directly to the respective family. However, costs that are administrative in nature cannot be directly allocated in the same way and instead an indirect method should be used. For non-personnel administrative categories, the costs collected at each site during a specific period of time could be divided among the families actively enrolled during that time. Personnel administrative categories can be allocated in a similar manner, but administrative personnel should report, during cost collection interviews, the proportion of their time specific to the home visit program that is not research in nature. This is of particular importance for sites in which more than one program is administered. This proportion of each individual's salary should then be allocated to families using the same method as those for non-personnel administrative categories.

Participant Costs

The final area of costs that would be important to the analysis is the cost to families in the home visiting program. A program's management information system can provide detail on home visits for each family by personnel, including the time spent at home visits by participants. To assess the value of this time, either the annual household income of the family can be used or the average wage in the community. To assess participants' costs outside of home visiting, participants could be surveyed about resource use related to the program as part of the six-month or twelve-month follow-up survey. The survey should address collection of other resources or time spent outside of the actual home visit. For example, activities similar to homework may be assigned by home visitors for the families to complete on their own. Also, some models encourage parents to attend events and the costs and time required by participants for traveling to these events could be estimated. This collection would only be necessary for programs implementing models where time and resources are consumed outside of visits to the home.

Other Considerations for Collecting Programmatic Costs

In addition to collecting program costs for the three areas described above, the evaluation should also collect costs from the societal perspective. The societal perspective is recommended for economic evaluations of interventions funded with societal resources.[62] That is, all costs should be included in the analysis regardless of to whom the costs accrue. This is the rationale

---

[62]Gold, Siegel, Russell, and Weinstein (1996); Haddix, Teutsch, and Corso (2003).

for including participant costs even though these are costs not borne by the provider or by the funding agency. Further, the inclusion of societal costs also means that one should include the value of resources for which there is no monetary exchange. That is, one should consider including the value of program resources that are free or donated, for example, donated space to house the home visitor provider staff. The rationale for including this cost, although it may not appear on a budget sheet, is that in attempting to assess the average cost for delivering the program in such a way that another site may be able to use for planning purposes, to exclude key costs may misrepresent the real cost of implementation. Second, if the donated space is used for the MIECHV program, that space cannot be used for some other program and thus represents an opportunity cost for that resource. Another example of an opportunity cost is the value of a family's time spent participating in the MIECHV program – else, the family could be participating in some other useful contribution to society.

Another consideration for collecting costs is the time period during which costs are being collected at the program level. Costs could be collected during start-up (pre-implementation), new implementation, or ongoing implementation. Because the evaluation will focus on programs that have been in operation for two or more years, start-up and new implementation costs would include the costs of expanding programs to use new MIECHV funds. As programs ramp up their implementation, it is expected that costs will vary, so this is an important consideration when comparing costs of one program to another. For example, during the pre-implementation period families are either not enrolled or the program runs far short of capacity. This means that most of the resources consumed are for training and purchasing of supplies and equipment important for building implementation capacity. A start-up period is required of any home visiting program and the total costs of this pre-implementation period should be annuitized, just as with the aforementioned equipment example. We suggest that some effort be made to collect budget information from the early years of sites involved in the evaluation to help estimate this cost.

Health Care Costs

The other major cost type necessary for the cost-effectiveness analysis is the cost of health care services used by participating families. Health care costs can include inpatient services, Emergency Department services, outpatient services, mental health services, drug and alcohol treatment or prevention services, medications, medical devices, and so on. A given intervention might reduce the costs of many of these types of services for the intervention group relative to the control group. Any cost savings would be subtracted from the programmatic costs to represent the savings from illness averted, effectively decreasing the difference in programmatic costs between the two groups (when the treatment costs more than the control). Since one of the goals of many home visiting models is to connect participant families to basic services such as primary health providers, home visiting group may increase health care costs, especially in the first year or two of the program. Thus, including health care costs could either increase or decrease the effective costs of home visiting programs in the short run.

