Variation in Pregnancy Outcomes Following Statewide Implementation of a Prenatal Home Visitation Program

David M. Rubin, MD, MSCE; Amanda L. R. O’Reilly, MPH; Xianqun Luan, MS; Dingwei Dai, MSc, PhD; A. Russell Localio, PhD; Cindy W. Christian, MD

Objective: To examine, following statewide dissemination, the influence of an evidence-based home visitation program for first-time mothers on reductions of subsequent pregnancies across time and different locations.

Design: Retrospective cohort study.

Setting: Replication sites for the Nurse-Family Partnership (17 urban sites and 6 rural sites) across the Commonwealth of Pennsylvania between January 1, 2000, and December 31, 2007.

Participants: A total of 3844 Nurse-Family Partnership clients matched by propensity score to 10,938 local-area controls.

Main Exposure: Program enrollment.

Main Outcome Measure: Time to second pregnancy resulting in a live birth within 2 years of the first infant’s birth.

Results: There were no program effects on time to first pregnancy in the early years of the program (2000-2003), but clients whose first infants were born after 2003 had fewer second pregnancies compared with controls (hazard ratio=0.87; 95% confidence interval, 0.80-0.96). This benefit occurred principally among mothers who were aged 18 years or younger (hazard ratio=0.73; 95% confidence interval, 0.61-0.89) and was twice as strong among mothers aged 18 years or younger from rural locations (hazard ratio=0.40; 95% confidence interval, 0.22-0.73) compared with those from urban locations (hazard ratio=0.79; 95% confidence interval, 0.65-0.95).

Conclusions: Program effects on pregnancy planning emerged after an implementation period of 3 years in both urban and rural locations, but they were particularly strong in rural locations and among younger mothers.


For editorial comment see page 278

The Nurse-Family Partnership (NFP), a program of prenatal, infancy, and toddler home visitation by nurses for low-income mothers bearing their first children, is designed to improve the outcomes of pregnancy, children’s health and development, and parents’ economic self-sufficiency. The latter goal is accomplished by helping parents plan the timing of subsequent pregnancies through the first child’s second birthday. Growing from its 1978 inception for 400 women in Elmira, New York,1 it has expanded throughout the United States and currently serves more than 20,000 families per year in 31 states.2 Its expansion is owing in large part to data from 3 randomized trials conducted in different locations revealing multiple benefits to the mother and child.1,3-11 For the mother, a focus on family planning has resulted in a reduction in rapid-succession second pregnancies within the 2-year period following birth of the first child.1,2,5 An outcome that can negatively affect outcomes for teenage mothers,1,2 let alone increase the risk for adverse perinatal outcomes.1 Long-term follow-ups of these trials have also found sustained benefit, including reduced welfare receipt9,11 and reduced antisocial behavior among adolescents born to program recipients many years later.7,8,14-16

Based on such evidence, the US Congress recently appropriated in the Patient Protection and Affordable Care Act (PubL No. 111-148) $1.5 billion to expand prenatal and early childhood home visitation programs over the next 5 years. Despite a large public investment in home visitation, however, it remains uncertain whether wider dissemination can reproduce the success of earlier trials. This concern is not trivial as prior prevention programs in maternal and child health...
have failed in some cases to maintain effectiveness following dissemination, reflecting the real-world difficulty in replicating on a larger scale earlier successes from controlled clinical trials or demonstration projects.17-19 Aside from the challenges of implementation, there are concerns that as the program disseminates, it will increasingly reach populations of women who were not well represented in the original trials. Such is the case for a state like Pennsylvania, which boasts the largest rural population in the country.20 Although the Elmina trial was conducted in a mixed urban and rural setting, the other trials of the NFP were conducted in urban locations; it is uncertain whether the benefits identified in those trials would be conferred equally to women served by sites in more rural areas.

This study sought to examine the success of the NFP program following statewide implementation in the Commonwealth of Pennsylvania. To do so, we matched clients served across the 23 NFP sites in Pennsylvania between January 1, 2000, and December 31, 2007, to local-area controls across the Commonwealth to understand agency-level variation in second-pregnancy outcomes resulting in live births and whether the benefits of earlier trials of the NFP were sustained after dissemination across disparate urban and rural locations. We sought to test 2 hypotheses: (1) that success following implementation would not be immediate but would grow over time, and (2) that women served in rural areas would have outcomes comparable to those of women served in urban areas.

