



SCHOOL of
PUBLIC POLICY

21

Nurse Family Partnership (Elmira)

Douglas J. Besharov
Peter Germanis
Caeli A. Higney
and
Douglas M. Call

September 2011



Maryland School of Public Policy
Welfare Reform Academy
www.welfareacademy.org

Part of a forthcoming volume
Assessments of Twenty-Six Early Childhood Evaluations
by Douglas J. Besharov, Peter Germans, Caeli A. Higney, and Douglas M. Call

Note: This report is open to public comments, subject to review by the forum moderator. To leave a comment, please send an email to welfareacademy@umd.edu or fill out the comment form at http://www.welfareacademy.org/pubs/early_education/chapter21.html.

21

Nurse Family Partnership (Elmira)

The Nurse Family Partnership (NFP) operated in Elmira, New York, a semi-rural area, from 1978 to 1982. It was designed to “help low-income, first-time parents start their lives with their children on a sound course and prevent the health and parenting problems that can contribute to the early development of antisocial behavior.”¹ The program had three main objectives: (1) to improve women’s health-related behaviors during pregnancy; (2) to aid parents in the attainment of parenting skills; and (3) to enhance the maternal life-course development of participating women by encouraging family planning, educational development, and self-sufficiency. This intervention was unique in that it used nurses on home visits to direct mothers toward specific behaviors.

David Olds, now professor of pediatrics at the University of Colorado Health Sciences Center and creator of the nurse-visitor design, and his colleagues (the “NFP team”) evaluated the NFP using random assignment that took place between 1978 and 1980. They reported large reductions in welfare use among participating mothers and delinquency rates among participating children, particularly for the subsample of low socioeconomic status (SES) families. Olds also conducted a benefit-cost analysis when the children reached age four, which was updated by Lynn A. Karoly and her colleagues at RAND. This analysis suggests that, after fifteen years, the program saved \$4 for every dollar spent. The savings were driven by reductions or delays in second births to the mothers and concomitant reductions in the receipt of welfare and food stamps among mothers. As the generalizability of the findings is an open question, the NFP team has recognized the need for replication of their findings. They have embarked on a strategy to examine the program’s effects in different settings and support independent evaluations by other researchers. One remaining concern about the evaluation is that it was conducted by the same group that designed the intervention and has yet to be independently evaluated.

Program Design

Program group. The NFP was targeted to young women (age nineteen and younger), who were less than twenty-five weeks pregnant, had no previous live births, were unmarried, and

¹David Olds, Peggy Hill, and Elissa Rumsey, “Prenatal and Early Childhood Nurse Home Visitation,” *Juvenile Justice Bulletin* (November 1998a): 1.

were of low-SES. Although designed for women with various sociodemographic risks, the program was open to any childless woman to avoid stigmatizing participants. At the time of enrollment, the average age of the women was nineteen, 62 percent were unmarried, 59 percent were of low-SES, and about 90 percent were white.²

Services. The program consisted of home visits made by trained, experienced nurses, beginning during pregnancy and continuing until the child reached two years. The nurse-visitors followed detailed program guidelines that addressed both the mother's personal health and development, as well as the development of the child. The services included parent education, social support, and referrals to other health and social services.

The parent education component focused on informing parents about fetal and infant development. Under the program, the nurses gave the mothers specific direction and behavioral guidance. The nurses helped mothers to improve their health behaviors that affect child well-being. They also encouraged mothers to finish school and to find employment, in addition to discouraging additional childbearing. They worked to strengthen the informal supports available to the women during pregnancy, birth, and while their children were young. They did this by encouraging boyfriends (or husbands), close friends, and relatives "to participate in the home visits, to help with household responsibilities, to accompany the women to the hospital at the time of delivery, to be present for the birth, to aid in the subsequent care of the child, and to reinforce the advice of the nurses in their absence."³ The nurses also referred participating women to health and human services agencies for family planning, mental-health counseling, and legal aid services as needed.