If Medicaid and SCHIP administrative data are collected, as presented in Appendix B, that information can be used to estimate the health care costs of participant families. Because most families in the evaluation will be low income, the majority of their health care resource usage could be captured this way. Ideally the information would include visits to physicians, dentists, emergency rooms, outpatient departments, and other health care providers; prescription medications filled; and other medical expenses. If the optional module is not funded, the follow-up survey can ask parents to report on their families' health care use. The latter approach is not preferred because individuals are unlikely to accurately recall all of their usage over the previous time period. In addition, it adds burden to the researcher and participant at the data collection visit or supplants other information that could be collected from the survey.

Economic Analysis

Program costs can be linked to program impacts at the site level to determine the cost-effectiveness of the program relative to a control group. The comparison can be made for whichever individual impacts estimates are deemed of most interest. The cost-effectiveness analyses will allow a comparison of home visiting with other programs that target similar outcomes or comparison of populations participating in a home visiting program to populations not participating in a home visiting program. The cost-effectiveness analyses can also be carried out on different subgroups of families or providers in the evaluation.

The relevant summary measure is called the incremental cost-effectiveness ratio (ICER) and is calculated as shown in equation (3):

$$\overline{\hspace{8cm}} \tag{3}$$

Program costs collected during the cost analysis are used in the numerator of the ICER. Costs averted would include reductions in medical and non-medical costs, and can also include reductions in productivity losses as might happen if home visiting helps parents avoid missing work. Outcomes can include any short-term or long-term outcomes that are directly collected in

the evaluation study. The inclusion of long-term outcomes not directly collected in the evaluation study would require modeling and are outside the scope of this evaluation. For each outcome deemed key in the impact analysis, separate cost-effectiveness ratios comparing the evidence-based program to the control can be determined. Separate ratios can also be determined for analyses of subgroups of providers or families.

Outcomes for the Cost-Effectiveness Analysis

Several child health measures could be used as outcomes in the cost-effectiveness analysis, including height and weight of children, low birth weight for newborns, cognitive development, and substantiated cases of child maltreatment. The height and weight of children whose families are participants in the home visiting program will be collected at baseline and follow-up data collections. Low birth weight would only be used for home visiting programs that enroll mothers several months before the birth of the child. Mothers who receive services for a period of time prior to having the child are more likely to have a healthier newborn and therefore less likely to have a child with low birth weight. Cognitive development is set on a scale based on the results of an instrument. Maternal depression, substance use, smoking cessation, and birth spacing are the primary non-child health measures that might be used in the cost-effectiveness analysis. Finally, the rating of parent-child interaction is the primary outcome of interest involving both the parents and children in the evaluation.

Summary Measure and Analysis Plan for the CEA

Separate cost-effectiveness analyses should be conducted for each of the outcomes of interest. Best practice guidelines prescribe calculating an ICER when comparing the cost-effectiveness of the home visiting program to the control group.[63] The ICER represents the incremental costs divided by the incremental effectiveness and describes the additional cost per additional health outcome prevented and can be interpreted as the incremental cost of producing effectiveness by one intervention compared to the control or usual services strategy. The formula for the ICER comparing 2 programs was outlined in Equation (3). The numerator of the ICER is derived from each program's net costs, or cost of delivering the intervention less the costs of the health conditions that are prevented as a result of the intervention. The results of the cost analysis outlined above will be used as the cost of delivering the home visiting programs. In a typical cost-effectiveness analysis, health costs averted include averted medical care treatment and averted economic monetary losses. The denominator in the ICER is the intervention-induced improvement of the outcome of interest. The final result is a measure of the net costs of the intervention for each one unit gain in the outcome measure, for example the cost per case of child maltreatment prevented.[64]

Subgroup Analysis

---

[63]Gold, Siegel, Russell, and Weinstein (1996); Haddix, Teutsch, and Corso (2003).

[64]Sensitivity analyses would also be conducted for the CEA and CUA estimates. If collected costs from this analysis and outcomes from the impact analysis are connected to participant families, many of the current gold standard methods could be used. Current gold standards for sensitivity analysis include confidence ellipses, net benefit regression, and cost-effectiveness acceptability curves.