The primary sources of data for this study included the following: (1) the enrollment history of clients participating in 23 NFP programs throughout the Commonwealth of Pennsylvania from January 1, 2000, to December 31, 2007; (2) welfare eligibility files from the Department of Public Welfare; and (3) birth certificate files provided by the Pennsylvania Department of Public Health. The target population, determined following a sequential linkage of these data sources, were NFP clients and potential controls who delivered a first-born singleton infant between January 1, 2000, and December 31, 2005, and who received any form of welfare assistance from the Commonwealth of Pennsylvania within 12 months prior to the infant’s birth. Because implementation across 23 sites commenced between 1999 and 2001, we divided our sample into 2 periods: the early clients from January 1, 2000, to December 31, 2003, and the later clients from January 1, 2004, to December 31, 2005. The periods were chosen to coincide with prior literature on implementation, which has suggested that program effectiveness may be delayed for several years following wider implementation of community-based prevention programs.19 Clients were matched to appropriate nonparticipating controls from across the Commonwealth who also delivered a first-born infant during the same period, were welfare enrolled within the 12 months prior to the infant’s birth, and were from the same NFP service region as the participating clients (Figure 1). To select the local-area controls for NFP clients, matching by propensity score21,22 was used to emulate as closely as possible randomized assignment to program and control groups, although we could match only on selection factors that were available from our data sources. The first step of the analysis modeled factors associated with a woman’s participation in the NFP for each agency. The factors included the following: maternal education (<8th grade, <12th grade, some college, and completed college or higher), maternal age, marital status (yes vs no), history of smoking during pregnancy (yes vs no), receiving Temporary Assistance for Needy Families vs other state welfare assistance (eg, to families whose incomes exceeded federal guidelines or who were not otherwise eligible for federal programs), history of having received food stamps, history of pregnancy-induced hypertension, history of gestational diabetes, and whether the mother resided in an urban or rural location.23 This latter variable was encoded using the rural-urban commuting area code 2006 version, a classification that has been applied in other health services and outcomes research.24,25

Logistic regression, modeled on the factors described earlier, was used to estimate an expected probability of participation in the NFP for each client and potential control.26,27 The propensity score models included interaction terms whenever possible to allow for effect modification across a range of characteristics. To ensure that controls were obtained from the same NFP service regions as the participating clients, we also included in these models a variable that encoded high-density zip codes within each agency catchment area to drive the selection of nonparticipating controls toward high-penetration neighborhoods of interest. This density variable was created by identifying zip codes that encompassed more than 5% of the NFP client population within the agency catchment area as high density.

Propensity score analyses were performed separately for each of the 23 agencies because we assumed that criteria for selecting clients would differ across agencies. In addition, matching was stratified within period (2000-2003, 2004-2005) and age (≤18 years, 19-21 years, 22-35 years) categories. Matched using SAS version 9.2 statistical software (SAS Institute, Inc, Cary, North Carolina),28 any nearest neighbor within a caliper of 0.05 who was also within the same age category as the NFP client was considered a match (up to a maximum of 4 matched controls per client).

The primary outcome was the length of time to conceiving a second pregnancy within 15 months post partum of the first pregnancy, assuming that pregnancy ultimately resulted in the live birth of an infant. This time frame was chosen because it roughly translated to the delivery of a second infant within 2 years of the first pregnancy, an interval found to increase the risk for adverse pregnancy outcomes in subsequent pregnancies.13 The variable was calculated by subtracting the second infant’s gestational age in weeks at delivery from his or her birth date and then calculating the remaining interval from the first infant’s birth date.

