

Mother and Infant Home Visiting Program Evaluation (MIHOPE) Impact Analysis

September 21, 2015

The legislation authorizing the Maternal, Infant, and Early Childhood Home Visiting Program (MIECHV also known as the Federal Home Visiting Program) required the evaluation (MIHOPE) to study the effectiveness of MIECHV-funded early-childhood home visiting programs on child and parent outcomes. The evaluation covers a number of areas, including maternal and infant health, child development, child maltreatment, parenting skills, school readiness and academic achievement, crime or domestic violence, and family economic self-sufficiency.¹ In addition, the study is supposed to assess the effects of the programs on different subgroups of families and the effects of the programs on health disparities and health care costs.

This memo provides our current thinking on the estimating the effects of MIECHV-funded home visiting programs. Specifically, it proposes outcomes to include in the impact analysis and subgroups of families that will be analyzed to address the following broad research question:

- What are the effects of home visiting programs across the range of outcomes specified in the authorizing legislation, both overall and for key subgroups?

In providing information about the effects of home visiting programs on family health care use and health outcomes, the impact analysis will also help address the legislative requirement that the evaluation analyze the effects of home visiting programs on health care systems and health disparities.

The memo is organized as follows. First, we present a set of questions that we would especially appreciate Committee input on, and which we would like you to keep in mind while you are reading the memo. Second, the memo briefly describes the types of information available for analysis and the principles by which the team made decisions as to what to propose. Finally, the memo presents a set of outcomes and subgroups that will be the focus of the impact analysis (“confirmatory outcomes”). Appendixes provide additional information on how the team arrived at these sets of outcomes and subgroups and shows the types of outcomes and subgroups that will be examined in exploratory analyses (“exploratory outcomes”).

¹The legislation required MIECHV grantees (states, territories, and tribal entities) to show improvement in six specified benchmark areas. In addition, the legislation required that MIECHV-funded programs be designed to improve individual outcomes for participating families in seven areas. Because there is considerable overlap between the six benchmark areas and the seven areas of individual participant outcomes, this report uses the term “outcomes” to refer to both lists. MIHOPE is designed to assess impacts relevant to both benchmarks and participant outcomes. See Table 1 of the overview memo to the Committee for a list of benchmark areas and outcome domains.

Questions for the Committee

While we welcome the Committee members' thoughts on all aspects of the plan, we would particularly appreciate the Committee's input and advice on the following:

- We are seeking advice from the Committee on our choice of outcomes. We ask that the Committee keep in mind that adding outcomes to the list without a strong empirical reason may increase the chance that MIHOPE tells a muddled story about the effects of home visiting.
- We are interested in your thoughts on the placement of the "any health care encounter for injury or ingestion" outcome. We are wondering whether it should be included in the child health domain or the child maltreatment domain (since it is a health-related outcome, but it has also been used as a proxy for child maltreatment). Prior studies of home visiting have included a broad "injuries or ingestions" outcome that has usually been characterized as a child health outcome. However, HHS's Home Visiting Evidence of Effectiveness (HomVEE) review lists this outcome in the child maltreatment domain.
- We are also seeking input on our strategy for defining and selecting subgroups. Do you agree with our recommendations for subgroups? Given the number of subgroups and outcomes, after adjustments are made for multiple hypothesis testing it is unlikely that any subgroup differences would be statistically significant. Should the number of subgroups be further reduced? Should subgroup differences be computed only for certain outcomes? Do you have input on definitions of particular subgroups?

Data available for the MIHOPE impact analysis

MIHOPE enrolled women at the time they were identified as being eligible for a local home visiting program operating one of four national models (Early Head Start—Home Based Program Option (EHS), Healthy Families America (HFA), Nurse-Family Partnership (NFP), and Parents as Teachers (PAT)). When they entered the study, women were at least 15 years old and were either pregnant or had a child no more than six months old. At that time, a one-hour survey was conducted with each enrollee, and study staff assessed the home environment – including the quality and amount of stimulation that the child could receive in the home – using the observational portion of the Infant-Toddler Home Observation for Measurement of the Environment (IT-HOME).

Information on outcomes will be available from a number of data sources around the time the child is 15 months old. These data sources include the following:

1. A **one-hour telephone interview** with the child's primary caregiver that asks about outcomes in all the domains mentioned in the authorizing legislation other than school readiness and academic achievement.
2. A **video-recording of an interaction between the child and caregiver** (the biological mother in the vast majority of cases so far) using the "Three Bags" and "Clean-up" tasks.

3. The **Preschool Language Scales, Fifth Edition (PLS-5), Auditory Comprehension scale**, to assess the child's ability to be attentive and respond to stimuli in the environment and to comprehend basic vocabulary or gestures.
4. The **child's weight and height** to provide information about whether the child's growth is within a normal range or the child exhibits early signs of being underweight or obese. In addition, we are measuring the mother's weight to assess the effects of home visiting on maternal weight and obesity.
5. The **IT-HOME** collected again at follow-up, including both observational and parent-reported measures.
6. **Administrative data** in four areas: (1) birth outcomes from children's birth certificates, (2) health care use from Medicaid and Children's Health Insurance Program data, (3) child maltreatment from state child welfare data, and (4) employment and earnings from the National Directory of New Hires.

The impact analysis will compare mean outcomes derived from these measures for the program and control groups, adjusting for baseline family characteristics to increase the statistical precision of the impact estimates. The main question for the subgroup analysis will be whether estimated effects are larger or smaller for one set of subgroups than another. Questions about the effects of actually participating in home visiting (rather than merely being assigned to the program group), dosage (that is, the frequency and intensity of services received), and the relationship between program features and impacts are addressed in a separate memo on analyzing impact variation.

Confirmatory Outcomes

Framework for choosing outcomes for the impact analysis

The wide range of information being collected about families at follow-up allows for the examination of many dozens of outcomes. Over the past few years, however, evaluation methodologists have begun to recommend focusing impact studies on a more streamlined set of outcomes in order to reduce the chances of a “false positive” finding in which an intervention with no true effect produces statistically significant impacts on at least one outcome.² The Institute of Education Sciences (IES) calls this streamlined set of outcomes “confirmatory outcomes”.