Given the nature of the underlying outcome evaluation, it would be possible to use this information to connect the total programmatic costs to specific subgroup outcome indicators. This process would be very similar to the process described above for the full sample, by splitting the sample of costs and outcomes into distinct groups of families or providers and then carrying out the economic analyses. However, rather than simply being interested in whether the outcomes differ among the different subgroups, it would be possible to test how the cost-effectiveness results differ across these subgroups. For example, the impact study may find that a specific subgroup showed a greater impact across some important domain. The subgroup analysis carried out for the economic evaluation may find that group to use more resources and be more expensive. The result of the cost-effectiveness analysis would then be the additional cost required to achieve that improved outcome. The net-benefit framework provides an excellent method for carrying out subgroup analysis.[65] A regression model with the treatment variable and covariates such as age of the child, race, sex, or socio-economic status could be constructed. The magnitude and significance of the coefficients on the interaction between the covariates and the treatment variables can then provide an estimate of cost-effectiveness by subgroup.

Benefit-Cost Analysis (Five-Year Plan)

A benefit-cost analysis has the added usefulness of allowing home visiting programs to be compared to programs that impact different health and non-health outcomes. A benefit-cost analysis involves measuring both the costs and the benefits of a program in dollar terms, producing a measure of a program's returns on investment or net social benefits. The outcomes collected in the impact analysis would determine the outcomes used in the benefit-cost analysis. The ideal benefit-cost analysis would consider both the health and non-health benefits of the programs, estimating non-monetary benefits in some capacity, and following participants long enough to capture potential lifetime benefits. Because this is not feasible for this study, a benefit-cost analysis could instead rely on five-year outcomes for the families. Longer-term benefits beyond five years could be estimated using a model to extrapolate five-year outcomes to lifetime benefits. For example school readiness gives an indication of the expected level of school completion and ultimately estimates potential future earnings.

The most difficult aspect of a benefit-cost analysis, especially when compared to the cost-effectiveness analysis, is converting outcomes into monetary values. For many outcomes, a review of the relevant literature provides the framework necessary for this analysis. This includes changes in productivity for the parents, less dependence on government benefits, a reduction in the use of services such as special education for the child, and the effect of improved educational outcomes on long-term productivity for the child.

---

[65]Claxton and Posnett (1996a, 1996b); Stinnett and Mullahy (1998).

<div align="center">

Chapter 7

Resource Estimates

</div>

Table 7.1 summarizes the resource estimates for conducting the evaluation described in this design report. These resources include those needed to meet legislated requirements as well as additional HHS goals. Details are throughout the report and assumptions for estimating these costs are found in the technical appendix to this document. These costs are contingent on what states describe in their state plans. Just as the design may need to be revised based on updated state plans, these resource estimates may also require revisions.

<div align="center">

Design Options for Home Visiting Evaluation Project

Table 7.1

Summary of Resource Estimates for the Designed Evaluation

</div>

Resource Estimates for Basic Design	Dollars	Labor Hrs
Impact Study[1]	$31,423,000	270,800
Implementation Study	$5,156,000	34,500
Cost Effectiveness Study	$1,101,000	6,380
Needs Assessment Study	$197,000	1,780
Other management & evaluation costs[2]	$2,070,000	11,180
Total estimate	**$39,947,000**	**324,640**

NOTES:

[1]This estimate assumes a 12 month follow up data collection effort. If a 6 month follow up is selected the estimate would be reduced to $30,681,000 and 264,460 hours.

[2]Includes ongoing consultation w/COTR, HHS, HV program developers & other HV experts, public use file.

Appendix A

Resource Estimates Technical Appendix

The purpose of this appendix is to present the basic assumptions underlying the calculation of resource estimates for the federal evaluation of home visiting programs. The current resource estimates total $39.9 million (see Table 7.1 for details).[66]

Assumptions about the sample

- The evaluation includes a total of 12 different home visiting grantees (either large geographic areas or states).
- Each grantee operates 10 local home visiting sites (or local program offices).
- The 12 grantees operate 7 different home visiting models.
- Teen mothers who need parental consent to enroll in the study make up 10% of the HIPPY and Family Check-Up sample and 20% of the five other program models' samples.
- The evaluation includes a total of 7,200 families.
- This means there are 3,600 program families and 3,600 control families (60 families per local site; 30 program and 30 control).
- Each site contains 4 home visitors (HV), 1 supervisor, and 1 program manager.
- The sample is drawn from the 48 contiguous states.
- All local sites are within 2 hours of an airport.

