

# Evaluation of Maternal and Child Home Visitation Programs: Lessons from Pennsylvania

Meredith Matone, Cara Curtis, Arina Chesnokova, Katherine Yun, Amanda Kreider, Meredith Curtis, David Rubin

## EXECUTIVE SUMMARY

Evaluating the effectiveness of maternal and child home visitation programs can be challenging. The complexity of program models, coupled with considerations of real-world program implementation, requires thoughtful and flexible evaluation strategies. Yet, meaningful evaluations of home visitation programs are not only feasible, but are essential as policymakers demand stronger linkages between funding decisions and program outcomes.

Within the current context of the scale-up of maternal and child home visitation programs under the Affordable Care Act's *Maternal, Infant, and Early Childhood Home Visitation* program (MIECHV), the importance and relevance of demonstrating program effectiveness has grown. Drawing from lessons learned through PolicyLab's evaluation of the Pennsylvania Nurse-Family Partnership (NFP) program, this *Evidence to Action* brief highlights four key program evaluation concepts to guide maternal and child home visitation administrators engaged in planning real-world program evaluation.



### KEY CONCEPT

- 1 Real-world evaluation results will reflect implementation environments with fewer supports, resources, and standardization in comparison to randomized trial environments.
- 2 Program performance is altered by the local context of the implementing site and the community it serves.
- 3 Program effectiveness increases over time following wide-scale implementation.
- 4 Engaging stakeholders enriches program evaluation.

### RESPONSE FOR EVALUATION

- ➔ Set realistic target outcomes, knowing that results from studies conducted under experimental conditions will likely demonstrate larger effects relative to evaluation results achieved under real-world conditions.
- ➔ Collect data at the site level, knowing that some sites will be more effective than others, depending upon local resources and baseline population differences. Learn from program outliers, both those that exceed expectations and those that underperform.
- ➔ Select evaluation benchmarks that account for expected changes in effectiveness over time, including the likely lag from implementation to effectiveness.
- ➔ Include program stakeholders in the evaluation process to inform the interpretation of findings for targeted quality improvement.

## INTRODUCTION

Evaluating the effectiveness of maternal and child home visitation programs can be challenging. It requires adaptation to complex program models and real-world implementation realities, such as resource constraints, site-specific infrastructure, and client variability. However, designing and conducting reliable evaluation studies is not only feasible, but essential at a time when policymakers are calling for greater connection between government funding decisions and program outcomes.

Within the field of maternal and child home visitation, there has been a concerted effort to ensure that state and federal dollars are spent on programs that demonstrate improved outcomes for families and children. The federal *Maternal, Infant, and Early Childhood Home Visitation Program* (MIECHV), established by the *Patient Protection and Affordable Care Act* (ACA)<sup>1</sup> in 2010, was allocated \$1.5 billion over five years to strengthen and expand home visitation programs.<sup>2</sup> Programs funded through MIECHV aim to promote maternal and child health, decrease child maltreatment, and promote school readiness and educational attainment.<sup>1</sup> The legislation specifies that states must spend at least three-quarters of their funds implementing or strengthening *evidence-based programs*—those that have been evaluated using rigorous scientific methods.<sup>3</sup> The remaining quarter can be used for *promising practices* without an established evidence base to date. For all funded programs, states must pursue data collection and monitoring in six benchmark areas.<sup>b</sup>

The MIECHV requirements are intended to strengthen the impact of home visitation by prioritizing programs that have demonstrated positive results.

However, the wide-scale implementation of evidence-based programs is challenging because it requires state and local program administrators to translate research concepts and tools into the practice environment. Through MIECHV, states must also establish the infrastructure for program monitoring and evaluation in a short time.<sup>4,5</sup> Many states have collaborated with academic institutions or private research agencies to meet their immediate evaluation requirements. In the long term, the maturation of a strong evidence base for home visitation and funding security will depend upon sustained, rigorous evaluation efforts by state and local government agencies.<sup>6,7</sup>

In 2008, PolicyLab at The Children’s Hospital of Philadelphia was engaged by the Commonwealth of Pennsylvania to evaluate the Pennsylvania Nurse-Family Partnership program. This brief distills the lessons learned from this evaluation and describes the strategies used to address some of the challenges tied to real-world program evaluation.

The brief proceeds in three parts: First, we offer an overview of public health program evaluation. Second, we provide a short description of what we learned from PolicyLab’s home visitation evaluation. Finally, we describe how our findings can be meaningful for the broader home visitation community, highlighting three key program evaluation concepts from our work. This *Evidence to Action* brief is intended to be useful to state and local maternal and child home visitation program administrators as they make decisions regarding ongoing or planned program evaluation.