Data were described using frequency distributions for categorical variables. Cartographic mapping was used to examine geographic representation of clients and controls across the Commonwealth. Using contingency tables and χ2 tests, characteristics for clients and matched controls in each agency were checked for balancing across all covariates.27 Through this analysis, we identified even modest imbalances in point estimates as a threat to residual bias. These imbalances were eliminated for each agency through adjustment of the propensity score models.29
Table 1. Characteristics of Nurse-Family Partnership Clients Compared With All Potential Controls, Welfare-Eligible Controls, and Final Matched Controls Across the Commonwealth of Pennsylvania

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>NFP Clients</th>
<th>All Pennsylvania Births</th>
<th>All Welfare Recipients</th>
<th>Matched Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aged ≤18 y</td>
<td>43%</td>
<td>12%</td>
<td>29%</td>
<td>41%</td>
</tr>
<tr>
<td>African American</td>
<td>25%</td>
<td>13%</td>
<td>41%</td>
<td>29%</td>
</tr>
<tr>
<td>≤12th grade education</td>
<td>85%</td>
<td>50%</td>
<td>79%</td>
<td>85%</td>
</tr>
<tr>
<td>Smoker</td>
<td>41%</td>
<td>16%</td>
<td>22%</td>
<td>38%</td>
</tr>
<tr>
<td>Unmarried</td>
<td>90%</td>
<td>43%</td>
<td>86%</td>
<td>91%</td>
</tr>
<tr>
<td>Receiving food stamps</td>
<td>67%</td>
<td>NA</td>
<td>61%</td>
<td>67%</td>
</tr>
<tr>
<td>Receiving TANF</td>
<td>53%</td>
<td>NA</td>
<td>53%</td>
<td>55%</td>
</tr>
<tr>
<td>Urban</td>
<td>84%</td>
<td>NA</td>
<td>93%</td>
<td>84%</td>
</tr>
</tbody>
</table>

Abbreviations: NA, not available; NFP, Nurse-Family Partnership; TANF, Temporary Assistance for Needy Families.

Table 2. Characteristics of Nurse-Family Partnership Clients Served in Urban vs Rural Agency Locations

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>All NFP Clients (n=3844)</th>
<th>Urban Agency (n=3296)</th>
<th>Rural Agency (n=548)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aged ≤18 y</td>
<td>43%</td>
<td>45%</td>
<td>32%</td>
</tr>
<tr>
<td>African American</td>
<td>25%</td>
<td>29%</td>
<td>5%</td>
</tr>
<tr>
<td>≤12th grade education</td>
<td>85%</td>
<td>86%</td>
<td>80%</td>
</tr>
<tr>
<td>Smoker</td>
<td>41%</td>
<td>39%</td>
<td>51%</td>
</tr>
<tr>
<td>Unmarried</td>
<td>90%</td>
<td>91%</td>
<td>86%</td>
</tr>
<tr>
<td>Receiving food stamps</td>
<td>67%</td>
<td>68%</td>
<td>62%</td>
</tr>
<tr>
<td>Receiving TANF</td>
<td>53%</td>
<td>95%</td>
<td>14%</td>
</tr>
</tbody>
</table>

Abbreviations: NFP, Nurse-Family Partnership; TANF, Temporary Assistance for Needy Families.

RESULTS

From 2000 to 2005, the 23 sites providing NFP services throughout Pennsylvania enrolled a total of 5721 clients. Of the 94% of clients who matched to vital statistics records for their infants, 94% were singleton first pregnancies, of which 88% were welfare eligible in the year preceding the infant’s birth (Figure 1). Local-area matching was successful for 86% of this restricted sample of women (n=3844 clients), yielding a total of 10 938 controls.

Matching greatly reduced the selection differences between women enrolled in the NFP and other local-area controls (Table 1). In aggregate, the majority of NFP clients were white, young, from urban environments, or with less than a high school education (Table 1). These characteristics differed considerably from other women giving birth across Pennsylvania during the period and even compared with all welfare-eligible births. Propensity score matching was successful in balancing these selection factors among the matched controls as well as balancing the geographic representation across the high-density neighborhoods served by the 23 sites in the Commonwealth (eFigure, http://www.archpediatrics.com).