---

[66]This estimate does not include a fee for the evaluator.

Resource Estimates Assumptions

Impact Study

Recruitment
- Finalize site selection criteria, review state plans, phone and email communications
- Three rounds of two-person trips to each grantee candidate

Baseline data collection
- Collect informed consent forms, conduct random assignment, and conduct a parent interview

Technical assistance (TA) and site development
- Yearly local site payments of $10,000 per local site for 3 years ($10,000*120 sites*3 yrs) totaling $3.6 million
- Provide TA for study-related assistance (evaluator will not to provide TA on program operations)
- Provide regular TA by phone and email by site liaisons
- One cross-site, 2-day conference for site program managers and state level representatives
- Two in-person technical assistance trips for each of the 12 grantees over the course of the study
- In-person data collection training and kick-off meeting for each grantee (program managers, supervisors and state level representatives)

Design, measurement & instrument development
- Finalize design plan and develop and test instruments

Follow-up data collection- survey data
- Conduct in-person parent interview at 6 or12 month follow-up

Follow-up data collection- direct child and parenting assessment data:
- Parent-child video interaction at 6 or 12 month follow-up
- 6 or 12 month follow-up child developmental assessments for older children only

Follow-up data collection- administrative data
- Collect child welfare records from 12 different state agencies
- Collect birth records from 12 different state agencies

Analysis and report writing
- Report to Congress in 2015 and a 6 or 12-month impact report

Resource Estimates Assumptions

Implementation Study

Design, measurement & instrument development
- Final design plan and development and testing of instruments

Follow-up data collection- survey data
- 2 rounds of annual program manager interviews
- Collect data on control service environment
- Home visitor skills, knowledge, perspectives

Follow-up data collection- videotaped observations
- One home visit videotape per family

Follow-up data collection- administrative/automated data
- Obtain electronic centralized MIS files from 3 grantees or programs
- Collect home visit and supervision

Field visits
- One visit to each grantee (12 total)

Analysis and report writing
- Implementation research portion of the Report to Congress and 6 or 12-month report

Cost Effectiveness Study

Design, measurement & instrument development
- Finalize design; develop and test instruments

Follow-up data collection- cost effectiveness specific data
- Collect program expenditure data

Analysis and report writing
- Cost-effectiveness research portion of the Report to Congress and the 6 or 12-month report

Needs Assessment Study

Design
- Finalize design

Analysis and report writing
- Needs assessment research portion of the Report to Congress

Resource Estimates Assumptions

Other Management and Evaluation Costs

General management
- Weekly check-ins with the COTR/HHS staff and annual in person trips to Washington DC
- Prepare monthly progress reports; agreements
- Outside consultants
- Create public use file and archive videos

Communication and data agreements with home visiting program developers
- One in-person visit to each of the 7 program model developers in addition to phone calls and emails
- Negotiate any data agreements needed

Appendix B

Supplemental Research Modules

A number of additional research activities could be conducted as part of either the national evaluation or an ongoing research agenda. These include the following:

- Perform a direct assessment of child outcomes for younger children when they are six months or one year old (depending on which follow-up period is chosen for the evaluation). The core evaluation would collect parent reports through interviews, would videotape parents interacting with their children, and would collect direct assessments of children's outcomes for the older cohort of children. Because prior research suggests that direct assessments of one-year old children may be less informative, such direct assessments were not included in the core evaluation, but are discussed as an optional task.

- Collect an additional round of follow-up data at 24 months. The core evaluation would collect one round of follow-up data on families at six or twelve months. Because many of the national program models provide home visiting for more than six or twelve months and because family and child outcomes might continue to evolve as the child gets older, the design includes an option for collecting and analyzing additional information at 24 months. Data collected at 24 months would include the same constructs as those measured at the six or twelve month point, using surveys of parents, direct observations of parents interacting with their children, direct assessments of children, and information on parents' participation in home visiting and associated services in the second year. For younger children, data would be collected around the time of the child's second birthday. For older children, data would be collected about two years after the family enters the study.