<sup>a</sup> Approved models are based on the evaluation performed by the U.S. Department of Health and Human Services’-sponsored Home Visiting Evidence of Effectiveness (HomVEE) study (<http://homvee.acf.hhs.gov/>). If a state wishes to suggest an alternative program or believes HomVEE inappropriately evaluated their program of choice, a formal appeals process must be followed.<sup>7</sup>

<sup>b</sup> States are required to monitor and evaluate outcomes in the following six benchmark areas, using prescribed constructs, and report on improvement on at least four areas in a 3-year period and on all six at the end of the 5-year period: (1) maternal and newborn health; (2) child injuries, abuse, neglect or maltreatment and reduction of emergency department visits; (3) school readiness and achievement; (4) crime or domestic violence; (5) family economic self-sufficiency; and (6) coordination and referrals for other community resources and supports.<sup>6,7</sup>

## OVERVIEW OF PUBLIC HEALTH PROGRAM EVALUATION

Evaluation is a systematic, planned process of gathering and analyzing information that answers program performance questions relevant to participating stakeholders.<sup>12</sup> Public health program evaluation can occur in two distinct settings: **experimental and real-world**. Programs being evaluated in experimental settings (e.g., a randomized controlled trial) are being tested for efficacy, i.e., whether the program has the potential to effect change when operating under ideal conditions. Programs being evaluated in a real-world setting are being evaluated for their effectiveness, i.e., whether the program effects change when implemented at a larger scale across multiple communities. Real-world program evaluation occurs in settings that are typically more complex, less controlled, and less resource-rich than an experimental setting.<sup>12, 13</sup>

Within the context of evidence-based home visitation, the most pressing research agenda is that of effectiveness. There are largely two categories of effectiveness research: implementation evaluation and outcome evaluation.

**Implementation evaluation** answers questions about whether a program was implemented as originally intended. It provides information about the types and number of services offered, the number of people served, staffing arrangements and recruitment (outputs).<sup>14</sup> In contrast, **outcome evaluation** is used to determine whether a program has had the intended effect (outcome) in the community it serves and whether the improvements can be attributed to the program.<sup>10</sup> Programs of a defined duration are often evaluated at the end of their activities. However, periodic, ongoing outcome evaluation is critical to program improvement and sustainability.<sup>15</sup>

A final consideration in program evaluation is evaluation design, which may be experimental, quasi-experimental, or non-experimental. Experimental designs, namely randomized trials, are associated with the highest strength of evidence for determining program efficacy and/or effectiveness. These designs are often employed to evaluate newly developed program models or program adaptations.

Quasi-experimental designs are another important mechanism used to assess program effect, but lack the randomization feature of experimental designs and therefore may have increased vulnerability to bias. Quasi-experimental designs enable large scale outcome evaluation in real-world settings. They typically compare program participants to demographically similar groups who have

not received program services. In real-world evaluation settings, randomization may be unfeasible or inapplicable for a host of reasons, including ethical concerns about restricting access to services. PolicyLab's evaluation of the Pennsylvania Nurse-Family Partnership is an example of a quasi-experimental outcome evaluation conducted in a real-world setting focusing on program effectiveness.

Non-experimental designs (e.g., benchmarks or qualitative efforts) typically monitor program inputs and outputs or explore contextual information related to performance. Non-experimental designs do not include a comparison group and do not allow evaluators to appraise whether changes in outcomes should be attributed to the program.

Evaluation designs can also be combined in mixed-methods evaluations, which often feature quasi-experimental designs in tandem with non-experimental qualitative research such as focus groups or interviews. Each evaluation design has utility depending on the research question of interest and the circumstances surrounding the evaluation.

FIGURE 1:

### KEY TERMS

**Efficacy** – the beneficial effects of a program or policy under optimal conditions of delivery.<sup>8</sup>

**Effectiveness** – the effects of a program or policy under more real-world conditions.<sup>8</sup>

**Program Evaluation** – the systematic collection of information about the activities, characteristics, and outcomes of programs to make judgments about the program, improve program effectiveness, and/or inform decisions about future program development.<sup>9</sup>

**Stakeholder** – an individual or group interested in or affected by the program.<sup>10, 11</sup>

**Outputs** – program structure and services delivered.

**Outcomes** – specific changes associated with program participation; outcomes are linked with program objectives and can be knowledge, behavior, or health status related.