²Unless otherwise noted, all findings come from the following references: David L. Olds, Charles R. Henderson, Jr., Robert Chamberlin, and Robert Tatelbaum, "Preventing Child Abuse and Neglect: A Randomized Trial of Nurse Home Visitation," *Pediatrics* 78, no. 1 (July 1986): 65–78; David L. Olds, Charles R. Henderson, and Harriet Kitzman, "Does Prenatal and Infancy Nurse Home Visitation Have Enduring Effects on Qualities of Parental Caregiving and Child Health at 25 to 50 Months of Life?" *Pediatrics* 93, no. 1 (January 1994): 89–98; David Olds, Charles R. Henderson, Jr., Robert Cole, John Eckenrode, Harriet Kitzman, Dennis Luckey, Lisa Pettitt, Kimberly Sidora, Pamela Morris, and Jane Powers, "Long-term Effects of Nurse Home Visitation on Children's Criminal and Antisocial Behavior: 15-Year Follow-up on a Randomized Controlled Trial," *Journal of the American Medical Association* 280, no. 14 (October 14, 1998b): 1238–1244; and John Eckenrode, Mary Campa, Dennis W. Luckey, Charles R. Henderson, Jr., Robert Cole, Harriet Kitzman, Elizabeth Anson, Kimberly Sidora-Arcoleo, Jane Powers, and David Olds, "Long-term Effects of Prenatal and Infancy Nurse Home Visitation on the Life Course of Youths: 19-Year Follow-up for a Randomized Trial," *Archives of Pediatrics & Adolescent Medicine* 164, no. 1 (January 2010): 9–15.

³David Olds, Charles R. Henderson, Jr., Harriet Kitzman, John Eckenrode, Robert Cole, Robert Tatelbaum, JoAnn Robinson, Lisa M. Pettitt, Ruth O'Brien, and Peggy Hill, "Parental and Infancy Home Visitation by Nurses: A Program of Research" (paper, Preventing Second Births to Teenage Mothers: Demonstration Findings conference, American Enterprise Institute, Washington, DC, March 6, 1998), <http://www.welfareacademy.org/conf/papers/olds/prenatal.htm> (accessed July 23, 2010).

The Evaluation. The NFP team conducted the random assignment evaluation. Five hundred women who met the project's criteria were recruited from prenatal clinics and the offices of private obstetricians. Four hundred (80 percent) enrolled in the research project before the thirtieth week of pregnancy and were randomly assigned between 1978 and 1980. There were no statistically significant differences in age, marital status, or educational attainment between the women who participated and those who refused. Nonwhite women, however, were more likely to enroll than white women (94 percent vs. 80 percent).

Participants were stratified by marital status, race and geographic region, and then randomly assigned to one of four groups. The first group received health and developmental screening when their children were ages one and two, with appropriate referral to services. The second group received the same screening but also received free transportation to health care. These two groups were later combined because there were no differences in the use of prenatal and well-child care when the child was age two. These two groups then constituted the control group. There were two intervention groups: one received home visits during pregnancy and the other received visits during pregnancy and for the first two years of the child's life. The findings reported here are generally based on the outcomes for this latter group compared to the control group.

Data were obtained from interviews with mothers and their children and various administrative records, including child abuse, welfare receipt, and criminal records.

Major Findings

The NFP team reported large reductions in welfare use among participating mothers and delinquency rates among participating children, particularly for the subsample of high-risk (unmarried and low-SES) families.⁴ A benefit-cost analysis suggests that, after fifteen years, the program saved \$4 for every dollar spent, the total savings were \$23,439 for every high-risk family.

Cognitive. There were no IQ gains (based on the Stanford-Binet test) for the full sample or the high-risk sample when the children were ages three and four. As a result, the NFP team concluded:

The absence of an enduring program effect of children's intelligence suggests that if the program is going to have any influence on the children's school achievement and adaptation it will be through parents' promotion of high standards, hard work, and

⁴David Olds, Charles R. Henderson, Jr., Robert Cole, John Eckenrode, Harriet Kitzman, Dennis Luckey, Lisa Pettitt, Kimberly Sidora, Pamela Morris, and Jane Powers, "Long-term Effects of Nurse Home Visitation on Children's Criminal and Antisocial Behavior: 15-Year Follow-up on a Randomized Controlled Trial," *Journal of the American Medical Association* 280, no. 14 (October 14, 1998b): 1238.

discipline rather than through the promotion of specific cognitive skills.⁵

After conducting further subgroup analyses, the NFP team found a statistically significant positive program effect on the intellectual development of children ages three and four born to women who smoked ten or more cigarettes per day during pregnancy. This effect even held for children whose mothers were visited only during pregnancy.⁶

School readiness/performance. At age nineteen, youth self-reports do not show any statistically significant effects on high school graduation. The researchers note, “Nineteen-year-old youths may simply be too young to observe patterns of educational achievement that include college attendance, or the program may not have affected educational outcomes for this population of youths in this semirural setting.”⁷

Socioemotional development. There were no program effects on parents’ reports of their children’s behavioral problems at child age two, three, or four.