  The evaluation of outcomes for children, particularly those in the younger cohort, would benefit from this later follow-up, because the instruments available to measure children's language development are likely to be more sensitive to intervention effects when the child is 24 months of age than at one year. (Similarly, there may be interest in long-term follow-up modules at 36 months and 60 months, which would allow for better understanding of long-term impacts on school readiness for the prenatal group, as well as long-term impacts on other key outcome domains.)

  In the event of a 24-month follow up, a number of things related to the cost analysis would change. This assumes that a follow-up collection of outcomes for all participants is made at 6 or 12, and 24, months for outcomes important to the economic evaluation. In addition to using the additional follow-up data described in Chapter 5, the following changes would need to be made:

  - The interview of all sites using the standard cost collection form would need to be carried out in the second year.

- o   Additionally, the final analysis should present all costs for a common base year, and would necessitate adjusting for inflation using the consumer price index.
- o   Participants would need to be surveyed at the 24-month follow-up to determine costs for the second year.
- o   Information would be collected about personnel time and transportation costs throughout the second year.

- Collect and analyze Medicaid and State Children's Health Insurance Program (SCHIP) claims data at six or twelve months (depending on which shorter-term follow-up period is chosen) and again through 24 months. Although surveys may be used to collect information on health care use for parents and children, Medicaid claims are likely to provide more accurate information on health care use and costs for Medicaid beneficiaries and children who are insured through SCHIP. In addition, administrative data add less burden to the participant than self-report data.

- Add a subgroup of families living in frontier areas to the evaluation. Because there is an interest in understanding the effects of home visiting for families living in frontier areas, the design includes an option to add enough of these sites to provide precise estimates of impacts in those sites.

- Study the implementation of new programs. As described earlier, the national evaluation would include only sites that had been in operation for two or more years. Although including new programs in the impact analysis would likely understate the effects of home visiting, information on the implementation of new sites could be valuable to the startup of future programs. It is important to understand the characteristics of communities in which grantees decide to locate brand new programs. In addition, results from such an analysis could provide insights into the features of implementation systems and communities that are associated with reaching optimal implementation capacity and fidelity as quickly as possible.

- Conduct in-depth qualitative interviews with mothers. In addition to surveys of parents that would be used for the impact analysis, this module would conduct three rounds of in-depth semi-structured interviews with 200 mothers across five states. The sample would include both teen and older mothers to learn about the experiences of both groups. Interviews would be conducted with both program group and control group mothers to gain a better understanding of how home visiting is viewed by mothers, how it may be benefiting program group mothers, and the unmet needs of control group mothers. Results from these interviews could be included in the 2015 report to Congress.

- Conduct in-depth qualitative interviews with fathers. The core evaluation would ask mothers about fathers' level of involvement in the lives of their children. To gain additional information on the perspective of fathers, this module would conduct three rounds of in-depth semi-structured interviews with 100 fathers across five states. Interviews would be conducted with both program group and control group fathers to

gain a better understanding of fathers' engagement in home visiting. Results from these interviews could be included in the 2015 report to Congress.

Resource Estimates Assumptions for Supplemental Research Modules

Module 1: Direct child assessments
- 6 or 12 month follow-up child developmental assessments for younger children

Module 2: 24 month follow-up
- Conduct a parent interview, child developmental assessments for older children, parent-child video interaction; update participation data for families who participate for more than 12 months; update cost-effectiveness study, and 24 prepare month report

Module 3: Medical records for 6 or 12 month follow-up
- Collect and analyze SCHIP and Medicaid records for 6 or 12 month follow-up

Module 4: Medical records for 6 or 12 and 24 month follow-up
- Collect and analyze SCHIP and Medicaid records for 6 or 12 month follow-up and 24 month follow-up

Module 5: Frontier subgroup
- Sample of 20 additional local home visiting sites, conduct all basic design activities