**Implementation Evaluation** – an assessment of program operations and activities.

**Outcome Evaluation** – an assessment of program effect on outcomes through experimental design.

## WHAT WE LEARNED

PolicyLab evaluated the Nurse-Family Partnership (NFP) following statewide implementation in Pennsylvania (PA). NFP is an evidence-based home visitation model serving more than 23,000 low-income, first-time mothers each year in 42 states ([www.nursefamilypartnership.org](http://www.nursefamilypartnership.org)). The program is designed to improve pregnancy outcomes, child health and development, and family economic self-sufficiency. Nurses visit mothers in their homes during pregnancy and for up to two years postpartum. At each visit, nurses provide education using a standardized curriculum. In PA, the program includes 24 agencies, operating in 40 of the state's 67 counties. These agencies serve a racially, ethnically, and geographically diverse client base, including women from historically marginalized racial and ethnic groups, as well as women in rural areas. Program sites are administered by local health departments, nursing associations, community-based organizations, health systems, or public-private partnerships. Sites also accept a wide range of referrals from community-based organizations, government agencies, and healthcare providers.

The PolicyLab evaluation compared PA NFP clients to women who were economically and demographically very similar but who had not enrolled in PA NFP. The study included approximately 6,000 NFP clients served by 24 sites from 2000 through 2007. PA NFP clients were compared to approximately 17,000 first-time mothers of similar age, ethnicity, educational attainment, marital status, income, and residential location.<sup>c</sup>

This quasi-experimental evaluation focused on the following outcomes: (1) pregnancy spacing, (2) prenatal smoking cessation, and (3) child injuries in the first two years of life.

PolicyLab's evaluation of PA NFP revealed four key lessons related to public health evaluation. While the concepts highlighted here are described in previous evaluation literature,<sup>8, 13, 15</sup> they present especially pertinent considerations for real-world evaluation of maternal and child home visitation programs. We illustrate these concepts with examples from PolicyLab's evaluation in PA, in order to inform future efforts to develop evaluation plans.

**The four real-world program evaluation concepts are as follows:**

- 1. Program success** is typically greatest under experimental conditions, a phenomenon that program evaluators and policymakers must consider and prepare for when evaluating programs in real-world settings.
- 2. Site-specific implementation** evaluation is necessary to understand the reasons for variation in program effectiveness across sites.
- 3. Outcome evaluation planning** should take into account a potential lag in effectiveness following implementation and the potential for increase in effectiveness over time.
- 4. Engaging stakeholders** enriches program evaluation. Additional qualitative data from program stakeholders can help evaluators interpret quantitative findings.

We discuss these concepts in greater detail and review findings from PolicyLab's evaluation in the next section. Most importantly, we offer recommendations for each concept to strengthen meaningful real-world evaluation efforts.

<sup>c</sup>Data for matching and subsequent analyses were provided by PA NFP and the Commonwealth of PA, including the Division of Vital Records, the Department of Public Welfare

**1 KEY CONCEPT:** Real-world evaluation results will reflect implementation environments with fewer supports, resources, and standardization in comparison to randomized trial environments.

**RESPONSE FOR EVALUATION:** Set realistic targets, knowing that results from evaluation studies conducted under experimental conditions are likely to have demonstrated larger effects relative to the results that will be achieved under real-world conditions.

While the results of randomized controlled trials provide valuable insight into program potential, they do not guarantee program success in real-world settings. Social and public health programs are typically less effective following wide-scale implementation than under experimental conditions. This phenomenon has been described as a “voltage drop” in effectiveness and is most likely due to changes associated with program expansion.<sup>16</sup> By definition, widely implemented programs reach communities that may not have been represented in randomized trials, each with distinct clientele demographics, staff proficiency and training, funding mechanisms, health and social service system infrastructure, and social, cultural, and behavioral norms. Moreover, program expansion efforts are typically not as well-resourced as programs being evaluated by an

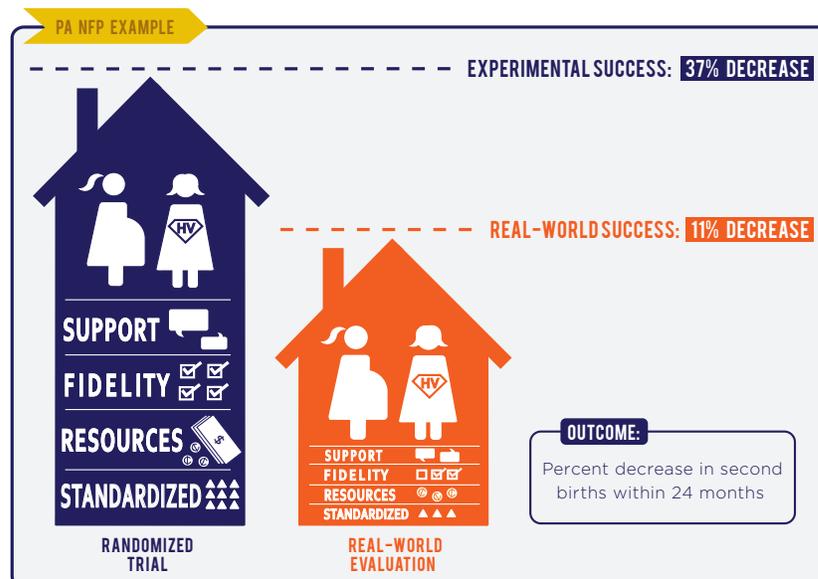
experimental trial, despite the host of new elements to manage in the scale up. As a result, there is often a mismatch between the characteristics of the original program and the needs and resources of the expansion site. Mismatch typically occurs in three areas: (1) program participants, (2) program staff, and (3) administrative and community factors.<sup>17</sup>