Health. Children born to mothers who smoked five or more cigarettes a day at registration and visited by nurses during pregnancy were 75 percent less likely to be delivered preterm (defined as less than thirty-seven weeks of gestation) than were their counterparts in the control group. During the first four years of the child’s life, nurse-visited children had fewer emergency room encounters and visits to the physician in which injuries or ingestions were detected.⁸

Behavior. At age fifteen, there were no statistically significant impacts for the entire sample on sixteen different measures of antisocial behavior. Subgroup analyses, however, revealed impacts on adolescents in the program group born to high-risk mothers. These young people had fewer sex partners (0.92 vs. 2.48), reported drinking alcohol fewer days (1.09 vs. 2.49 in the

⁵David L. Olds, Charles R. Henderson, and Harriet Kitzman, “Does Prenatal and Infancy Nurse Home Visitation Have Enduring Effects on Qualities of Parental Caregiving and Child Health at 25 to 50 Months of Life?” *Pediatrics* 93, no. 1 (January 1994): 97.

⁶David L. Olds, Charles R. Henderson, and Robert Tatelbaum, “Prevention of Intellectual Impairment in Children of Women Who Smoke Cigarettes During Pregnancy,” *Pediatrics* 93, no. 2 (February 1994): 231.

⁷John Eckenrode, Mary Campa, Dennis W. Luckey, Charles R. Henderson, Jr., Robert Cole, Harriet Kitzman, Elizabeth Anson, Kimberly Sidora-Arcoleo, Jane Powers, and David Olds, “Long-term Effects of Prenatal and Infancy Nurse Home Visitation on the Life Course of Youths: 19-Year Follow-up for a Randomized Trial,” *Archives of Pediatrics & Adolescent Medicine* 164, no. 1 (January 2010): 14.

⁸David L. Olds, Charles R. Henderson, and Harriet Kitzman, “Does Prenatal and Infancy Nurse Home Visitation Have Enduring Effects on Qualities of Parental Caregiving and Child Health at 25 to 50 Months of Life?” *Pediatrics* 93, no. 1 (January 1994): 89–98.

preceding six months), and ran away less often (0.24 vs. 0.60). They experienced no statistically significant effects on the remaining behavioral measures.

At age nineteen, youth self-reports continued to suggest that there were no statistically significant impacts for a range of measures of antisocial behavior, but as with the age fifteen follow-up, there were impacts on youth in the program group born to high-risk mothers. The youth were more likely to use condoms. Girls were significantly less likely to have had a child (11 percent vs. 30 percent) and were also less likely to use Medicaid. However, the boys born to these high-risk mothers reported an average higher number of sexual partners (3.07 vs. 1.53) than the control group.⁹

Crime/delinquency. There were statistically significant reductions in criminal activity among adolescents born to high-risk mothers at the time of registration in the study, and who received nurse home visits during pregnancy and infancy—compared with their control-group counterparts. Children’s self-reports at age fifteen showed a lower incidence of arrests (0.20 vs. 0.45) and a much lower incidence of convictions and probation violations (0.47 vs. 0.9).¹⁰ These findings were also statistically significant for the full program sample.

There were inconsistencies among some measures, however, depending on the data source. The NFP team cautions that “the arrest and conviction data were based primarily on the children’s and parents’ reports, which may be subject to treatment-related reporting bias.”¹¹ For example, the incidence of arrests reported by mothers, while statistically significant, was considerably lower (0.04 vs. 0.19). And, there was no statistically significant difference in arrests among those who had lived in the same county since birth. This latter finding, however, may have occurred because the nurse-visited mothers who remained in the county were more disadvantaged on average than control group mothers, which tended to understate the differences in arrests. In addition, self-reports by adolescents themselves suggested little difference in major delinquent acts (3.99 vs. 4.09, a difference that is not statistically significant). The NFP team reports that the intervention was not as successful in reducing relatively minor delinquent behavior that was not serious enough to lead to arrest.

The group that received just prenatal home visits experienced slightly larger reductions in the incidence of arrests, convictions, and probation violations, implying that home visits during

⁹John Eckenrode, Mary Campa, Dennis W. Luckey, Charles R. Henderson, Jr., Robert Cole, Harriet Kitzman, Elizabeth Anson, Kimberly Sidora-Arcoleo, Jane Powers, and David Olds, “Long-term Effects of Prenatal and Infancy Nurse Home Visitation on the Life Course of Youths: 19-Year Follow-up for a Randomized Trial,” *Archives of Pediatrics & Adolescent Medicine*, 164, no. 1 (January 2010): 13.