We propose to use such an approach in MIHOPE. A relatively streamlined set of pre-specified confirmatory outcomes will be shown in an Executive Summary table and used to assess whether MIECHV-funded home visiting programs were effective at improving outcomes for families. In order to reduce the chance of a “false positive” finding — that is, a statistically significant impact when the true effect is zero — the study would use the Westfall-Young

²See, for example, Schochet (2008).

stepdown method to assess whether the pattern of impacts across the confirmatory outcomes is likely to be due to chance.³ The method proposed for MIHOPE would be similar to that used by other studies conducted for OPRE.

Other outcomes would be considered exploratory (again, following the terminology used in IES studies) and would be shown in tables in the body of the report but not in Executive Summary tables. These outcomes could be discussed in the text of the Executive Summary if they helped answer an important policy or research question or shed light on findings related to the confirmatory outcomes.

Since the confirmatory outcomes would provide succinct information on whether MIECHV-funded programs affected family outcomes, it is important to choose outcomes that are sensitive to the intervention of home visiting but to also include outcomes that could inform policy decisions. For this reason, possible outcomes were assessed using three criteria: (1) whether home visiting is likely to show effects on the outcome, (2) whether the outcome addresses an important issue related to home visiting policy, and (3) whether the outcome was measured in way to minimize bias (for example, through observation or administrative records).

The first criterion was addressed by looking at evidence of effectiveness in prior studies of the four national models. If prior studies somewhat consistently found effects for families with young children on a particular outcome, we were more likely to include it on our list of confirmatory outcomes. By contrast, if prior studies had examined an outcome but rarely found statistically significant effects, it was unlikely to be proposed as a confirmatory outcome. We believe this approach makes sense because home visiting programs aim to improve a wide range of outcomes, and prior evidence provides useful guidance on which family outcomes are most likely to be influenced by home visits.

Although the literature presents a rich set of information about where prior studies of home visiting have found effects, there may be reasons to expect effects under MIECHV for other outcomes. In particular, earlier studies were conducted as far back as several decades ago, and MIECHV-funded programs may be placing special emphasis on some areas precisely because prior studies had not found effects. Thus, as noted below, several outcomes are being proposed as confirmatory despite no clear evidence of effects from prior studies.

Policy and research relevance also factored into our decisions. For example, one principal parenting outcome was chosen because it is a strong predictor of subsequent child development, even though only one prior study of home visiting had examined the specific measure being used and it did not find a statistically significant estimated effect. Since there will not be many measures of child development when children are 15 months old, this outcome can provide important information about whether the programs are likely to have positive effects on child development in the future.

³Although choosing a streamlined set of confirmatory outcomes reduces the chance of a false positive finding compared to having a larger set of outcomes, an adjustment for multiples comparisons is proposed because the chance of a false positive finding would still exceed the critical value set for an individual outcome (typically 5 or 10 percent). For a discussion of the general problem of multiple comparisons, see Westfall, Tobias, and Wolfinger (2011).

Finally, the way the outcome was measured played a role in choosing confirmatory outcomes. In particular, preference was given to measures that were based on observations (the Three Bags task and the IT-HOME) or direct assessments (the PLS-5) since observations and direct assessments may provide more objective ratings of parenting than parent reports.

Proposed confirmatory outcomes

Table 1 shows the list of proposed confirmatory outcomes, each of which is in a MIECHV benchmark domain. The table also briefly describes information regarding the three criteria used to choose the list of confirmatory outcomes. The table's second column provides an indicator of prior evidence of effectiveness, namely the number of statistically significant estimates and total number of estimates in prior studies in families with children under two years old. The third column provides an indicator of the quality of measurement. Outcomes measured through observations, direct assessments, or administrative records are considered good. Those measured through parent reports from surveys are considered adequate. The fourth column shows how many of the four national models indicated to the MIHOPE team that the outcome was a high priority.⁴ The final column shows the source of information for the outcome.

Following is a brief discussion of each outcome and why it was chosen for the list of confirmatory outcomes, organized by domain. More information about how these decisions were made along with a list of exploratory outcomes is shown in the appendixes, which are organized by domain.⁵

Maternal and child health

Improving maternal health and child health are goals of many home visiting programs. All four national models aim to improve child health and well-being, but vary in their explicit emphasis on maternal health. One of NFP's key goals is to improve infant health by improving maternal preventive care and prenatal practices. HFA's goals include promoting child safety, encouraging healthy parent-child interactions, and increasing access to prenatal care. Both EHS and PAT emphasize promoting child health and development by providing parents with relevant knowledge, resources, and tools.⁶

In addition to achieving the goals set out by the national models they follow, local MIECHV programs are expected to collect and report to the state data on a number of health indicators. For example, in the benchmark area of maternal and newborn health, states are reporting data from local programs on prenatal care; parental use of alcohol, tobacco, or illicit drugs; inter-birth intervals; screening for maternal depressive symptoms; insurance status of the

⁴Michalopoulos et al. (2015).

⁵Appendix A provides information about maternal and child health, Appendix B about child development, Appendix C about child maltreatment, Appendix D about parenting, Appendix E about domestic violence, and Appendix F about family economic self-sufficiency.

⁶Information on each of the four national models was taken from the HomVEE website: <http://homvee.acf.hhs.gov/programs.aspx> and national model websites for PAT (<http://www.parentsasteachers.org/>), NFP (<http://www.nursefamilypartnership.org/>), HFA (<http://www.healthyfamiliesamerica.org/home/index.shtml>), and EHS (<http://eclkc.ohs.acf.hhs.gov/hslc/tta-system/ehsnrc>).

mother and child; and well-child visits. The reduction of emergency department (ED) visits for both mother and child is also included in a benchmark area along with reductions in child injuries and child maltreatment.