The challenges stemming from this mismatch are typified by the results of a recent PA effectiveness evaluation, which found no difference in the number of emergency department (ED) visits for serious childhood injuries between the children of NFP clients and those in the comparison group.<sup>18</sup> In contrast, the first randomized trial of NFP in Elmira, New York, in 1977 found significantly fewer ED visits for accidents and poisoning in the second year of life among nurse-visited children who received NFP.<sup>19</sup>

Similarly, while the PolicyLab evaluation demonstrated a lower rate of rapid second pregnancies (defined as a second birth within 24 months of the first birth) among NFP clients than the comparison group, the effectiveness of the program in PA (17% of NFP clients had rapid second pregnancies, compared to 19% of the comparison group, a relative decrease of 11%.) was lower than the efficacy observed in the randomized trials of NFP (See Figure 2).<sup>20, 22</sup> During the randomized trial in

## FIGURE 2: Randomized Trials: Fortified by Design and Resource Supports

For further information on data used in Figures 2–5, see references 18, 20, and 24.



Denver, Colorado (1994), 12% of NFP clients had a second birth within 24 months, compared to 19% among the comparison group, a relative decrease of 37%.<sup>21</sup>

Although randomized trials produce evidence of a program's potential in a particular setting, program success is influenced by many factors following real-world implementation, and positive results are not a certainty. The knowledge that real-world effects are likely to be smaller than those observed in efficacy studies should empower programs to continue monitoring outcomes.

“...effectiveness research studies programs under typical, not optimal conditions...”<sup>23</sup>

As programs broaden their impact by reaching new communities, ongoing outcome evaluation is critical to ensure that programs are proactive in identifying areas for targeted quality improvement.

**2 KEY CONCEPT:** Program performance is altered by the local context of the implementing site and the community it serves.

**RESPONSE FOR EVALUATION:** Collect data at the site level, knowing that some sites will be more effective than others, depending upon local resources and baseline population differences. Learn from program outliers, both those that exceed expectations and those that underperform.

As programs expand into highly diverse communities and are disparately administered and staffed, cross-site variability in program performance is also likely to rise. For example, PolicyLab's PA NFP evaluation found variation across sites in rates of superficial injuries to young children of NFP clients; specifically, sites' injury rates ranged from 3.5% to 24.6%. Further, there was a large degree of variation in how much NFP client injury rates differed from rates in the comparison group. Nineteen sites experienced more injuries among NFP clients than those in the comparison group, and at five sites there were more injuries among the comparison group than NFP clients.

Variation between program sites may be attributed to differences in *program implementation* and in *local context*. Program implementation is an umbrella construct most commonly covering: fidelity to program model, quality of administration of program components, dosage of program administered, and program reach or coverage of the targeted population. As an example of the frequency of variability in this construct, a review of child and youth health prevention and promotion interventions noted

that among assessments of program implementation (measured by fidelity, dosage, quality, and reach), 20% to 40% variation in implementation levels between sites or providers within a single program model was common.<sup>13</sup>

The local environment in which a program is set to operate is also inextricably linked to the level of its success. Local context matters because home visitation programs primarily operate as behavior change interventions, wherein a home visitor provides education and support to facilitate the adoption of healthy behaviors by families. While behavior change interventions are a proven tool in public health, they often compete against or work alongside community-level norms and behaviors.

For instance, the evaluation of PA NFP success in achieving maternal smoking cessation found prenatal smoking cessation behavior to be strongly associated with local levels of community prenatal smoking. For each 10% increase in community smoking, we observed a 30% decrease in the likelihood of a woman quitting smoking, irrespective of whether or not she was served by NFP (See Figure 3).<sup>24</sup>

Given inter-site variability in outcomes, it is crucial that programs do not limit themselves to aggregate evaluations, which summarize program effectiveness across multiple sites and large regions. Such aggregation will often omit important successes and failures. As such, meaningful evaluation should take into account performance and community norms at a site level.