Demographic characteristics also differed considerably among women served by the 6 rural agencies compared with the 17 urban agencies (Table 2). Women served by rural locations were more likely to be white, to be married, to have completed some postsecondary education, and to smoke. Conversely, women served by urban agencies were more likely to be younger and to receive food stamps or Temporary Assistance for Needy Families within the portfolio of welfare services.

After an initial implementation period in which there was no difference in second pregnancies resulting in live births by 2 years (19.3% for NFP clients vs 18.9% for controls), fewer NFP clients (16.8%) than controls (19.0%) had such pregnancies by 2004 to 2005 (P=.05). A reduction in second pregnancies among mothers aged 18 years or younger from this latter period (17.9% for NFP clients vs 23.3% for controls; P=.003) was the principal contributor to this overall difference.

Kaplan-Meier curves were constructed to describe the time from the first infant’s birth to a second pregnancy across period, age groups, and rural and urban agency locations. A Cox proportional hazards model, clustered by agency catchment area, was used to evaluate the time to conception of a second pregnancy resulting in a live birth between clients and matched controls. Models were tested for the assumption of proportional hazards. A nested variable encoded the 3-way interaction of period, age group (≤18 years vs >18 years), and rural or urban agency location. This latter classification was based on the clients served by the agency. In all but 3 agencies, clients were entirely from urban or rural locations; in the 3 remaining agencies, 85% of clients were from an urban location, so the agencies were encoded as urban. At the same time, we also constructed a separate model including these 3 agencies only and using individual-level urban or rural designation as an interaction term with treatment status (client vs control) to evaluate the rural vs urban contrast within those agencies. Finally, Cox regression stratified on agency was also performed to check for the sensitivity of results to model specification. As no differences were observed from the clustered Cox model, only that latter model is reported.

Kaplan-Meier and Cox analyses were conducted using Stata version 10.0 and 11.0 statistical software (StataCorp LP, College Station, Texas). Approval for the study was granted by the Department of Public Welfare for the Commonwealth of Pennsylvania and the Institutional Review Board at the Children’s Hospital of Philadelphia.
second pregnancies resulting in live births between clients and controls, there was a 13% reduction in such pregnancies between NFP clients and controls who were enrolled in the program between 2004 and 2005 (hazard ratio [HR] = 0.87; 95% confidence interval [CI], 0.80-0.96). Second, during this period, the program effect was enhanced among mothers aged 18 years or younger, with a 27% reduction in the rate of such pregnancies compared with similarly aged controls (HR = 0.73; 95% CI, 0.61-0.89). Finally, younger mothers demonstrated greater benefit in rural agency locations compared with urban locations. Although younger mothers in programs in both locations had lower rates of second live births than their matched controls, the effect was twice as strong among younger women participating in rural locations (HR = 0.40; 95% CI, 0.22-0.73) compared with those in urban locations (HR = 0.79; 95% CI, 0.65-0.95). Even among the 3 urban agencies with 15% of clientele who were from rural locations, rural mothers tended to have a lower rate of second pregnancies resulting in live births across time than urban mothers at the same site (HR = 0.28; 95% CI, 0.05-1.63; P = .15).

**COMMENT**

Building on the evidence of earlier trials, this study reveals that following an initial start-up period across Pennsylvania in which program effects were muted, significant program-control differences in subsequent pregnancy outcomes began to emerge by 2004. The delay in achiev-
grams. With the exception of the original Elmira trial, surprising given the infrequent reports of enhanced prog- ining positive outcomes after an initial implementation launch period was also not unexpected and is consistent with interventions experience, which forecasts a start-up period in which program effectiveness is delayed as inter- ventions move from the controlled laboratory to community settings. Additionally, investment in monitoring fidelity and program standardization by the NFP National Service Office and its Pennsylvania state partner Public/Private Ventures—critical ventures for achieving successful implementation—have likely aided program maturation across the Commonwealth and reduced second pregnancies resulting in live births over time.