Table 1
Proposed Confirmatory Outcomes

Domain and outcome	Criteria considered for selection			
	Prior evidence of favorable findings ^a	Measurement quality	Number of models rating high priority	Source ^b
<u>Maternal and child health</u>				
New pregnancy after study entry	Yes (2 of 9)	Adequate	1	FFI
Mother has health insurance coverage	No (0 of 1)	Adequate	1	FFI
Number of well-child visits	Limited (1 of 14)	Adequate-Good	4	MD, FFI
Child has health insurance coverage	Yes (2 of 4)	Adequate-Good	4	MD, FFI
Any child ED use	Limited (2 of 12)	Adequate-Good	4	MD, FFI
<u>Child development</u>				
Behavior problems total score	Yes (4 of 14)	Adequate	4	FFI (BITSEA)
Language skills in normal range	Yes (4 of 23)	Good	4	PLS-5
<u>Child maltreatment</u>				
Frequency of minor physical assault	Limited (3 of 12)	Adequate	4	FFI (CTSPC)
Frequency of psychological aggression	Limited (2 of 14)	Adequate	4	FFI (CTSPC)
Any health care encounter for injury or ingestion	Yes (4 of 13)	Adequate-Good	4	MD, FFI
<u>Parenting</u>				
Quality of home environment	Yes (4 of 13)	Good	4	IT-HOME
Parental supportiveness	No (0 of 8)	Good	4	Three Bags
<u>Family economic self-sufficiency</u>				
Receiving education or training	Yes (9 of 25)	Adequate	4	FFI

SOURCE: U.S. Department of Health and Human Services website HomVEE: <http://homvee.acf.hhs.gov/>.

NOTES: ^aThe evidence base presented in this table includes only findings for the four MIHOPE models and for children 2 years of age and younger.

^bFFI = Family follow-up interview; MD = Medicaid files; BITSEA = Brief Infant Toddler Social Emotional Assessment; PLS-5 = Preschool Language Scale, Fifth Edition; CTSPC = Conflict Tactics Scale — Parent-Child version; IT-HOME = Infant-Toddler Home Observation for Measurement of the Environment.

Finally, since all MIHOPE sample members are low socioeconomic status, improvements in the outcomes presented here as well as the exploratory outcomes described in Appendix A to this memo will provide an indication that the programs reduced health care and

health disparities. These results will thus help address the legislative requirement that the evaluation study the potential of MIECHV to affect the health care system.

With that background, the team proposes five measures of maternal and child health:

- New pregnancy after study entry. Prior studies of home visiting programs have found impacts on maternal reproductive health, including reductions in the number of subsequent pregnancies as well as the time interval between subsequent pregnancies (birth spacing), with potentially long-term implications for family self-sufficiency and health care system expenditures. Specifically, both the Memphis and Denver NFP evaluations found reduced rates of mothers giving birth within 24 months of their previous birth (with effect sizes in the 0.30 range). Other evaluations of EHS, HFA, and PAT did not find significant effects on this outcome, however.

Although effects on birth spacing have only been found in evaluations of one of the four national models in MIHOPE thus far, other evidence-based home visiting programs, such as Oklahoma's Community-based Family Resources and Support Program and Mind the Baby, have found statistically significant impacts on pregnancy intervals.⁷ Moreover, the national models and local programs in MIHOPE generally place a high priority on improving family planning and birth spacing. For example, NFP ranked this outcome as a high priority and both EHS and HFA ranked it as a moderate priority. Even though PAT did not rate it as a high priority at the national level, about half of all local PAT programs in MIHOPE and about 30 percent of local EHS and HFA programs reported increasing their emphasis on family planning and birth spacing since MIECHV began.⁸

Our proposed measure of reproductive health is rapid repeat pregnancy, which we define as a subsequent pregnancy occurring between baseline and follow-up, using mothers' self-reports from the surveys.

- Mother has health insurance coverage. Insurance coverage for the mother is included as a benchmark indicator of maternal health, and provides a measure of the mother's access to and stability of insurance coverage for herself. This outcome is not commonly examined. Only one study of the four national models has examined this outcome and it did not find a statistically significant effect. However, maternal eligibility for Medicaid benefits has increased under the Affordable Care Act so programs may be better able to influence this outcome if they are able to help parents apply or reapply for benefits. Finally, insurance coverage is a gateway to accessing affordable care,⁹ so this outcome may be a predictor of later maternal health outcomes.

Despite the limited research on this outcome in prior home visiting studies, the MIHOPE Report to Congress found that 75 percent of home visitors have reported that their

⁷Culp, Culp, Anderson, and Carter (2007); Sadler et al. (2013).

⁸Michalopoulos et al. (2015).

⁹Centers for Disease Control and Prevention (2012); Majerol, Newkirk, and Garfield (2014); Newacheck et al. (1998).

program expects them to help mothers gain or maintain access to health care coverage.¹⁰ Moreover, this expectation was fairly consistent across the four national models.

We thus recommend including health insurance coverage at follow-up as a confirmatory outcome, using a combination of self-reports from surveys and maternal Medicaid records when available. We suggest creating a binary indicator of whether the mother has any coverage versus no coverage, since we are less interested in the type of coverage (for example, public or private) than we are in the presence of coverage.

- Number of well-child health care visits. As indicated by its inclusion in the MIECHV benchmarks, home visiting programs can encourage parents to maintain a check-up schedule for their child, which ensures that a health care provider is monitoring healthy growth and development. In addition, as noted above, all four national models place a high priority on infant or child health and promoting child preventive care. Nevertheless, the evidence from the four national models is weak on this outcome: of the four national models, only HFA has been found to increase the number of well-child visits among children under the age of 3 (although studies of Early Start in New Zealand and Healthy Steps in the United States have also found improvements in this outcome).¹¹ However, prior home visiting studies suggest that maintaining adequate well-child visits is a parental health practice where there is room for improvement, and this may be an area of greater focus under MIECHV.¹² In fact, almost all home visitors (94 percent) surveyed in MIHOPE around the time of the Report to Congress reported that they are expected to make sure children are up-to-date on shots and well-child care.

In MIHOPE, the number of well-child visits will be measured using a combination of Medicaid files (using specific encounter or claims codes for well-child or Early Periodic Screening, Diagnosis, and Treatment [EPSDT] visits), and parental reports in the follow-up survey (if the child is not on Medicaid or is on Medicaid managed care in a state without high-quality managed care data). Depending on the proportion of children who do not have Medicaid records, we may need to use a binary indicator for this outcome (for example, at least three visits in the first 15 months), since the follow-up survey asks about whether the child had a well-child visit at three points in time (that is, first week, at 12 months, and at 15 months).

- Child has health insurance coverage. Although not often examined, the evidence suggests that home visiting programs may be able to help families obtain insurance coverage for their children. Insurance coverage ensures that health care services are affordable for families, and should, in principle, increase the use of preventive health care services and screenings.¹³ The MIHOPE follow-up survey asks parents to report insurance coverage for their child, but Medicaid records also provide some information on child health care coverage.