Module 6: New site implementation substudy
- Collect preliminary information on grantees with new home visiting program sites
- Conduct two rounds of in-person visits to 3 of the grantees with new programs
- Analysis of data and report writing for inclusion in the Report to Congress

Module 7: Participant interviews of mothers
- Conduct three rounds of in-person interviews with a subsample of 200 mothers across 5 grantees
- Sample includes both teenage mothers and non-teenage mothers from the program and control groups
- Analysis of data and report writing for inclusion in the Report to Congress

Module 8: Interviews of fathers
- Conduct three rounds of in-person interviews with a subsample of 100 fathers across 5 grantees
- Sample includes fathers from the program and control groups
- Analysis of data and report writing for inclusion in the Report to Congress

Design Options for Home Visiting Evaluation Project

Table B.1

Summary of Resource Estimates for Supplemental Modules

Resource Estimates for Additional Modules	Dollars	Labor Hrs
Module 1: Direct child assessments	$4,153,000	35,890
Module 2: 24 month follow up	$14,143,000	150,840
Module 3: 6 or 12 month medical records	$1,103,000	8,310
Module 4: 6 or 12 & 24 month medical records	$1,801,000	12,650
Module 5: Frontier subgroup	$5,594,000	50,920
Module 6: New site implementation substudy	$658,000	43,380
Module 7: Participant interviews- mothers	$2,255,000	13,450
Module 8: Interviews of fathers	$1,374,000	8,260

Reference List

Agency for Healthcare Research and Quality. March 2010. *National Healthcare Disparities Report 2009*. Rockville: AHRQ Publication.

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. 1996. "Identification of causal effects using instrumental variables." *Journal of the American Statistical Association.* 91, 434: 444-455.

Barnett, W. Steven. 1993. "Economic evaluation of home visiting programs." *The Future of Children.* 3, 3: 93-112.

Brooks-Gunn J, Duncan GJ, Klebanov PK, Sealand N. 1993. Do neighborhoods influence child and adolescent development? *American Journal of Sociology.* 99,2: 353–395.

Berry, J. G., S. Bloom, S. Foley, and J.S. Palfrey. 2010. "Health inequity in children and youth with chronic health conditions." *Pediatrics.* 126 Suppl 3: S111-9.

Bloom, Howard S., and Charles Michalopoulos. 2010. *When is the Story in the Subgroups? Strategies for Interpreting and Reporting Intervention Effects on Subgroups*. New York: MDRC.

Carroll, C., M. Patterson, S. Wood, A. Booth, J. Rick, and S. Balain. 2007. "A conceptual framework for implementation fidelity." *Implementation Science.* 2: 40.

Carter-Pokras, O., and C. Baquet. 2002. "What is a "health disparity"?" *Public Health Reports.* 117, 5: 426-34.

Centers for Disease Control and Prevention, and Health Resources and Services Administration. 2007. "Healthy People 2010 Midcourse Review." Web site: http://www.healthypeople.gov/2010/Data/midcourse.

Claxton, K., and J. Posnett. 1996a. "An economic approach to clinical trial design and research priority-setting." *Health Economics.* 5, 6: 513-524.

Claxton, K., and J. Posnett. 1996b. "An economic approach to clinical trial design and research priority-setting." *Health Economics.* 5, 6: 513-524.

Dane, A.V. and B.H. Schneider. 1998. "Program intensity in primary and early secondary prevention: Are implementation effects out of control?" *Clinical Psychology Review*. 18: 23-45.

Duggan A. K., A. M. Windham, E. Y. McFarlane, D. Salkever, C. Rohde, L. Rosenberg S. Buchbinder, L. Fuddy, and C. Sia. 1999. "Evaluation of Hawaii's Healthy Start Program." *The Future of Children: Home Visiting, Recent Program Evaluations.* 9,1:66-90.

Duggan, A.K., L. Fuddy, L. Burrell, S. Higman, E. McFarlane, A. Windham, and C. Sia. 2004. "Randomized Trial Of A Statewide Home Visiting Program To Prevent Child Abuse: Impact In Reducing Parental Risk Factors." *Child Abuse & Neglect*. 28, 6:6: 23-643.