FIGURE 3: Community Context Affects Program Outcomes



**3 KEY CONCEPT:** Program effectiveness increases over time following wide-scale implementation.

**RESPONSE FOR EVALUATION:** Select evaluation benchmarks that account for expected changes in effectiveness over time, including the likely lag from implementation to effectiveness.

Public health program implementation must pass through several stages before new programs can function sustainably and effectively within a community.<sup>25</sup> Programs typically pass through stages of: *installation*, when the program acquires resources; *initial implementation*, when program staff are learning their roles and how to fill them most efficiently; and *full operation*, when staff skills are fully developed and the program has become integrated into the target community.<sup>25</sup>

The time required for new programs to progress through the stages of implementation is unique to each program and each community. However, all programs must pass through these stages in order to achieve maximum effectiveness. Programs that have not yet reached full operation will likely demonstrate poor outcomes or outcomes not applicable to the program's future success.

In the PolicyLab evaluation of PA NFP, we discovered that program effectiveness increased considerably between the initial implementation period and later stages of program operation. More specifically, we observed no significant differences between NFP clients and their comparison group in achieving recommended pregnancy spacing during the initial four years following program expansion (2000-2003), yet noted significant differences between the groups during the fifth and sixth year (2004-2005), with 19% of the comparison group having a second birth within 24 months compared to 16.8% of NFP clients.<sup>20</sup> Had our outcome evaluation focused on the initial implementation phase, or combined the entire period following dissemination, we would have erroneously concluded the program to be unsuccessful and undermined its potential to achieve a positive maternal health outcome (See Figure 4).

Setting realistic expectations for program evaluation results includes recognition of the delay in any new program's ability to show effectiveness. With sparse resources for evaluation, it may be useful for evaluation efforts to defer outcome evaluation and prioritize implementation evaluation during program roll-out periods. Programs should disaggregate evaluation not only across sites (*Key Concept 2*), but also across time, to paint the most genuine portrayal of the program's success and potential.

**FIGURE 4: Program Effectiveness Increases With Time**



**4 KEY CONCEPT:** Engaging stakeholders enriches program evaluation.

**RESPONSE FOR EVALUATION:** Include program stakeholders in the evaluation process to inform the interpretation of findings for targeted quality improvement.

Even in the most rigorous evaluations that employ a quasi-experimental design with an appropriate comparison group, it is impossible to overcome the limitation that, by nature, observational data on programs over time are always subject to potential bias in interpretation. Unlike randomized trials, where the process of randomization inherently selects program recipients who share all the characteristics of non-program recipients, the real world

of evaluation offers no such ease of comparison. For this reason, stakeholder engagement and qualitative research are useful evaluation resources that can add meaning and context to quantitative findings.

For example, prior trials and systematic reviews had revealed home visiting programs to have significant benefit in reducing early childhood injuries.<sup>19, 22, 26</sup> In contrast, following implementation of NFP across Pennsylvania, injury rates, as measured by emergency department visits and hospitalizations, were detected to be higher overall among infants of women enrolled in the program compared to infants of women who were not enrolled in the program.<sup>18</sup> The greatest differences were for superficial injuries (see Figure 5) but, even so, no program impact was detected for more serious injuries and for child abuse.

*continued...*

**FIGURE 5: Qualitative Data Can Supplement Findings**



At face value, it would be easy to conclude that, following implementation, the NFP program had failed to achieve the results that were demonstrated in efficacy trials. However, it is difficult to reach this conclusion with confidence using quantitative data alone. Ultimately, a reason for such conflicting data may be found, but arriving at this reason requires a careful approach with deeper qualitative study at the local level that can put the quantitative data into perspective. Interviews or focus groups at local program sites can be used to confirm whether additional selection differences between women who enrolled in the program versus women who did not enroll and were chosen for the comparison group might have biased the results (for example, higher rates of juvenile justice involvement among enrollees).

Alternatively, such an approach might also reveal whether the curriculum was being delivered as intended (intervention fidelity), whether poor retention to the program may be interfering with program success, whether increased surveillance of families by home visitors may have influenced health care behaviors, or whether the return of mothers to school or work may have created unintended challenges to appropriate supervision and childcare for young infants. When such in-depth qualitative analyses are conducted among positive or negative outlier sites, best practices and quality improvement needs can be identified.