¹⁰Olds et al., 1998b, 1238.

¹¹Olds et al., 1998b, 1243.

infancy and early childhood had no added impact on delinquency.¹² There is, however, a large discrepancy in impacts for this group depending on whether the number of arrests is based on self-report (0.45 vs. 0.15) or on the mother's report (0.16 vs. 0.19).

At age nineteen, youth self-reports indicate that there were statistically significant reductions in criminal activity among youth born to mothers who received nurse home visits during pregnancy and infancy-compared with their control group counterparts. These effects, however, were limited to girls who were less likely to be arrested (10 percent vs. 30 percent) and convicted (4 percent vs. 20 percent.) in the preceding year. There were no statistically significant differences for boys.¹³

Early/nonmarital births. Data apparently either not collected or not reported.

Economic outcomes. Data apparently either not collected or not reported.

Effects on parents. For the high-risk group, at the fifteen-year follow-up there were statistically significant reductions in subsequent childbearing (1.1 vs. 1.6 additional children), welfare receipt (60 vs. 90 months), food stamp receipt (47 vs. 84 months), behavioral impairments due to substance use (0.41 vs. 0.73 incidences), and arrests (0.18 vs. 0.58 self-reported arrests).¹⁴ A sixteen-month increase in months employed (96 months vs. 80 months) was not statistically significant. These mothers were also less likely to be identified as perpetrators of child abuse and neglect (0.11 vs. 0.53 verified reports). The program impact on child abuse and neglect extended to the full sample of women (0.29 vs. 0.54 verified reports), but none of the other aforementioned positive impacts extended to the full sample.

Benefit-cost findings. The NFP team conducted a benefit-cost analysis of the home visiting program based on outcomes when the children were age four.¹⁵ All costs and benefits are

¹²This finding does not rule out the possibility that home visits during infancy and early childhood, but not during pregnancy, could also have an impact on these measures.

¹³John Eckenrode, Mary Campa, Dennis W. Luckey, Charles R. Henderson, Jr., Robert Cole, Harriet Kitzman, Elizabeth Anson, Kimberly Sidora-Arcoleo, Jane Powers, and David Olds, "Long-term Effects of Prenatal and Infancy Nurse Home Visitation on the Life Course of Youths: 19-Year Follow-up for a Randomized Trial," *Archives of Pediatrics & Adolescent Medicine* 164, no. 1 (January 2010): 13.

¹⁴David L. Olds, John Eckenrode, Charles R. Henderson, Jr., Harriet Kitzman, Jane Powers, Robert Cole, Kimberly Sidora, Pamela Morris, Lisa M. Pettit, and Dennis Luckey, "Long-term Effects of Home Visitation on Maternal Life Course and Child Abuse and Neglect: 15-Year Follow-up of a Randomized Trial," *Journal of the American Medical Association* 278, no. 8 (August 27, 1997): 637-643.

¹⁵David L. Olds, Charles R. Henderson, Charles Phelps, Harriet Kitzman, and Carole Hanks, "Effect of Prenatal and Infancy Nurse Home Visitation on Government Spending," *Medical Care* 31, no. 2 (February 1993):

reported in 2005 dollars and discounted at a 3 percent rate.

Costs. The NFP team examined several cost categories: the direct costs of nurse home visits (such as salaries, travel and supplies, and overhead); the costs of other services (such as the Special Supplemental Nutrition Program for Women, Infants, and Children [WIC], to which participants were referred); and transportation costs. The net cost of the nurse home visiting program was calculated as the per-family difference between the program group cost and the control group cost. The average, per-family incremental cost was estimated at \$7,784 for the sample as a whole and \$7,513 for low-income mothers. These incremental costs consisted primarily of the costs of the nurse program, since there was little difference in the use of community services and transportation costs were very low.¹⁶

Benefits. Estimated benefits focused on improvements in parental outcomes that produced savings. Benefits to children were excluded because the cost-benefit analysis was performed when the children were age four. Thus, no savings were observed in school-related impacts (such as grade retention or special education placement) or in the children's own subsequent employment, welfare receipt, or other outcomes.

The estimated benefits for the entire sample was \$3,990 per family and \$7,945 per low-income family. As a result, the net cost (that is, total cost minus savings) was \$3,794 per family for the sample as a whole. The low-income sample produced a net savings of \$432 per family.