¹⁰Michalopoulos et al. (2015).

¹¹Fergusson, Horwood, Grant, and Ridder (2005); Guyer et al. (2003).

¹²Duggan et al. (1999).

¹³McMorrow, Kenney and Goin (2014); Yu et al. (2002).

- Any child ED use. Home visiting, by encouraging regular preventive care or connecting families to a medical home, may decrease the number of ED visits.¹⁴ In addition, home visiting may reduce injuries due to child maltreatment that might require ED visits. At the same time, home visiting may encourage families to use ED services for many reasons, including if they reside in medically underserved areas where pediatric care is lacking. Perhaps because of these counteracting forces, the studies reviewed by HomVEE do not generally show significant changes in this outcome, although the Elmira NFP trial showed decreases in the number of ED visits in the child's first and second years of life.¹⁵ Despite the weak prior evidence, it is an important measure for policy making purposes because receiving care in an ED setting is among the most expensive of health care interactions and often leads to even more expensive hospitalizations. While some ED use is expected and unavoidable, prior research shows that low-income parents of young children are more likely to use EDs for conditions that are not urgent or are potentially avoidable. An analysis of California hospital data revealed that 60 percent of ED visits for Medicaid-covered children under 1 year of age were avoidable.¹⁶ This suggests that home visiting programs may be well-positioned to address an important issue by providing information and education on appropriate ED use (in addition to connecting families to a medical home).

In MIHOPE, we propose to capture this outcome by examining whether the child has been to the ED at least once from birth to follow-up, using both Medicaid records and follow-up survey reports.

Child development

Child development is another outcome area of importance to both MIECHV and the four national models. In the MIECHV benchmark area of improvements in school readiness and achievement, states are expected to report annually on the following dimensions of child development: (1) child's communication, language, and emergent literacy; (2) child's general cognitive skills; (3) child's positive approaches to learning, including attention; and (4) child's social behavior, emotion regulation, and emotional well-being.

Each of the four home visiting models places a high priority on child development, although there is some variation in the areas they target. EHS focuses squarely on enhancing child development while strengthening families.¹⁷ HFA seeks to promote child health and development.¹⁸ One of NFP's three broad goals is to improve children's health and development.¹⁹ Finally, PAT seeks to provide early detection of developmental delays and health issues, prevent child abuse and neglect, and increase school readiness.²⁰

¹⁴American College of Physicians (2014) provides one definition of a medical home: a "care delivery model in which patient treatment is coordinated through a primary care physician to ensure they receive the necessary care, where they need it, and in a manner they can understand."

¹⁵Olds, Henderson, Chamberlin, and Tatelbaum (1986).

¹⁶McConville and Lee (2008).

¹⁷Love et al. (2001).

¹⁸Mitchell-Herzfeld et al. (2005).

¹⁹Olds et al. (2004).

²⁰Wagner and Spiker (2001).

As Table 1 shows, we propose two principal child development outcomes: behavior problems (total score) and language skills in the normal range. Although there is only thin evidence that home visiting programs have affected child development in the child's first two years, the existence of some evidence of favorable impacts combined with the strong focus of all four national models on this area suggest the importance of assessing MIECHV programs in this domain.

- Behavior problems total score. Behavior problems include behaviors that are part of children's typical development (for example, aggression, sadness, and fear) but become problematic when the frequency or intensity with which they occur is much higher or lower than expected. Behavior problems also include deviant behaviors that are never developmentally appropriate, such as those seen in children with developmental disorders (for example, self-injurious behaviors). Behavior problems are typically characterized along two dimensions: *externalizing problems*, which include aggression, acting out, and hyperactivity, and *internalizing problems*, which include anxiety, sadness, and social withdrawal. Having behavior problems in early childhood is a risk factor for mental health problems and academic difficulties throughout childhood and into adulthood.²¹

Each of the four national models has been assessed on its ability to reduce children's behavior problems, with some evidence of positive effects of HFA and NFP but no statistically significant impacts of EHS or PAT.

Behavior problems will be measured using the parent-reported Brief Infant-Toddler Social Emotional Assessment (BITSEA). Although none of the studies included in the HomVEE review used the parent-reported BITSEA, three of the four favorable findings are based on the Child Behavioral Checklist (CBCL). The BITSEA has been validated using the CBCL in a sample of 2- and 3-year-olds, suggesting that the two measures are comparable.²² We propose using the raw behavior problems subscale from the BITSEA based on guidance from Margaret Briggs-Gowan – the developer of the measure – and team members' prior experience that suggests the subscale can be sensitive to program effects.²³

- Language skills in normal range. Children's early language development has been linked to later cognitive and language outcomes, as well as school readiness and later achievement,²⁴ and HFA and NFP have shown some improvements in children's cognitive or language development at an early age.²⁵ As noted earlier, follow-up data collection includes a direct assessment of children's receptive language skills through the Preschool Language Scales, Fifth Edition, Auditory Comprehension Scale (PLS-5). Using the PLS-5, we propose a binary measure that would indicate the proportion of

²¹Broidy et al. (2003); McClelland, Morrison, and Holmes (2000).

²²Briggs-Gowan et al. (2004).

²³Hsueh, Jacobs, and Farrell (2011).

²⁴Prior, Bavin, and Ong (2011).

²⁵Caldera et al. (2007); Landsverk et al. (2002); Olds et al. (2002).

children who have language skills in the normal range. We propose a binary measure because prior evidence indicates that a binary measure is sensitive to program impacts.²⁶

Child maltreatment

The authorizing legislation specifies that outcomes for individual families should include improvements in the prevention of child injuries and maltreatment. In accordance with that stipulation, states are expected to report data on five indicators specifically related to child injuries, child abuse, neglect, or maltreatment as part of their annual MIECHV benchmark reporting.²⁷ The five child injury and maltreatment indicators in the benchmarks are:

1. Information provided or training on prevention of child injuries
2. Incidence of child injuries requiring medical treatment
3. Reported suspected maltreatment for children in the program
4. Reported substantiated maltreatment
5. First-time victims of maltreatment

In addition to being included in the MIECHV benchmarks and participant outcomes, preventing child maltreatment is a specified goal or outcome of many home visiting programs. Three of the four national models being studied in MIHOPE name the prevention of child maltreatment as their mission or as a program outcome, but vary in the mechanisms they use to achieve this goal. One of the short-term outcomes of NFP is to decrease child injuries or ingestions and verified cases of child abuse and neglect by helping parents to provide sensitive and competent caregiving. HFA's mission is to promote child well-being and prevent the abuse and neglect of children by strengthening parent-child relationships, promoting childhood growth and development, and enhancing family functioning by reducing risk and building protective factors such as resources and coping strategies that promote effective parenting under stress. One of the stated outcomes of PAT is to prevent child abuse and neglect by equipping families with the information and tools necessary to provide a healthy and safe home using protective factors to strengthen families. EHS does not identify the reduction or prevention of child maltreatment as an explicit program outcome, but a cornerstone of the program is to support the healthy development (for example, physical and emotional) of each child through parent education and the support of positive parent-child relationships, and EHS at the national level has indicated to MIHOPE that they consider reductions in child maltreatment to be a high priority for the model.