While a start-up implementation time lag related to outcomes was not unexpected, the particularly strong program effects demonstrated in rural locations—with a comparative 2-fold reduction in second pregnancies resulting in live births compared with urban locations—was surprising given the infrequent reports of enhanced program effects in rural areas from other home visitation programs. With the exception of the original Elmira trial, which occurred in a mixed urban and rural location, the other NFP trials were conducted principally in urban locations and therefore yielded insufficient inferences as a whole for rural locations. The Pennsylvania implementation, in contrast, included a large number of rural programs and rural participants. The risk reduction demonstra- nded in rural locations substantiates efforts to expand the program beyond additional urban locations that were well represented among program recipients in earlier trials.

This is not meant to suggest that urban locations in the Pennsylvania NFP program did not achieve successful outcomes. Rather, the rural effect demonstrated in this study surpassed what was already a significant effect in the urban locations. The reasons for such a difference are uncertain but likely multifactorial. For example, rural locations might have been better suited to meet community need (through smaller caseloads and greater community penetration) and also avoided the overwhelming caseloads (in need and volume) that often challenge urban practice settings. Nurses might have been aided by more informal family and community networks in rural locations that could have potentiated their benefit. It is also possible that nonparticipating controls in rural locations were less likely to have received other community-based interventions that might be more common in urban locations (eg, family planning or other home visitation programs). A reduced safety net within the rural community would therefore magnify the effect of the NFP program in such locations.

While the reduction in second pregnancies was particularly strong among younger women and those living in rural locations, in aggregate the reduction in second pregnancies during the second period was smaller compared with prior trial data. In the Elmira trial, nurse-visited, low-income, unmarried women had a mean of 0.17 subsequent pregnancies compared with 0.51 for their control counterparts within 22 months following birth of the first child; in the Memphis trial, 36% of the nurse-visited women had a second pregnancy by the first child’s second birthday compared with 47% in the control group; and in the Denver trial, 29% of those visited by nurses had a second pregnancy by the first child’s second birthday compared with 41% in the control group.

There are several explanations for the more muted aggregate effect in this study. A relatively high proportion of older mothers in our study (mean [SD] age, 19.6 [3.2] years), in whom program effect was expected to be lower, likely reduced the observed aggregate effect on second pregnancies. Also, there were fewer second pregnancies overall in our study, reflecting secular trends toward reduced second pregnancies in the overall population during this time. Finally, it is possible that remaining differences reflect the challenges following implementation, emphasizing the need for continuous quality improvement efforts to enhance program outcomes.

Several limitations merit additional caution in interpreting our findings. First, we describe only 1 of several outcome measures reported in prior controlled trials, so it is important to see whether program effects hold for other sentinel outcomes in replication sites as well. Second, our birth certificate data were limited in identifying only second pregnancies that resulted in a live birth of an infant. Prior trial data, however, suggest that terminated pregnancies or miscarriages either do not vary between recipients and controls or may be less common among NFP recipients, potentially

| Table 3. Hazard Ratios for Second Pregnancies Delivering by 2 Years Among Nurse-Family Partnership Clients Compared With Locally Matched Controls Across the 2000 to 2003 and 2004 to 2005 Periods |
|-----------------|-----------------|-----------------|-----------------|
| **Comparison Group** | **2000-2003** | **2004-2005** | **Across Time, Ratio of HRs (95% CI)** |
| Overall effect | 1.09 (0.95-1.26) | 0.87 (0.80-0.96) | 0.80 (0.66-0.96) |
| NFP clients vs controls | 1.07 (0.85-1.35) | 0.73 (0.61-0.89) | 0.69 (0.48-0.99) |
| Subgroup | 1.11 (0.95-1.29) | 1.00 (0.85-1.18) | 0.90 (0.74-1.11) |
| Mothers aged <18 y | 1.16 (0.91-1.49) | 0.83 (0.66-1.04) | 0.71 (0.47-1.08) |
| Rural mothers | 1.08 (0.93-1.27) | 0.88 (0.70-0.98) | 0.82 (0.67-1.00) |
| Urban mothers | 1.07 (0.83-1.38) | 0.79 (0.65-0.95) | 0.73 (0.50-1.08) |
| Nested subgroup | 1.04 (0.79-1.38) | 0.40 (0.22-0.73) | 0.39 (0.20-0.74) |
| Urban mothers aged <18 y | 1.09 (0.92-1.29) | 0.98 (0.81-1.20) | 0.90 (0.73-1.12) |
| Urban mothers aged >18 y | 1.23 (0.83-1.84) | 1.07 (0.66-1.04) | 0.87 (0.49-1.56) |

Abbreviations: CI, confidence interval; HR, hazard ratio; NFP, Nurse-Family Partnership.