Among low-income mothers, nurse-visited mothers reported fewer months of welfare receipt.¹⁷ Although data on food stamps and Medicaid enrollment were limited, the NFP team estimated savings to these programs by relying on available data. Reductions in welfare and food stamp costs accounted for 82 percent of total estimated savings. The remaining 18 percent of savings was attributed to reductions in Medicaid spending and the costs associated with child abuse and neglect (for example, initial investigation, ongoing supervision, and foster care), and the increase in tax revenues. They based estimates of tax revenue gains on increases in employment and estimated earnings.

A large part of these savings was achieved by reductions or delays in subsequent childbearing. The NFP team observes that the prospects for long-term net savings depended

155–174.

¹⁶The NFP team noted that these costs may be somewhat higher than future implementations, since part of the cost involved development of the home-visit protocol.

¹⁷Most welfare data came from administrative records. These were supplemented with survey data only when individuals moved out of state or when the interview data produced higher rates of Welfare use than did the administrative data.

heavily on whether the reduction in subsequent pregnancy results in fewer lifetime births for the women in the project.

A reduction in completed family size would augur for continued cost-savings to government over time, while an increase in child spacing without a decrease in completed family size would produce long-term savings at a reduced rate, and most likely only if the increased spacing allowed women to find and retain employment.¹⁸

Some uncertainties remain with the estimates of savings.¹⁹ Moreover, the finding that the program for high-risk families pays for itself within four years is unusual, compared to the estimated savings associated with other early childhood intervention programs, which tend to produce savings sufficient to offset costs only decades after the intervention.

RAND update. Several years later, Lynn Karoly and her colleagues at RAND updated the benefit-cost analysis to include the benefits accrued to the participants through the age of fifteen.²⁰ They followed procedures similar to those used by the NFP team, but were able to build in savings from additional outcomes. (All dollar values are converted to 2005 dollars and discounted at a 4 percent rate.) As with the earlier work by the NFP team, Karoly and her colleagues report costs and savings as the difference between the program and control groups. They also estimate the costs separately for “lower-risk” and “higher-risk” families, that is families in which the mother was unmarried and from a low SES family at registration (high-risk) compared to the remainder of the sample (low-risk).

Karoly and her colleagues estimated the discounted cost of the nurse home visiting program at \$7,661 per family (about the same as the figure derived by the NFP team after adjusting for inflation). The longer follow-up period, however, made it possible to see whether savings continued to accrue and to include savings from other outcomes. Table 1 summarizes the costs and benefits for both higher- and lower-risk families.

For higher-risk families, the difference in costs (\$7,661) and savings (\$31,100) was \$23,439, or about \$4 saved for every dollar spent (in 2005 dollars). Most of the savings (80 percent) were attributable to changes in the mother’s behavior, related to increased employment

¹⁸David L. Olds, Charles R. Henderson, Charles Phelps, Harriet Kitzman, and Carole Hanks, “Effect of Prenatal and Infancy Nurse Home Visitation on Government Spending,” *Medical Care* 31, no. 2 (February 1993): 165.

¹⁹These caveats are described in detail in Olds et al., 1993, 155–174.

²⁰Lynn A. Karoly, Peter W. Greenwood, Susan S. Everingham, Jill Hoube, M. Rebecca Kilburn, C. Peter Rydell, Matthew Sanders, and James Chiesa, *Investing in Our Children: What We Know and Don’t Know About the Costs and Benefits of Early Childhood Interventions* (Santa Monica, CA: RAND, 1998).

and reduced welfare receipt. In contrast, the program did not produce net savings for lower-risk mothers, with just 62 cents saved for each dollar spent. Karoly and her colleagues also caution that “the uncertainty caused by small sample sizes must be considered.”²¹ They estimate that the true savings for the higher-risk had a two-thirds chance of being between about \$24,000 and \$36,000. Using the more traditional 95 percent confidence interval produced a range of approximately \$18,000 to \$42,000. Such wide intervals suggest that considerable caution should be used in making claims about cost effectiveness.²²

In addition, the RAND analysis involved a number of assumptions and projections, reducing the certainty of the findings. For example, Karoly and her colleagues estimated the amount saved due to decreased criminal activity of the child in the future. Nevertheless, the findings are consistent with the earlier, short-term findings found by the NFP team. In a 2000 update, the NFP team adds:

Given that welfare reform has changed the lifetime limits that women may receive welfare, the government savings in the welfare arena will be smaller than those estimated from the Elmira study. It is important to note, however, that the Rand analysis did not monetize the savings caused by the reduction in rates of subsequent pregnancy and child abuse and neglect, so in this way the analysis has underestimated the likely benefits.²³

²¹Karoly et al. 1998, 94.