Because of the strong focus of MIECHV and the four national models on reductions in child maltreatment, we recommend including three principal child maltreatment outcomes based on parent reports.²⁸ For each outcome, prior studies of home visiting have shown favorable impacts.

²⁶Olds et al. (2002).

²⁷These indicators are included under Benchmark 2, which also encompasses child and maternal emergency department visits for all causes, measures that we currently address in the health domain.

²⁸Although the study is collecting state administrative data on child maltreatment, administrative data generally show few incidences of child maltreatment for children under two years old. The low incidence of formal reports makes it unlikely that estimated impacts on administrative data reports would be statistically significant. In addition, state administrative data may suffer from surveillance bias, wherein program group families might be more likely to be reported to the state because home visitors are frequently seeing those families. As discussed in Appendix C, the report will include an exploratory analysis of child welfare data.

These behaviors may also be direct targets of home visiting programs across all program models, and may be precursors to more serious behaviors that can manifest later in the child's life.²⁹

The three proposed confirmatory outcomes are:

- Frequency of minor physical assault.³⁰ This is a measure of behaviors such as spanking, pinching, or hitting on the bottom with a hard object, as derived from parent reports from the Conflict Tactics Scale – Parent-Child version (CTSPC). Minor physical assault has only been examined in studies of HFA. Of the three statistically significant impacts found, two were on the frequency of minor physical assault and one was on a measure of corporal punishment.³¹
- Frequency of psychological aggression. This measure covers behaviors such as yelling, screaming, or swearing at the child, or calling the child names. It is also derived from the CTSPC. Psychological aggression has been commonly examined in studies of HFA, and statistically significant improvements in psychological aggression have been found in two such studies.³²
- Health care encounter for injuries and ingestions. Home visiting studies have examined several different outcomes related to health care encounters for injuries or ingestions. These include broader outcomes such as any injury requiring medical care and any health care encounter for injury or ingestion. They also include more specific outcomes such as hospitalization for injury or ingestion, ED visit for injury or ingestion, and outpatient visits for injury or ingestion. Studies of both NFP and HFA have examined impacts on health care encounters for injuries and ingestions, and four favorable impact estimates have been found, all for NFP.³³ Because hospitalizations or emergency room visits for injuries would likely be rare events in a sample this young, we propose to use Medicaid claims data and parent reports to define this outcome as any type of medical care received to treat an injury or ingestion.

Parenting

Parenting is one of the primary focuses of both MIECHV and the four national models. In the MIECHV benchmark area focused on improvements in school readiness and achievement, states are expected to report annually on the following dimensions of parenting: (1) parent support for children's learning and development, (2) parent knowledge of child development and of their child's developmental progress, (3) parenting behaviors and parent-child relationship, and (4) parent emotional well-being or parenting stress.

²⁹Lee, Grogan-Kaylor, and Berger (2014).

³⁰Because more studies using the CTS have found favorable effects on frequency rather than prevalence, and because these behaviors may be relatively common, we propose to analyze frequency of minor physical assault and psychological aggression.

³¹Caldera et al. (2007); Duggan et al. (2007); DuMont et al. (2008).

³²Caldera et al. (2007); DuMont et al. (2008).

³³Kitzman et al. (1997); Mitchell-Herzfeld et al. (2005); Olds, Henderson, Chamberlin, and Tatelbaum (1986).

In addition, each of the four national models included in MIHOPE targets at least one aspect of parenting. For example, one of the main goals of EHS is to support the development of close, supportive relationships between parents and their children, both because these relationships are important in their own right and because they lay the groundwork for enhancing children's development.³⁴ By comparison, HFA seeks to enhance parent-child interactions and promote child health and development.³⁵ One of NFP's three broad goals is to improve children's health and development by helping parents provide more competent care.³⁶ Finally, PAT emphasizes parent behavior as the mechanism through which the program benefits children and teaches parents about good parenting practices and principles of child development.³⁷ Given these emphases of the national models, it is not surprising that studies of all four models have found improvements in parenting practices. In addition, each of the four national models reported to the MIHOPE team that they place a high priority on parenting behavior.³⁸

Because of the prior evidence and the focus of the national models and MIECHV on parenting, we consider parenting to be one of the most important areas for the impact analysis to examine. We therefore propose two observational summary measures as confirmatory outcomes.

- Quality of home environment. This is the total score derived from the Infant-Toddler Home Observation for Measurement of the Environment (IT-HOME), which assesses the general quality of the home environment as well as the quality and amount of stimulation that the child receives in the home. The measure provides information on several aspects of parenting behavior, including social-emotional responsiveness, cognitive stimulation, and harsh parenting.

We propose this indicator as a confirmatory outcome for several reasons. First, studies of all four national models have found statistically significant effects on the IT-HOME total score.³⁹ In addition, the IT-HOME total score has been linked to children's cognitive development in the first two years of life, suggesting a possible pathway through which home visiting programs might influence child development.⁴⁰ Finally, the IT-HOME is appealing because it is largely an observational measure.

There is less consistent evidence of effectiveness for any of the IT-HOME subscales, which could include areas such parental warmth, parental verbal skills, parental lack of hostility, learning and literacy, home interior, developmental advance, and activities and

³⁴Love et al. (2001).

³⁵Mitchell-Herzfeld et al. (2005).

³⁶Olds et al. (2004).

³⁷Wagner and Spiker (2001).

³⁸Michalopoulos et al. (2015).

