Reprinted from...

PEDIATRICS®

OFFICIAL JOURNAL OF THE AMERICAN ACADEMY OF PEDIATRICS

Effects of Nurse Home Visiting on Maternal and Child Functioning: Age-9 Follow-up of a Randomized Trial

David L. Olds, PhD, Harriet Kitzman, PhD, RN, Carole Hanks, DrPH, RN,
Robert Cole, PhD, Elizabeth Anson, MS, Kimberly Sidora-Arcoleo, PhD, MPH,
Dennis W. Luckey, PhD, Charles R. Henderson, Jr, John Holmberg, PsyD,
Robin A. Tutt, BA, Amanda J. Stevenson, BS, and Jessica Bondy, MHA

American Academy of Pediatrics

DEDICATED TO THE HEALTH OF ALL CHILDREN™

ARTICLE

Effects of Nurse Home Visiting on Maternal and Child Functioning: Age-9 Follow-up of a Randomized Trial

David L. Olds, PhD[a], Harriet Kitzman, PhD, RN[b], Carole Hanks, DrPH, RN[c], Robert Cole, PhD[b], Elizabeth Anson, MS[b],
Kimberly Sidora-Arcoleo, PhD, MPH[d], Dennis W. Luckey, PhD[e], Charles R. Henderson, Jr[f], John Holmberg, PsyD[a], Robin A. Tutt, BA[a],
Amanda J. Stevenson, BS[e], Jessica Bondy, MHA[e]

Departments of [a]Pediatrics and [e]Preventive Medicine and Biometrics, University of Colorado, Denver, Colorado; [b]University of Rochester School of Nursing, Rochester,
New York; [c]Baylor University Louise Herrington School of Nursing, Rochester, New York; [d]Arizona State University College of Nursing and Healthcare Innovation, Phoenix,
Arizona; [f]Department of Human Development, Cornell University, Ithaca, New York

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

OBJECTIVE. Our goal was to test the effect of prenatal and infancy home visits by nurses on mothers' fertility and children's functioning 7 years after the program ended at child age 2.

METHODS. We conducted a randomized, controlled trial in a public system of obstetric and pediatric care. A total of 743 primarily black women <29 weeks' gestation, with previous live births and at least 2 sociodemographic risk characteristics (unmarried, <12 years of education, unemployed), were randomly assigned to receive nurse home visits or comparison services. Primary outcomes consisted of intervals between births of first and second children and number of children born per year; mothers' stability of relationships with partners and relationships with the biological father of the child; mothers' use of welfare, food stamps, and Medicaid; mothers' use of substances; mothers' arrests and incarcerations; and children's academic achievement, school conduct, and mental disorders. Secondary outcomes were the sequelae of subsequent pregnancies, women's employment, experience of domestic violence, and children's mortality.

RESULTS. Nurse-visited women had longer intervals between births of first and second children, fewer cumulative subsequent births per year, and longer relationships with current partners. From birth through child age 9, nurse-visited women used welfare and food stamps for fewer months. Nurse-visited children born to mothers with low psychological resources, compared with control-group counterparts, had better grade-point averages and achievement test scores in math and reading in grades 1 through 3. Nurse-visited children, as a trend, were less likely to die from birth through age 9, an effect accounted for by deaths that were attributable to potentially preventable causes.

CONCLUSIONS. By child age 9, the program reduced women's rates of subsequent births, increased the intervals between the births of first and second children, increased the stability of their relationships with partners, facilitated children's academic adjustment to elementary school, and seems to have reduced childhood mortality from preventable causes.