²²Authors’ estimate based on Karoly et al. 1998, figure 3.5.

²³David Olds, Peggy Hill, JoAnn Robinson, Nancy Song, and Christina Little, “Update on Home Visiting for Pregnant Women and Parents of Young Children,” *Pediatrics* 30, no. 4 (April 2000): 116.

Table 1. Prenatal/Early Infancy Project: Benefit-Cost Findings

	Higher-risk families	Lower-risk families
Savings		
Reduction in emergency room visits	\$145	\$135
Increase in taxes	\$7,157	\$1,441
Decrease in welfare	\$17,716	\$1,599
Decrease in arrests/jail	—	—
Mother	\$799	\$140
Child	\$5,283	\$1,440
Total benefits (savings)	\$31,100	\$4,754
Program cost	\$7,661	\$7,661
Net present value	\$23,439	-\$2,907
Benefit-cost ratio	\$4.10/1	\$0.62/1

Source: Lynn A. Karoly, Peter W. Greenwood, Susan S. Everingham, Jill Hoube, M. Rebecca Kilburn, C. Peter Rydell, Matthew Sanders, and James Chiesa, *Investing in Our Children: What We Know and Don't Know About the Costs and Benefits of Early Childhood Interventions* (Santa Monica, CA: RAND, 1998).

Note: Dollar amounts converted to 2005 dollars, discounted at 4 percent.

Overall Assessment

The NFP project was carefully evaluated using random assignment. The NFP team has recognized the need for replication of the findings. They have embarked on a strategy to examine the program's effects in different settings (see chapters 19–20, this volume) and have supported independent evaluations by other researchers.

Program theory. According to Olds, the conceptual framework guiding the design of the program and its evaluation was based on evidence that “suggests that parental behavior is the most immediate, powerful, and potentially alterable influence on child health during pregnancy and the early years of the child’s life.”²⁴ Through nurse visits, this program sought to improve women’s health-related behaviors, parenting skills, and personal development. The evaluation, designed to assess improvements in child health and development, parental health-related behaviors, and qualities of parenting skills, was appropriate within this context.

Program implementation. No serious implementation problems were reported. Nurses

²⁴Olds, 1992, 705.

completed an average of nine visits (range of zero to eighteen) during pregnancy and twenty-three (range of zero to fifty-nine) from birth to the child's second birthday.²⁵

Assessing the randomization. Although conceived as a random assignment experiment, there were two departures from random assignment. First, six women who enrolled were living with another woman in the household who was already in the program. These women were assigned to the same intervention status as their housemate. (An alternative would have been not to randomly assign such women at all.)

In addition, during the last six months of the thirty-month enrollment period, the sampling ratio was changed to boost enrollment in the program group and increase the statistical power of the design. This adjustment could have introduced bias if the women in this later period were systematically different than those assigned in earlier periods.²⁶ The NFP team notes, "Analysis of selected dependent variables confirmed that this slight confounding of treatments with time did not alter the pattern of treatment effects."²⁷

Despite these departures, women in the nurse-visited and control groups were comparable on a range of sociodemographic characteristics. The proximity of many of the mothers, however, may have led to some problems with "crossover." David Greenberg, professor of economics at the University of Maryland-Baltimore County, and Mark Shroder, economist at the U.S. Department of Housing and Urban Development, explain in a 1997 analysis:

Of the sample (both treatment and control) 20-25 percent reported that they discussed pregnancy or child-care matters with nurse-visited women. In other words, people in groups 3 and 4 talked about the program with women in groups 1 and 2. This could be a source of horizontal diffusion.²⁸

²⁵Olds et al., 2000, 111. According to Olds, "A significant number of . . . home visits were cancelled, forgotten, or ignored, which was often very frustrating. Our nurses found that they sometimes had to overcome a lot of initial skepticism or resistance on the part of the parents. A number of parents were apathetic to toward what the nurses were doing and the materials they were presenting." David Olds, "An Intervention Program for High-Risk Families," in *A Round Table on Minimizing High-Risk Parenting*, ed. Robert A. Hoekelman (Media, PA: Harwal Publishing, 1983), 260.

²⁶The NFP team found no such differences. David Olds, University of Colorado Health Sciences Center, e-mail message to Peter Germanis, May 17, 2001.