³⁹The IT-HOME is a version of the HOME used for families with infants and toddlers. The HomVEE review, which was this memo's main source on evidence of the prior effects of home visiting, did not always indicate which version of the HOME was used in a study. Since this memo presents evidence from HomVEE only for children 2 years of age and younger, the memo assumes that the studies cited had used the IT-HOME.

⁴⁰Bradley et al. (1989); Holditch-Davis et al. (2000); Totsika and Sylva (2004).

outings.⁴¹ For that reason, we propose presenting the IT-HOME subscales as exploratory outcomes, as discussed in Appendix D.

- Parental supportiveness. Parental supportiveness will be measured through the Three Bags task, which is a video-recorded interaction between the primary caregiver and her/his child. In the task, participants are given three bags of objects, such as board books and building blocks. Parents are asked to use these toys to play with their child for about 10 minutes. The task and various adaptations of the task have been successfully administered and coded in a variety of large-scale experimental and longitudinal studies of toddlers, including the National Evaluation of Early Head Start,⁴² the Early Childhood Longitudinal Study—Birth Cohort (ECLS-B), and Baby FACES. The task has been used with children as young as 14 months of age.⁴³ In addition to its use in studying EHS, the parental supportiveness measure and many of the Three Bags scales are conceptually similar to dimensions of parenting that have been examined in studies of other models reviewed by HomVEE, such as maternal teaching, as measured by the observational Nursing Child Assessment Satellite Training (NCAST).⁴⁴

The proposed measure of parental supportiveness will be a composite of three scales of the Three Bags coding system that capture positive dimensions of parenting: parental sensitivity, parental positive regard, and parental stimulation of cognitive development.⁴⁵ This approach is consistent with the way in which parenting has been assessed in prior studies that used the Three Bags task, including the EHS national evaluation.

This measure is being proposed for several reasons. First, the EHS study found a marginally statistically significant impact estimate ($p < 0.10$) with an effect size of 0.14, which is reasonably large for this type of measure and intervention and large enough to produce statistically significant impact estimates in MIHOPE. Second, parental supportiveness has been linked to children’s cognitive and social-emotional development in the first two years of life.⁴⁶ Finally, as an observational measure, the Three Bags Task is less subject to bias than self-reported parenting measures.

School readiness and academic achievement

Although this is one of the outcome areas identified in the authorizing legislation, no measures are available for children under two years old. It is thus not an outcome area in the MIHOPE final report.

⁴¹These subscales are proposed by Linver, Martin, and Brooks-Gunn (2004). Other researchers using the IT-HOME have used a different set of subscales. The team may undertake measurement work to determine whether one set of subscales or another is more consistent with the data that are collected for MIHOPE.

⁴²Brady-Smith et al. (2000).

⁴³Love et al. (2001).

⁴⁴Kitzman et al. (1997).

⁴⁵If measurement work does not support the creation of this composite measure, we will consider raising to confirmatory status some or all of the individual scales.

⁴⁶Cabrera, Fagan, Wight, and Schadler (2011); Ryan, Martin, and Brooks-Gunn (2006).

Crime or domestic violence

We propose no confirmatory outcomes from the area that includes crime or domestic violence because there is little evidence from prior studies that home visiting programs affect these outcomes at such an early stage. With regard to domestic violence, HomVEE indicates that studies of HFA and NFP have provided 41 estimates of the effects of the two models on domestic violence, but only two of those estimates have been statistically significant, and neither one was for children age 2 and younger. This area may be examined in longer-term follow-up of the MIHOPE sample. With regard to crime, prior evidence from home visiting has found effects on the children's criminal activity when they are older, but not on the caregiver's criminal activity.

Family economic self-sufficiency

All four national models place a high priority on improving family economic self-sufficiency.⁴⁷ In addition, the MIECHV benchmarks include improving outcomes in the areas of household income and benefits, and employment or education of adults. Based on evidence from prior studies of home visiting, we recommend one confirmatory outcome in the area of economic self-sufficiency.

- Currently receiving education or training. This outcome may be a precursor to subsequent improvements in employment, earnings, and income, and it has shown fairly consistently positive findings for mothers of young children in prior studies of home visiting.^{48,49 50,51} Moreover, other possible measures of economic self-sufficiency, such as employment and earnings, have been studied extensively but no effects by home visiting programs for mothers with children under 2 years old have been observed.

Confirmatory Subgroups

Framework for selecting subgroups for the impact analysis

In analyzing impacts of home visiting for subgroups of families, the study also proposes a distinction between confirmatory and exploratory subgroups. As for outcomes, the confirmatory subgroup analysis will include statistical adjustments for multiple hypothesis tests. Thus, we aim to be conservative in specifying the number of subgroup-outcome combinations designated as confirmatory. Table 2 summarizes our proposed distinctions between confirmatory or exploratory subgroup-outcome combinations.

⁴⁷Michalopoulos et al. (2015).

⁴⁸Love et al. (2002).

⁴⁹Love et al. (2001).

⁵⁰Landsverk et al. (2002).

⁵¹LeCroy and Krysik (2011).

We used three different lenses to identify the most relevant subgroups for MIHOPE: (1) subgroup constructs used in prior home visiting research, (2) policy relevance, and (3) national model relevance.

Lens 1: Subgroup constructs used in prior research. We focused on subgroups used in studies with follow-up periods of five years or fewer. For each of the national models included in MIHOPE, we reviewed studies included in the HomVEE review that were rated as “high” or “moderate.”⁵² We also conducted a literature search focused on moderators of home visiting program impacts, to ensure we collected information on other subgroup analyses as well. Subgroup constructs in the studies we reviewed were numerous and fell into three general categories (though a particular characteristic might fall under multiple categories):

Demographic characteristics of the mother:

- maternal age
- partner status (for example, whether the mother was unmarried, whether the mother lived with the child’s father)
- whether the mother had other children
- race and ethnicity
- primary language
- socioeconomic status
- education level
- employment status

Physical or mental health of the mother:

- depression
- attachment style
- psychological resources
- family stress
- whether the mother had given birth at the time of program enrollment
- substance abuse

Domestic violence or child abuse:

- domestic violence (in the current home or in the mother’s home when she was a child)
- child abuse or neglect (in the current home or in the mother’s home when she was a child)

⁵²Subgroup findings for an outcome domain were included in the HomVEE review only if they were replicated in at least two different studies using different analytic samples and if only one type of family characteristic subgroup was identified in HomVEE.