---

## CONCLUSION

The complexity of maternal and child public health programs, coupled with the myriad internal and external forces that influence implementation, participation, and sustainability, makes evaluation challenging. At the same time, these attributes of public health programs make evaluation a linchpin to success. It is only through thoughtful, multi-faceted evaluation that programs can reach optimum performance.

On an operational level, it is important to recognize that no evaluation can answer every question. Monitoring program outputs and the level of implementation processes alone does not inform stakeholders about the effectiveness of a program, and even the best single-design outcome evaluations do not capture the experiences of program staff and the families they serve.

**Even if limited in scope, well-planned evaluations can reap useful rewards.** Using the meaningful evaluation principles of this brief would suggest, for example, that programs prioritize implementation evaluation measures over outcome evaluation measures during early implementation and expansion phases. This strategy provides information for quality improvement initiatives that can be rolled out early if implementation efforts are

struggling. Similarly, this brief provides rationale for a mixed-method site-specific evaluation approach, inclusive of a qualitative component in addition to quantitative outcome metrics, to uncover how community norms (e.g., smoking and prenatal care utilization) or health-system infrastructure (e.g., referrals and funding) may affect program success at the site level. The PA effectiveness study featured in this brief underscores the power of quasi-experimental designs in answering a multitude of evaluation questions. Finally, this brief should help policy and program stakeholders set reasonable expectations for overall success at every stage. Programs can account for potentially weaker results than those witnessed during efficacy trials and site-to-site variation in performance, while simultaneously identifying areas of strength for reinforcement and target areas for improvement.

**MIECHV offers an exciting opportunity to implement valuable evaluation of maternal and child home visitation efforts in states across the country. Beyond the confines of the MIECHV timeline, prioritizing sustained evaluation goals for home visitation programs is critical to building programs that flourish, rather than falter, as they serve increasing numbers of families and communities across the nation.**

*For those interested in data sources for the quasi-experimental methodologies discussed in this brief, we include an appendix with additional information.*

## APPENDIX: DATA SOURCES FOR QUASI-EXPERIMENTAL EVALUATION

Reliable data are crucial to evaluation. While data on program outcomes can be extracted from program records, generating meaningful comparison group outcomes and community indicators can present a challenge in quasi-experimental outcome evaluations. This appendix suggests some potential sources for these data. In conducting evaluations, PolicyLab has often turned to administrative data sources and governmental data warehouses to generate such measures. Vital statistics, welfare and public insurance enrollment records, and publicly available community health indicator databases are the major sources of information utilized by the PolicyLab research team in home visitation evaluation. A number of other maternal-child home visiting quasi-experimental evaluations have also used the above data sources to both extract additional information about program enrollees as well as to create comparison groups of nonparticipants.<sup>27-29</sup> It is important to note that administrative data sources external to the program contain data that can be used in evaluating benchmarks prescribed by the MIECHV evaluation requirements.

### Vital Statistics

Vital statistics are records, collected by each state, that include birth and death records. Birth certificate files are of particular interest to evaluators of maternal and child health programs as they contain detailed parental demographic and health information as well as birth outcome information. While individual-level data on births and deaths are highly protected,<sup>d</sup> states often make aggregate statistics available to the public. Demographic information, prenatal tobacco use, and inter-birth intervals are examples of the information extractable from birth certificate data.

### Welfare and insurance Enrollment Records

Each state maintains welfare records and records of public insurance enrollment among its residents. This data is inclusive of programs such as TANF, SNAP, SSI, Medicaid and state-financed public programs. Such records can be used to build comparison groups of women not receiving home visitation services via the extraction of public-service eligibility to serve as an indicator of risk. Additionally, these data provide information on critical child health outcomes, such as injury-related emergency department visits, which can be extracted from Medicaid claim files.

### Public Health Databases

Finally, while individual-level data are an essential component to quasi-experimental designs, they often are not useful to assess the risk profile of a community as a whole. A number of resources report public health measures at the county and state levels and present timely data for evaluating the health and well-being of communities served by home visitation programs. These public health databases include the U.S. Department of Health and Human Services (HHS) Community Health Status Indicators database,<sup>e</sup> the HHS Leading Health Indicators database,<sup>f</sup> and the federal, state, and industry-run H-CUP KIDS database.<sup>g</sup>

Though this is a broad sampling of data sources useful in quasi-experimental evaluation of home visitation programs, it is not an exhaustive list. Home visitation programs should carefully assess their unique target populations and desired outcomes in order to determine the best sources for comparison data.