The ratio of HRs is the effect of the intervention in the 2004 to 2005 period compared with that in the 2000 to 2003 period, overall and within subgroups.

The ratio of HRs is the effect of the intervention in the 2004 to 2005 period compared with that in the 2000 to 2003 period, overall and within subgroups.

The ratio of HRs is the effect of the intervention in the 2004 to 2005 period compared with that in the 2000 to 2003 period, overall and within subgroups.

The ratio of HRs is the effect of the intervention in the 2004 to 2005 period compared with that in the 2000 to 2003 period, overall and within subgroups.
understanding the differences we observed. \(^4,6,9,37,38\) Finally, while we used an analysis that balanced on selection factors likely to influence client participation in the NFP, it is possible that other selection factors remain unmeasured. If there was heavier participation of women with histories of receiving child welfare, being incarcerated, or having other high risks within the NFP program, for example, such unmeasured selection factors might falsely reduce the differences in outcomes we observed within and across agencies. Alternatively, if those women who chose to participate in the NFP program had higher self-efficacy to participate than those who might have refused to participate, our results could overestimate the benefit of the program. However, it is reassuring that in supplementary analyses (not shown), we did not observe enhanced and potentially biased program outcomes in neighborhoods with high NFP participation, where the higher likelihood of refusers among the control population would have exacerbated any selection effect that might have been present.

Nevertheless, despite concerns of unmeasured selection factors, these data provide evidence that the NFP continues to be successful after statewide implementation. The finding of potentiated benefit in rural locations seems to justify implementation in such locations in addition to urban areas, greatly expanding the reach of the program. As that expansion occurs, investments in infrastructure to measure fidelity and ensure that outcomes continue to improve across time and sites will be critical to the development of better programs and even better outcomes for women and children over time.

Accepted for Publication: August 20, 2010.
Published Online: November 1, 2010. doi:10.1001/archpediatrics.2010.221

Correspondence: David M. Rubin, MD, MSCE, Division of General Pediatrics, Children’s Hospital of Philadelphia, 34th Street and Civic Center Boulevard, CHOP North, Room 1533, Philadelphia, PA 19104-4399 (rubin@email.chop.edu).

Author Contributions: Study concept and design: Rubin, O’Reilly, Luan, Localio, and Christian. Acquisition of data: Rubin, O’Reilly, Luan, and Dai. Analysis and interpretation of data: Rubin, O’Reilly, Luan, Dai, Localio, and Christian. Drafting of the manuscript: Rubin and Localio. Critical revision of the manuscript for important intellectual content: Rubin, O’Reilly, Luan, Dai, Localio, and Christian. Statistical analysis: Rubin, Luan, Dai, Localio, and Christian. Obtained funding: Rubin. Administrative, technical, and material support: Rubin, O’Reilly, and Luan. Study supervision: Rubin, Luan, and Christian.

Financial Disclosure: None reported.

Funding/Support: This project was supported by a grant from the Department of Public Welfare, Commonwealth of Pennsylvania.

Online-Only Material: The eFigure is available at http://www.archpediatrics.com.

Additional Contributions: David Olds, PhD, and Kathleen Noonan, JD, provided critical review during the preparation of the manuscript. We thank the Pennsylvania Department of Public Welfare, the Pennsylvania Department of Health, the 23 NFP agencies across Pennsylvania, and Public/Private Ventures for contributing data for this study.

REFERENCES


---

**Announcement**

The Archives is introducing a rolling theme issue on comparative effectiveness research for the care of children and adolescents. We are interested in comparative effectiveness research across the spectrum of care that child health care providers deliver. Priority will be given to studies that use rigorous methodological designs, are generalizable well beyond an individual institution’s walls, and are real-world effectiveness studies. Please consult our Website at http://www.archpediatrics.com for instructions on manuscript preparation and submission.