HOME VISITING BY nurses for low-income, at-risk families has been promoted as a promising strategy for preventing child abuse and neglect, children's mental health problems,[1–3] and infant mortality.[4] Recent evidence suggests that the benefits hoped for from such programs do not hold for all types of home-visiting programs.[5] A program of home visiting by nurses known as the Nurse-Family Partnership (NFP) has produced consistent effects on several aspects of maternal and child health through the preschool period when tested in randomized trials with various racial and ethnic groups, in various living contexts, and at different historical periods.[6–8] Many of the apparent benefits of the program that have captured the attention of policy makers, however, have been derived from the adolescent follow-up of the first trial of the NFP, tested with a primarily white sample in Elmira, New York.[9,10]

The first replication trial of the NFP was conducted in Memphis, Tennessee, and focused on low-income black individuals. Results of the Memphis trial through child age 4 corroborated many of the early effects of the program on maternal life course observed in the first trial that focused on white individuals.[11,12] A recent study of program effects in Memphis through child age 6 found that the program also produced positive effects on children's cognition, mental health, and internal representations of relationships.[8] Our study was designed to examine the enduring impact of the program on mothers' life course, on children's academic and behavioral functioning in early elementary school (grades 1–3), and on mothers' reports of their children's mental health through child age 9.

For the current phase of follow-up, we hypothesized that the program would produce enduring effects consistent with those observed either earlier in this trial or in the first trial conducted in Elmira, New York, on primary maternal life-course outcomes: the intervals between births of first and second children, rates of subsequent births (operationalized at this phase of follow-up as the cumulative number of subsequent children born per year), use of welfare (Temporary Assistance for Needy Families [TANF] and food stamps), substance use, behavioral impairments as a result of substance use, arrests and number of days incarcerated, marriage, and duration of partner relations, as well as the biological father's involvement in the family. Better pregnancy planning, maternal employment, sense of mastery, and father involvement, along with reductions in substance abuse, were expected to improve family economic self-sufficiency. To understand fully the clinical and economic impact of these hypothesized changes in maternal life course, we examined the following variables as secondary outcomes: counts of subsequent miscarriages, abortions, still births, and low birth weight newborns; maternal symptoms of depression; and mothers' employment, use of Medicaid, being partnered with men who were unemployed, and experience of domestic violence.

We hypothesized that the program would produce effects on the following primary child outcomes: grade-point averages (GPAs) in reading, math, and conduct; the counts of failures in academics (reading and math) and conduct, as well as disruptive behavior, anxiety, and depressive disorders; and teachers' reports of antisocial behavior. We also examined as secondary outcomes children's special education placements and grade retentions and teachers' ratings of children's academically focused behavior and peer affiliation. Given limited statistical power, we did not hypothesize program effects on the mortality of firstborn children. We nevertheless examined program effects on this outcome given emerging treatment differences in the rates and causes of mortality.

Earlier reports on trials of this program have found consistent effects on child outcomes concentrated among children who were born to mothers who were more psychologically vulnerable.[7] We therefore predicted that program benefits for children would be concentrated on those who were born to mothers with low psychological resources (limited intellectual functioning, poor mental health, and low sense of control over their life circumstances). We examined whether program effects on maternal fertility and welfare outcomes were greater for women with initially higher psychological resources, given greater program effects on fertility for this segment of the sample in earlier phases of this trial.[9,12]

METHODS

We conducted interviews with the children's mothers by telephone ($n = 16$) and in the study offices at approximately the child's ninth birthday (mean age: 9.73 years; SD: 0.42). We reviewed children's school records in grades 1 to 3 and obtained teachers' (primarily third-grade) reports of children's classroom behavior. The details of basic study design and its implementation have been reported previously[11,12] and are summarized here.

Table 1 provides the numbers of eligible patients who were (1) invited to participate, (2) randomly assigned, and (3) evaluated at each follow-up assessment. As this table indicates, of mothers who were randomly assigned and had no fetal or child death, follow-up assessments at child age 9 were completed with 91% of the mothers, school records were abstracted for 88% of the children, teacher report forms were completed for 81% of the sample, and achievement-test scores were abstracted for 83%.