<sup>d</sup>Individual-level (i.e. non-aggregate) vital statistics, welfare and public insurance enrollment data are highly protected. Access to such records is contingent on research need and ability to safeguard data. Often, data use agreements with appropriate state agency charged with maintaining such data are the mechanism of data access. Relevant state agencies for inquiries about the appropriate protocols for obtaining data records include Departments of Health, Departments of Public Welfare, or the equivalent state bodies.

<sup>e</sup><http://www.communityhealth.hhs.gov>

<sup>f</sup><http://www.healthypeople.gov/2020/LHI/default.aspx>

<sup>g</sup><http://www.hcup-us.ahrq.gov>

## REFERENCES

- 111th United States Congress. Patient Protection and Affordable Care Act. P.L. 111-148. Washington, DC; 2010.
- Patient Protection and Affordable Care Act of 2010. Pub L. No. 111-148. Sec. 2951 (j) (1). <http://thomas.loc.gov/cgi-bin/query/F?c111:7:/temp/~c111dmH9i6:e698181>.
- Patient Protection and Affordable Care Act of 2010. Pub L. No. 111-148. Sec. 2951 (d) (3) (A). <http://thomas.loc.gov/cgi-bin/query/F?c111:7:/temp/~c111dmH9i6:e698181>.
- U.S. Department of Health and Human Services. News Release: Health care law expands support for children and families. Washington, DC. April 3, 2012. <http://www.hhs.gov/news/press/2012pres/04/20120403b.html>.
- U.S. Department of Health and Human Services. HRSA-11-179 Funding Opportunity Announcement. Washington, DC. June 1, 2011. <https://grants3.hrsa.gov/2010/Web2External/Interface/FundingCycle/ExternalView.aspx?&fCycleID=75773544-C311-43E1-8668-7DAD95696629&ViewMode=EU&GoBack=&PrintMode=&OnlineAvailabilityFlag=True&pageNumber=3>.
- Patient Protection and Affordable Care Act of 2010. Pub L. No. 111-148. Sec. 2951 (d) (1) (A). <http://thomas.loc.gov/cgi-bin/query/F?c111:7:/temp/~c111dmH9i6:e698181>.
- Affordable Care Act Maternal, Infant, and Early Childhood Home Visiting Programs – Supplemental Information Request for Submission of the Updated State Plan for a State Home Visiting Program. Section 3 (p. 10) <http://www.hrsa.gov/grants/manage/homevisiting/sir02082011.pdf>.
- Flay B, Biglan A, Boruch R, et al. Standards of evidence: Criteria for efficacy, effectiveness and dissemination. *Prevention Science*. 2005/09/01 2005;6(3):151-175.
- Patton MQ. Utilization-focused evaluation. 3rd ed. Thousand Oaks, California: Sage Publications; 1997.
- Wholey JS, Hatry HP, Newcomer KE. Handbook of practical program evaluation. Vol 19. San Francisco, CA: Jossey-Bass; 2010.
- Cooney SM, Huser M, Small S, O'Connor C. Evidence-based programs: An overview. *What Works, Wisconsin – Research to Practice Series*: University of Wisconsin; 2007.
- Introduction to program evaluation for public health programs: A self-study guide. Atlanta, GA: Centers for Disease Control and Prevention; 2011.
- Durlak J, DuPre E. Implementation matters: A review of research on the influence of implementation on program outcomes and factors affecting implementation. *American Journal of Community Psychology*. 2008/06/01 2008;41(3-4):327-350.
- The Program Manager's Guide to Evaluation. 2nd ed. Washington, DC: US Department of Health and Human Services; 2010.
- Fixsen D, Naoom S, Blase D, Friedman R, Wallace F. Implementation research: A synthesis of the literature. University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network (FMHI Publication# 231): FL Tampa; 2005.
- Kilbourne AM, Neumann MS, Pincus HA, Bauer MS, Stall R. Implementing evidence-based interventions in health care: application of the replicating effective programs framework. *Implementation Science*. 2007;2.
- Castro FG, Barrera M, Martinez CR. The cultural adaptation of prevention interventions: resolving tensions between fidelity and fit. *Prevention Science*. 2004;5(1):41-45.
- Matone M, O'Reilly AL, Luan X, Localio AR, Rubin DM. Emergency department visits and hospitalizations for injuries among infants and children following statewide implementation of a home visitation model. *Maternal and Child Health Journal*. 2011.
- Olds DL, Henderson CR, Jr., Chamberlin R, Tatelbaum R. Preventing child abuse and neglect: a randomized trial of nurse home visitation. *Pediatrics*. Jul 1986;78(1):65-78.
- Rubin D, O'Reilly A, Luan X, Dingwei D, Localio AR, Christian CW. Variation in pregnancy outcomes following statewide implementation of a prenatal home visitation program. *Archives of Pediatrics & Adolescent Medicine*. 2010(In Press, November 2010).
- Olds DL, Robinson J, O'Brien R, et al. Home visiting by paraprofessionals and by nurses: a randomized, controlled trial. *Pediatrics*. Sep 2002;110(3):486-496.
- Kitzman H, Olds DL, Henderson CR, Jr., et al. Effect of prenatal and infancy home visitation by nurses on pregnancy outcomes, childhood injuries, and repeated childbearing: A randomized controlled trial. *JAMA*. Aug 27 1997;278(8):644-652.
- Green LW, Glasgow RE. Evaluating the relevance, generalization, and applicability of research. *Evaluation & the Health Professions*. 2006;29(1):126-153.
- Matone M, O'Reilly ALR, Luan X, Localio R, Rubin DM. Home visitation program effectiveness and the influence of community behavioral norms: A propensity score matched analysis of prenatal smoking cessation. *BMC Public Health*. 2012;12(1):1016.
- Fixsen DL, Blase KA, Naoom SF, Wallace F. Core implementation components. *Research on Social Work Practice*. 2009;19(5):531-540.
- American Academy of Pediatrics. The role of preschool home-visiting programs in improving children's developmental and health outcomes. *Pediatrics*. 2009;123(2):598-603.
- Gessner BD. The effect of Alaska's home visitation program for high-risk families on trends in abuse and neglect. *Child Abuse & Neglect*. 2008;32(3):317-333.
- Arima Y, Guthrie BL, Rhew IC, De Roos AJ. The impact of the First Steps prenatal care program on birth outcomes among women receiving Medicaid in Washington State. *Health Policy*. 2009;92(1):49-54.
- Carabin H, Cowan LD, Beebe LA, Skaggs VJ, Thompson D, Agbangla C. Does participation in a nurse visitation programme reduce the frequency of adverse perinatal outcomes in first-time mothers? *Paediatric and Perinatal Epidemiology*. 2005;19(3):194-205.