Duggan, A., D. Caldera, K. Rodriguez, L. Burrell, C. Rohde, and S.S. Crowne. 2007. "Impact of a Statewide Home Visiting Program to Prevent Child Abuse." *Child Abuse & Neglect.* 31, 8: 29-52.

Duncan, G. J, J. Ludwig, and P. Hirschfield. 2001. "Urban poverty and juvenile crime: Evidence from a randomized housing-mobility experiment." *The Quarterly Journal of Economics.* 116, 2: 655-679.

Duncan, Greg J. and Stephen Raudenbush, eds. 2001. Neighborhoods and adolescent development: how can we determine the links? In Does It Take a Village? "Community Effects on Children, Adolescents, and Families" ed. A. Booth and A. C. Crouter. State College, Pennsylvania: Pennsylvania State University Press. 105-36.

Durlak, J. A., and E. P. DuPre. 2008. "Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation." *American Journal of Community Psychology.* 41, 3-4: 327-50.

Elliot, D., and S. Mihalic. 2004. Issues in disseminating and replicating effective prevention programs. *Prevention Science, 5*, 47-53.

Ennett, S.T., C.L. Ringwalt, J. Thorne, L.A. Rohrbach, A. Vincus, A. Simons-Rudolph, and S. Jones. 2003. "A comparison of current practice in school-based substance use prevention programs with meta-analysis findings." *Prevention Science.* 4:1-14.

Fixsen, D. L., S. F. Naoom, K. A. Blase, R. M. Friedman, and F. Wallace. 2005. *Implementation Research: A Synthesis of the Literature*. Louis de la Parte Florida Mental Health Institute. The National Implementation Research Network.

French, M. T., L. J. Dunlap et. al. 1997. "A structured instrument for estimating the economic cost of durg abuse treatment - The Drug Abuse Treatment Cost Analysis Program (DATCAP)." *Journal of Substance Abuse Treatment.* 14(5): 445-455.

French, M. T. and K. A.McGeary. 1997. "Estimating the economic cost of substance abuse treatment." *Health Economics*. 6(5): 539-544.

Gennetian, Lisa A., Pamela A. Morris, Johannes M. Bos, and Howard S. Bloom. 2005. *Constructing Instrumental Variables from Experimental Data to Explore How Treatments Produce Effects*. New York: NY: MDRC.

Gold, M. R., J. E. Siegel, L. B. Russell, and M. C. Weinstein. 1996. Cost-Effectiveness in Health and Medicine. New York: Oxford. University Press.

Greenberg, D. H, R. Meyer, Chuck Michalopoulos, and M. Wiseman. 2003. "Explaining variation in the effects of welfare-to-work programs." *Evaluation Review.* 27, 4: 359-394.

Haddix, Anne C, Steven M Teutsch, and Phaedra S Corso. 2003. *Prevention effectiveness: a guide to decision analysis and economic evaluation.* New York: Oxford University Press.

Hagan, Joseph F., Judith S. Shaw, and Paula M. Duncan. 2008. *Bright Futures: Guidelines for Health Supervision.* Elk Grove Village, IL: American Academy of Pediatrics 3:3.

Hallfors, D. and H. Cho. 2007. "Moving behavioral science from efficacy to effectiveness." *International Journal of Behavioral and Consultation Therapy, 3*, 236-250.

Homer, J. B., and G. B. Hirsch. 2006. "System dynamics modeling for public health: background and opportunities." *American Journal of Public Health* 96, 3: 452-8.

Howell, E., S. Decker, S. Hogan, A. Yemane, and J. Foster. 2010. "Declining child mortality and continuing racial disparities in the era of the Medicaid and SCHIP insurance coverage expansions." *American Journal of Public Health* 100, 12: 2500-6.

Institute of Medicine. 1990. *Medicare: A strategy for quality assurance. Vol. 1.* Washington DC:

Irwin, Lori G., Arjumand Siddiqi, and Clyde Herzman. 2007. *Early Child Development: A Powerful Equilizer.* World Health Organization.