Table 2
Distinctions Between Confirmatory and Exploratory Subgroup Impact Estimates

Criteria	Confirmatory subgroup impact estimate	Exploratory subgroup impact estimate
Statistical tests of subgroup impact differences <i>will be adjusted for multiple comparisons</i>	✓	
The subgroup construct, definition of its subgroups, and confirmatory outcomes for which impacts will be estimated must be <i>specified in advance</i> of any impact estimation	✓	
Each subgroup must have <i>sufficient sample size</i> for analysis (approximately 1,000 observations per subgroup)	✓	
Subgroup impact estimates will be reported in the <i>main report</i>	✓	
The subgroup construct has an <i>empirical base</i> with regard to at least one primary outcome in the MIHOPE impact analysis	✓	
The subgroup construct is of <i>policy or program significance</i>	✓	✓

Our review indicated that no common set of subgroup constructs or particular outcomes were examined across studies. Some studies defined subgroups based on one of these dimensions (for example, “teenage mother”) while others defined subgroups based on composite dimensions (for example, “poor, unmarried teenage mother”). Some studies reported subgroup impacts whether or not they were statistically significant, while other studies reported only statistically significant results. Further, studies tended to test whether subgroup impacts were statistically significantly different from zero, not whether subgroup impacts were statistically significantly different from each other.

Lens 2: Policy relevance. Our working definition of “policy-relevance” emphasizes individual or family characteristics that are of particular interest to policymakers and interest groups focused on home visiting and child well-being. The primary source we consulted for policy relevance was the authorizing legislation of MIECHV: Section 2951 of the Affordable Care Act, in particular the subsection that specifies characteristics of “high-risk populations.”⁵³

⁵³The high risk populations identified in authorizing legislation are as follows: (A) Eligible families who reside in communities in need of such services; (B) Low-income eligible families; (C) Eligible families who are pregnant women who have not attained age 21; (D) Eligible families that have a history of child abuse or neglect or have had interactions with child welfare services; (E) Eligible families that have a history of substance abuse or need substance abuse treatment; (F) Eligible families that have users of tobacco products; in the home; (G) Eligible families that are or have children with low student achievement; (H) Eligible families with children with developmental delays or disabilities; (I) Eligible families who, or that include individuals who, are serving or formerly served in the Armed Forces, including such families that have members of the Armed Forces who have had multiple deployments outside of the United States.

We also consulted websites (primarily policy priorities sections) of a number of organizations concerned with child well-being such as Zero to Three and the American Academy of Pediatrics; the characteristics we identified through these organizations were broadly consistent with those identified in the authorizing legislation.

Lens 3: National model relevance. We gathered information on the extent to which each of the four national models emphasized outcomes or target populations that suggested a focus on family-level subgroups (Appendix G). Our primary information sources were the MIHOPE Report to Congress and the HomVEE review. Not surprisingly, most national models emphasized maternal and child health and development. Other areas such as domestic violence were emphasized to varying degrees by the national models.

Proposed confirmatory family subgroups

Based on our review using these three lenses, and applying the criteria described in Table 2, the team is proposing six confirmatory subgroups, listed briefly here, and described more fully in Table 3 and below:⁵⁴

1. Pregnancy status at time of enrollment
2. Whether or not the mother is a first-time mother
3. Whether or not the mother has low psychological resources
4. Maternal attachment style and depression (explained below)
5. Presence of domestic violence
6. Whether or not the mother exhibits high demographic risk (explained below)

Each of the subgroups is described briefly below:

- Pregnancy status at time of enrollment. The subgroups in this construct include: enrolled up to 28th week of pregnancy, enrolled after 28th week of pregnancy, and enrolled after baby's birth. Some confounding with national model will result from this subgroup: NFP only enrolls mothers before the 28th week of pregnancy. The other three national models in the study allow (but do not require) participants to enroll during pregnancy, but they do not specify a particular point during the pregnancy when mothers can enroll and local programs can set specific eligibility criteria. The Report to Congress found that the majority of women participating in MIHOPE were pregnant at the time they enrolled in the program, even for EHS, HFA, and PAT. The authorizing legislation reflects part of this construct when it emphasizes "pregnant women who have not attained age 21."

⁵⁴Appendix G provides further detail on the team's process of identifying and selecting potential subgroups for confirmatory and exploratory analyses, as well as more detail on each of the characteristics and their measurement.

Table 3**Proposed Confirmatory Subgroup Constructs**

Confirmatory subgroup construct	Dimensions of subgroup construct (measured at baseline)	Policy significance (from ACA authorizing legislation)	National program model significance (theory, goals, or target population)
<u>Pregnancy status at time of enrollment</u>			
Type: Ordinal (enrolled before 28th week of pregnancy, enrolled after 28th week of pregnancy, enrolled after birth of child)	Pregnancy status at time of enrollment	Pregnant women who have not attained age 21.	HFA, NFP, EHS, PAT
<u>Whether the mother is a first-time mother</u>			
Type: Binary (first-time mother, not first-time mother at enrollment)	Number of children at time of enrollment		NFP
<u>Whether the mother has low psychological resources</u>			
Type: Binary (high or low psychological resources)	Intelligence (WAIS); Mental health (CES-D and GAD-7); Sense of mastery (Pearlin Mastery Scale)		
<u>Maternal attachment style and depression</u>			
Type: Binary (whether mothers are both depressed and exhibit relationship insecurity)	Attachment Style Questionnaire (ASQ); Center for Epidemiologic Studies Depression Scale (CES-D)		HFA, NFP
<u>Presence of domestic violence</u>			
Type: Binary (presence of any physical violence toward the mother or perpetrated by the mother)	Conflict Tactics Scale (CTS)		HFA

(continued)

Table 3 (continued)

Confirmatory subgroup construct	Dimensions of subgroup construct (measured at baseline)	Policy significance (from ACA authorizing legislation)	National program model significance (theory, goals, or target population)
<u>High demographic risk</u>			
Type: Binary (high risk and lower risk)	1) Mother is younger than age 21.	1) Pregnant women who have not attained age 21.	HFA, NFP
	2) Mother does not live with child's father.	2) N/A.	HFA, PAT
	3) Mother receives any form of public assistance.	3) Low income families.	HFA, NFP, EHS
	4) Work or school status: a) If the mother is younger than 19: not currently enrolled in school, not working or not in job training. b) If the mother is older than 19: does not hold a high school diploma.	4) Are or have children with low student achievement.	