## THE AUTHORS

**MEREDITH MATONE, M.H.S.** is a research scientist with PolicyLab at The Children’s Hospital of Philadelphia Research Institute and project director for a statewide home visitation program evaluation.

**CARA CURTIS** is a research associate with PolicyLab at The Children’s Hospital of Philadelphia Research Institute.

**ARINA CHESNOKOVA, M.P.H.** is a research associate with PolicyLab at The Children’s Hospital of Philadelphia Research Institute.

**KATHERINE YUN, M.D., M.H.S.** is a faculty member with PolicyLab at The Children’s Hospital of Philadelphia Research Institute, an assistant professor of pediatrics at the University of Pennsylvania Perelman School of Medicine and an attending physician in the Division of General Pediatrics at The Children’s Hospital of Philadelphia.

**AMANDA KREIDER** is a research associate with PolicyLab at The Children’s Hospital of Philadelphia Research Institute.

**MEREDITH CURTIS** is a medical student at the University of Pennsylvania Perelman School of Medicine.

**DAVID RUBIN, M.D., M.S.C.E.** is co-director of PolicyLab at The Children’s Hospital of Philadelphia, an associate professor of pediatrics at the Perelman School of Medicine at the University of Pennsylvania, and an attending physician in the Division of General Pediatrics at The Children’s Hospital of Philadelphia.

*We would like to thank the Pennsylvania Nurse Family Partnership program and the Pennsylvania Office of Child Development and Early Learning for their collaboration and support of the research that informed this brief. PolicyLab’s Kathleen Noonan and Sarah Zlotnik acted as editorial and content advisors. Support for the brief was provided by The Pew Charitable Trusts.*



The aim of PolicyLab at The Children’s Hospital of Philadelphia is to achieve optimal child health and well-being by informing program and policy changes through interdisciplinary research.

PolicyLab develops evidence-based solutions for the most challenging health-related issues affecting children. We partner with numerous stakeholders in traditional healthcare and other community locations to identify the programs, practices, and policies that support the best outcomes for children and their families. PolicyLab disseminates its findings beyond research and academic communities as part of its commitment to transform evidence to action.







---

*PolicyLab Evidence to Action briefs highlight PolicyLab research areas in the context of local and national policy issues to advance child health and well-being.*

[policylab.chop.edu](http://policylab.chop.edu)

**PolicyLab**

The Children's Hospital of Philadelphia  
34th Street and Civic Center Boulevard  
CHOP North, Room 1528  
Philadelphia, PA 19104  
Phone: 267-426-5300  
Fax: 267-426-0380  
[PolicyLab@email.chop.edu](mailto:PolicyLab@email.chop.edu)