²⁷David L. Olds, Charles R. Henderson, Jr., Robert Chamberlin, and Robert Tatelbaum, "Preventing Child Abuse and Neglect: A Randomized Trial of Nurse Home Visitation," *Pediatrics* 78, no. 1 (July 1986): 68.

²⁸David Greenberg and Mark Shroder, *The Digest of Social Experiments* (Washington, DC: The Urban Institute Press, 1997), 239.

Such diffusion of information would tend to reduce measured impacts.

In addition, pregnant women in the control group were offered free screenings for their children for sensory and developmental problems, and those with possible problems were referred to other specialists. Some also received free transportation by taxi to regular prenatal and well-child care visits. The NFP team points out, “The transportation and screening may have elevated the functioning of the comparison group and thus reduced the range over which the program might produce positive effects.”²⁹

Assessing statistical controls in experimental and nonexperimental evaluations. The evaluation was based on random assignment, so selection bias should not be a serious problem. Additionally, the statistical model used controlled for interactions among six covariates: maternal age, education, locus of control, husband or boyfriend support, mother’s employment status, and father’s public-assistance status, all measured at registration.

Sample size. Although the overall sample of the Elmira intervention was relatively large at about 400 families, the sample of high-risk women, which has the most significant findings, is much smaller at just 100 families (with 38 in the program group and 62 in the control group). With a small sample, large impacts are needed for statistically significant findings. Thus, the absence of impacts in some areas does not mean that the program did not affect some outcomes, but that the impacts may have been too small to be detected with the sample size. A small sample also means that differences in baseline characteristics would have had to be very large to be statistically significant, making it more difficult to assess the comparability of the program and control groups.

Attrition. Attrition was very low throughout the project. After fifteen years, 81 percent of women completed an assessment, as did 79 percent of the participating children, now adolescents. There were no statistically significant differences in attrition across the program and control groups in terms of either the overall rate of attrition or the characteristics of those dropping out. According to the NFP team:

For those families for which 15-year assessments were completed, the treatment groups were essentially equivalent on background characteristics for both the sample as a whole and for women who were unmarried and from high-risk households. Small differences on some background variables (such as paternal receipt of public assistance) led us to include them as covariates.³⁰

²⁹Olds et al., 1993, 165.

³⁰Olds et al., 1998, 1241.

At age nineteen, 78 percent of the youth participated in the follow-up survey, representing 88 percent of those where the child was still alive.³¹

Data collection. The data collection relied on a various standardized tests, administrative data, and parent surveys.

Measurement issues. Many of the outcomes measured in the NFP were validated by comparisons across two or more data sources. Olds and Henderson, for example, describe how the various data sources with respect to child abuse and neglect reinforced each other:

The second reason for caution in interpretation is that these findings may be the result of systematic reporting bias. For each of the sources of data, one can construct an alternative explanation for specific findings. For example, the nurse-visited women may have been taught by the nurses to give more socially desirable answers and to behave more appropriately in front of the interviewers. Their friends, neighbors, and relatives may have been more reluctant to report them for abuse and neglect than in the case of women in the comparison group. The picture of superior caregiving for the highest risk nurse-visited women over their comparison-group counterparts comes from a variety of sources, however, and it is unlikely that each of these sources could have its own idiosyncratic bias and still produce the same pattern of results overall.³²

For some important outcomes, such as welfare receipt and employment histories, the main source of information is self-reports.³³ Given the role these outcomes play in the benefit-cost analysis, it would be helpful for future replications to include validation from administrative records. Indeed, the early findings from the Memphis replication are based on administrative records (see chapter 19).

Generalizability. Olds summarizes the generalizability of the findings as follows:

³¹John Eckenrode, Mary Campa, Dennis W. Luckey, Charles R. Henderson, Jr., Robert Cole, Harriet Kitzman, Elizabeth Anson, Kimberly Sidora-Arcoleo, Jane Powers, and David Olds, "Long-term Effects of Prenatal and Infancy Nurse Home Visitation on the Life Course of Youths: 19-Year Follow-up for a Randomized Trial," *Archives of Pediatrics & Adolescent Medicine*, 164, no. 1 (January 2010): 10.

³²David L. Olds and Charles R. Henderson, "The Prevention of Maltreatment," in *Child Maltreatment: Theory and Research on the Causes and Consequences of Child Abuse and Neglect*, ed. Dante Cicchetti and Vicki Carlson (Cambridge, MA: Cambridge University Press, 1989), 752.