Kaminski, J. W., L.A. Valle, J.H. Filene, and C.L. Boyle. 2008. "A Meta-analytic Review of Components Associated with Parent Training Program Effectiveness." *Journal of Abnormal Child Psychology. 36: 567-589.*

Leatherman, Sheila, and Douglas McCarthy. 2004. *The Commonwealth Fund.* "Quality of health care for children and adolescents: A chartbook." Web site:

MacDorman, Marian F., and T. J. Mathews. 2011. "Infant Deaths - United States, 2000-2007." *Mortality and Morbidity Weekly Report* 60, Supplement: 49-51.

McGlynn, E. A. 1997. "Six challenges in measuring the quality of health care." *Health Affairs.* 16, 3: 7-21.

Medical Home Initiatives for Children With Special Needs Project Advisory Committee. 2002. "The medical home." *Pediatrics* 110, pt1: 184-186.

Molnar, B.E., S.L. Buka, R.T. Brennan, J.K. Holton, and F. Earls. 2003. "A multilevel study of neighborhoods and parent-to-child physical aggression: Results from the project on human development in Chicago neighborhoods." *Child Maltreatment.* 8: 84-97.

Mowbray, C.T., M.C. Holter, G. B. Teague, and D. Bybee. 2002. "Fidelity criteria: Development, measurement, and validation." *American Journal of Evaluation.* 24: 315-

340.

Norman, C. D. 2009. "Health promotion as a systems science and practice." *Journal of Evaluation in Clinical Practice* 15: 868-872.

Paulsell, Diane, Sarah Avellar, Emily Sama Martin, and Patricia Del Grosso. 2010. *Home Visiting Evidence of Effectiveness Review: Executive Summary*. Princeton, New Jersey: Mathematica Policy Research.

Rubin, David M., Amanda L. R. O'Reilly, Xianqun Luan, Dingwei Dai, A. Russell Localio, and Cindy W. Christian. 2010. "Variation in Pregnancy Outcomes Following Statewide Implementation of a Prenatal Home Visitation Program" *Archives of Pediatrics & Adolescent Medicine.*

Seid, M., J. W. Varni, and P. S. Kurtin. 2000. "Measuring quality of care for vulnerable children: challenges and conceptualization of a pediatric outcome measure of quality." *American Journal of Medical Quality.* 15, 4: 182-8.

Stavrakos, J.C., G. Summerville, and L.E. Johnson. 2009. *Growing what works: Lessons learned from Pennsylvania's Nurse-Family Partnership Initiative*. Philadelphia, PA: Public/Private Ventures.

Stinnett, A. A., and J. Mullahy. 1998. "Net health benefits: A new framework for the analysis of uncertainty in cost-effectiveness analysis." *Medical Decision Making.* 18, 2: S68-S80.

Sweet, M. A. and M. I. Appelbaum. 2004. "Is Home Visiting an Effective Strategy? A Meta analytic Review of Home Visiting Programs for Families with Young Children." *Child Development.* 75,5:14, 35-56.

The PEW Center on the States. 2010. "Home Visiting Inventory." Web site: http://www.pewcenteronthestates.org/initiatives_detail.aspx?initiativeID=61054.

Truman, Benedict I., C. Kay Smith, Kakoli Roy, Zhuo Chen, Ramal Moonesinghe, Julia Zhu, Carol Gotway Crawford, and Stephanie Zaza. 2011. "Rationale for Regular Reporting on Health Disparities and Inequalities — United States." *Mortality and Morbidity Weekly Report.* 60, Supplement: 3-10.

Weiss, Heather B. 1993. "Home Visits: Necessary but Not Sufficient." *The Future of Children.* 3, 3: 113-128.

Wise, P. H. 2004. "The transformation of child health in the United States." *Health Affairs.* 23, 5: 9-25.

Xue, Y., T. Leventhal, J. Brooks-Gunn, and F. J. Earls. 2005. "Neighborhood residence and mental health problems of 5- to 11-year-olds." *Archives of Genetic Psychiatry.* 62: 554-563.