Participants

From June 1990 through August 1991, we invited to participate 1290 patients who met study inclusion criteria and were seen consecutively at the obstetric clinic of

TABLE 1　Sample Composition Over Time by Treatment

Parameter	Treatment Group				Total
	1	2	3	4	
No. allocated to treatment	166	515	230	228	1139
No. of postrandomization drops[a]	1	4	4	4	13
No. (range, SD) of completed prenatal home visits, average			7 (0–18, 4.0)	7 (0–18, 4.0)	14
No. (range, SD) of completed postnatal home visits, average				26 (0–71, 14.7)	26
No. of miscarriages	6	19	6	8	39
No. of stillbirths	0	5	3	2	10
No. of infant/child deaths (through age 9)	NA	10	NA	1	11
No. available for follow-up	NA	477	NA	213	690
No. of 9-y maternal interviews completed	NA	436	NA	191	627
No. of teacher reports obtained		387		171	558
No. of school record abstracted		416		188	604
No. of cases with achievement test scores		388		182	570

Unless otherwise specified, the difference between the number of women who were randomly assigned and assessed is attributable to missed assessments. Number of eligible patients invited to participate: 1290; number of refusals: 151; number randomly assigned: 1139. NA indicates not applicable.

[a] Eleven of the 13 women who dropped out were cases in which the mother refused additional participation after randomization. One woman (assigned to treatment 2) was dropped from the study when it was learned that she was registered and randomly assigned a second time after an earlier registration and miscarriage; a second woman (assigned to treatment 1) was not followed because of clerical error after a staff member identified her as not meeting the inclusion criteria.

the Regional Medical Center in Memphis, Tennessee. We recruited women who were of low income and unmarried because this group benefited the most in the Elmira trial. Women who were at <29 weeks of gestation were recruited when they had no previous live births, no specific chronic illnesses that are thought to contribute to fetal growth retardation or preterm delivery, and at least 2 of the following sociodemographic risk conditions: (1) unmarried, (2) <12 years of education, and (3) unemployed. Eighty-eight percent (1139) of the 1290 eligible women completed informed consent and were randomly assigned to 1 of 4 treatment conditions described in "Randomization" (2 of which were followed after delivery and form the basis of this report). Ninety-two percent of the women enrolled were black, 98% were unmarried, 64% were 18 years or younger at registration, and 85% came from households with incomes at or below the federal poverty level.

Statistical Power and Assignment Ratios

Sample sizes were established when the trial was first designed from power calculations conducted for pregnancy and infancy outcomes. For all power calculations, we set $\alpha = .05$ and $\beta = .20$ and specified 2-tailed tests. We chose to enroll fewer women in the postnatal phase of this trial than in the prenatal phase because treatment effects (in SD units) in the Elmira trial were larger for postnatal outcomes than for prenatal outcomes. These calculations also indicated that with very little loss of statistical power for normally distributed dependent variables, we could assign half as many women to the relatively expensive nurse-visitation intervention as to

the comparison condition. These calculations led to a total target sample of 750 for the postnatal phase of the study, assuming 20% attrition, and 743 were enrolled. The differences in prenatal and postnatal sample sizes and in proportion assigned to nurse and comparison conditions were accomplished by disproportionately assigning participants to 4 treatment conditions outlined in "Randomization."

Given the sample enrolled and retained at the 9-year follow-up and assuming for normally distributed variables that 10% of the variance is accounted for by other terms in the model, we estimated the smallest detectable treatment main effect size for key postnatal outcomes to be 0.24 SD and the smallest detectable effect for children who were born to low-resource mothers to be 0.33 SD. The smallest detectable program effect for dichotomous outcomes varies depending on the control group prevalence rates. We had power to detect an increase in marriage from 19.3% to 29.7% at the main effect level. Details of the design and assignment ratios are presented elsewhere.[12]