³³The NFP team uses administrative records whenever possible. Through child-age four, they used AFDC administrative records, but by the time the child was fifteen, many of the records had been destroyed. At age fifteen, the researchers did use data from the New York State Division of Criminal Justice System on maternal arrests and convictions. These data indicated larger impacts than self-reported data. David Olds, University of Colorado Health Sciences Center, e-mail message to Peter Germanis, October 24, 2003.

In interpreting the findings from the Elmira trial, it is important to keep in mind that the results were derived from one study carried out in a small, semirural community with a white sample in the late 1970s and early 1980s. We do not yet know whether the findings apply to minorities living in major urban areas in the 1990s.³⁴

Moreover, the positive impacts were concentrated among children born to high-risk women. Importantly, the program was limited to women having their first child. This limits the generalizability of the findings, as Olds and Henderson explain:

We reasoned that women having first children would be more receptive to the nurses' offers of help, that the skills and resources that women developed in coping with their first pregnancies and children would be carried over to subsequent childbearing and childrearing experiences, and that reducing future pregnancies and returning to school and work would be more feasible if we limited the program to women having first children. By limiting the program to first-time parents, we reasoned that we would increase the long-term impact of the program and its potential cost-effectiveness."³⁵

Replication. The nurse visiting program has been replicated in Memphis, with promising results, by the NFP team (see chapter 19), as well as in Denver (see chapter 20). But, further replication and evaluation by an independent evaluator following the program protocol is highly desirable.

Home visiting using paraprofessionals or other protocols has been widely tested. One comprehensive review of home visiting programs found that such programs have been less effective than the nurse home visiting model applied in Elmira.³⁶ Thus, it is important to bear in mind that the program model is very important. As Richard Berk, a professor of Criminology and Statistics at the University of Pennsylvania, commented, "Replications of a given evaluation may be used to incrementally define the boundaries within which generalization is possible."³⁷

Evaluator's description of findings. The NFP team has been careful in describing their findings, cautioning readers that the positive impacts were limited to low-income, unmarried, and

³⁴David L. Olds, "Home Visitation for Pregnant Women and Parents of Young Children," *American Journal of Diseases of Children* 146, no. 6 (June 1992): 706.

³⁵Olds and Henderson, 1989, 735.

³⁶See David L. Olds and Harriet Kitzman, "Review of Research on Home Visiting for Pregnant Women and Parents of Young Children," *The Future of Children: Home Visiting: Recent Program Evaluations* 3, no. 3 (Winter 1993): 53-92, <http://www.futureofchildren.org> (accessed July 11, 2002).

³⁷Richard A. Berk and Peter H. Rossi, *Thinking About Program Evaluation 2*, (Thousand Oaks, CA: Sage Publications, 1999), 25.

generally white mothers in a semirural area. They have suggested that their findings may represent an upper-bound estimate:

In evaluating the results of the Elmira trial, it is important to note that this was an efficacy trial, given that the investigators were intensely involved in monitoring the implementation of the program, and the same set of nurses worked with their families for the entire duration of the program. This means that the results obtained in Elmira are probably an upper-bound estimate of what might be accomplished with a program like this if it were disseminated on a large scale.³⁸

They also describe in detail the limitations of their analysis and call for further replication in other kinds of sites.

Evaluator's independence. The NFP team both developed and evaluated the NFP model. They support independent evaluations of the model, however, assuming they would be implemented with fidelity. Moreover, the evaluation findings have undergone extensive critical review and have been published in high-quality peer-reviewed journals.

Statistical significance/confidence intervals. Statistical significance was reported for variables with p values at 5 percent or lower.

Effect sizes. Apparently effect sizes were either not calculated or not reported.

Sustained effects. The evaluation examined the impacts through nineteen years after the birth of the women's first children.

Benefit-cost analysis. A benefit-cost analysis was conducted by the researchers, as well as in a separate analysis by researchers at the RAND corporation.

Cost-effectiveness analysis. Apparently not performed.

³⁸David Olds, Charles Henderson, Jr., Harriet Kitzman, John Eckenrode, Robert Cole, and Robert Tarelbaum, "The Promise of Home Visitation: Results of Two Randomized Trials," *Journal of Community Psychology* 26, no. 1 (1998): 7.

Commentary

Editor's Note: David Olds comments on the Nurse Family Partnership trials and the role of evaluation in the field of early intervention in the “Commentary” section of Chapter 19.

Note: This report is open to public comments, subject to review by the forum moderator. To leave a comment, please send an email to welfareacademy@umd.edu or fill out the comment form at http://www.welfareacademy.org/pubs/early_education/chapter21.html.