- Whether or not the woman is a first-time mother. This is binary indicator, comprised of two subgroups: women who are first-time mothers and those who are not. As with pregnancy status, there is some confounding with national model for this subgroup because NFP only enrolls first-time mothers. This focus, in part, reflects both theory and some prior evidence that first-time mothers may be more receptive to home visiting services than mothers who have an older child. This receptivity, in turn, may translate to larger impacts for first-time mothers. The Report to Congress indicated that about one-third of women in the sample already had a child living in the home; and just over half of women enrolling after the birth of their child already had an older child in the home.
- Whether or not the mother has low psychological resources. Two subgroups would be created for this construct, based on a composite measure created from three dimensions: maternal intelligence (based on the Wechsler Adult Intelligence Scale (WAIS)), mental health (based on measures of depressive symptoms and anxiety using the Center for Epidemiologic Studies Depression Scale (CES-D) and the Generalized Anxiety Disorder Scale (GAD-7), and sense of mastery (based on the Pearlin mastery scale). Although this construct did not appear in our scan of policy or program significance, it was the only family subgroup to be included in the HomVEE review because it was the only one that was replicated in the same outcome domain in at least two studies using different analytic samples (the studies also had to meet the other HomVEE requirements).⁵⁵ This suggests that the empirical evidence is strong for maternal psychological resources related to certain outcomes.
- Maternal attachment style (relationship insecurity) and depression. We propose creating two subgroups for this construct, based on two dimensions: relationship insecurity as measured by the Attachment Style Questionnaire (ASQ), and depression as measured by the CES-D. On the one hand, mothers with co-occurring mental health risks may be the most difficult for programs to reach; on the other hand they may differentially benefit from home visiting.⁵⁶

We propose two subgroups for this construct: one subgroup of mothers who are at risk for depression and who exhibit relationship insecurity; the other subgroup of all other mothers.

We seek input from the Committee on this subgroup construct: Relationship insecurity encompasses both attachment anxiety and attachment avoidance. However, if we were to further divide the subgroups (there are eight possible groups resulting from two-group splits of depression, attachment anxiety, and attachment avoidance), sample sizes per group will be too small. Further, depression is captured in the “Low psychological resources” measure described above. Even though depression and relationship insecurity have been examined together in recent research, is it advisable to have two different confirmatory subgroup constructs include a construct?

⁵⁵<http://homvee.acf.hhs.gov/Review-Process/4/Assessing-Evidence-of-Effectiveness/19/7>.

⁵⁶Berlin et al. (2011); Cluxton-Keller et al. (2014); Duggan et al. (2009); Robinson and Emde (2004).

- Presence of domestic violence. A number of studies have looked at subgroups defined by the presence or absence of domestic violence at baseline. These studies have typically defined domestic violence in one of two ways: (1) any incident of physical assault by either the mother or partner (four studies),⁵⁷ or (2) psychological or physical abuse toward the mother by her partner (two studies).⁵⁸ These studies have generally found that baseline domestic violence attenuates home visiting program impacts in a number of outcome domains, including parenting, child development, child maltreatment, and measures of father engagement and the quality of the couple's relationship. We can create subgroups in MIHOPE that capture the presence or absence of intimate partner violence (physical violence toward the mother or physical violence perpetrated by the mother, as measured using the Conflict Tactics Scale). The MIHOPE Report to Congress indicates that 10 percent of MIHOPE families reported the presence of any violence toward the mother and 25 percent reported any violence perpetrated by the mother. The relatively small sample sizes suggest that it might be appropriate to create a measure that captures the presence of any physical violence, either toward the mother or perpetrated by the mother.⁵⁹
- Whether or not the mother exhibits high demographic risk. A measure of high demographic risk would be based on having all four of the following characteristics: the mother is younger than age 21; the mother does not live with the child's father; the mother receives any form of public assistance; work or education status based on whether the mother is younger than age 19 (if younger than age 19, then if the mother is not currently enrolled in school, not working, or not in job training; if the mother is at least 19, then if the mother does not hold a high school diploma). As shown in Appendix Table G.1 and G.2, prior studies have used these characteristics either singly or as part of indexes. Each of the elements is either specifically mentioned in the authorizing legislation, by one of the national models, or both. "High-risk" would be defined as having all four risk factors, while having 0-3 risk factors would be considered "Lower-risk." **We seek Committee input on this index**, considering another cutpoint (for example, a score of 0 would be "Low risk" and scores 1 through 4 would be "Higher Risk;" or more than two categories.

Strategy for estimating confirmatory subgroup impacts

Following is an outline of the empirical strategy that will be used in examining confirmatory subgroup impact estimates:

- The main focus of the confirmatory subgroup impact analyses will be tests of whether estimated subgroup impacts within a construct are statistically different from each other for particular outcomes.

⁵⁷Caldera et al. (2007); Duggan et al. (2007); Duggan et al. (2004a); Duggan et al. (2005).

⁵⁸Duggan et al. (2004b); Mitchell-Herzfeld et al. (2005).

⁵⁹The subgroup would not be based on psychological abuse because most studies have defined subgroups based only on physical violence and because MIHOPE did not collect a valid and reliable measure of psychological abuse at baseline.

- A multiple comparison adjustment will be applied to all statistical tests for confirmatory subgroup impacts. We will likely use the Westfall-Young adjustment (the same adjustment used for estimating impacts using the full sample).
- All statistical tests will (for example, a *t*-test) will be two-tailed, since this is the standard approach in social policy research.
- Primary subgroup impacts will be estimated using both a split-sample approach and an approach that uses the pooled sample with full interactions. In the split sample approach, the full sample will be split into various mutually exclusive subgroups and impacts estimated for each part of the sample. The split-sample results will be reported in the main report; results from the full-interaction method will be reported in an appendix. Use of the two approaches and the reporting locations has been used in prior MDRC studies. The two methods should produce results similar results, though some readers prefer one presentation to the other.
- We will include statistical controls for each national model when estimating primary subgroup impacts. This empirical strategy allows us to measure whether subgroup differences are present within national model. (As discussed in the memo to the Committee on impact variation, MIHOPE will also investigate whether impacts differ across the four national models.)