Randomization

After completion of baseline interviews, identifying information on the participants was sent to the University of Rochester, where it was entered into a computer program that randomly assigned individual women to 4 treatment conditions by using methods that are extensions of those given by Soares and Wu.[13] This procedure concealed the randomization from individuals who were directly involved with the participants in Memphis. The randomization was conducted within strata from a

model with 5 classification factors: maternal race (black versus nonblack), maternal age (<17, 17–18, and ≥19 years), gestational age at enrollment (<20 vs ≥20 weeks), employment status of head of household (employed versus unemployed), and geographic region of residence (4 regions). Women who were randomly assigned to the home-visitation groups subsequently were assigned at random to a nurse home visitor.

Treatment Conditions

Women in treatment 1 ($n = 166$) were provided free round-trip taxicab transportation for scheduled prenatal care appointments; they did not receive any postpartum services or assessments. Women in treatment 2 ($n = 515$) were provided the free transportation for scheduled prenatal care plus developmental screening and referral services for the child at 6, 12, and 24 months of age. Women in treatment 3 ($n = 230$) were provided the same services as those in treatment 1 plus intensive nurse home-visiting services during pregnancy, 1 postpartum visit in the hospital before discharge, 1 postpartum visit in the home, but no postpartum research assessments. Women in treatment 4 ($n = 228$) were provided the same services as those in treatment 3; in addition, they continued to be visited by nurses through the child's second birthday. For evaluation of postnatal outcomes, treatment 2 was contrasted with treatment 4. To reduce cost of the study, only these 2 groups were assessed after delivery of the child.

Program Plan and Implementation

The program was conducted by the Memphis/Shelby County Health Department. The nurses completed a mean of 7 home visits (range: 0–18) during pregnancy (same mean number of prenatal visits for groups 1 and 2) and 26 home visits (range: 0–71) during the first 2 years postpartum. They followed detailed visit-by-visit guidelines in their efforts to (1) improve the outcomes of pregnancy by promoting women's healthy prenatal behaviors; (2) improve the health and development of the child by promoting parents' competent care of their children; and (3) enhance parents' life-course development by encouraging parents to plan subsequent pregnancies, complete their education, and find work. The nurses helped families make use of needed health and human services and attempted to involve other family members and friends (particularly the children's fathers and grandmothers) in the pregnancy, birth, and early care of the child. Program protocols were grounded in epidemiology and theories of human ecology, human attachment, and self-efficacy.[7,14]

Masked Data Gathering

Interviews with the mothers were conducted by staff members who were masked to women's and children's treatment assignments. Some of the outcomes were based on teacher report and reviews of the children's school records. So far as we can tell, teachers were unaware of the families' treatment assignment. Although principal investigators and statisticians had access to participants' treatment assignments, all decisions about coding of interview responses and construction of variables were made explicitly without this information.

Assessments and Definitions of Variables

We assessed 2 broad domains of maternal and child outcomes: maternal life course and child behavioral, academic, and mental health functioning. Assessments for the current phase of follow-up were conducted after children had completed at least 7 months of third grade (through March) and were based on interviews with mothers, questionnaires that were completed by teachers, and reviews of the children's school records through third grade. Restrictions in funding made it impossible to test directly the children's academic achievement or to interview them regarding their psychosocial adjustment. Previous interviews with participating women were conducted at registration (before their assignment to treatments), at the 36th week of pregnancy and at the 6th, 12th, 24th, 54th, and 72nd month of the child's life. Whenever possible, we used data from these earlier phases of follow-up to estimate program impacts over time. We also reviewed Tennessee State administrative records to determine on a month-by-month basis after birth of the first child women's use of Aid to Families With Dependent Children (AFDC)/TANF (Families First in Tennessee after 1996), Medicaid (TennCare in Tennessee), and food stamps through child age 4.5 years. These administrative records were supplemented with interview data to estimate use of these services for women who moved out of Tennessee and for all participants from child age 4.5 to 9 years.

Baseline Assessments

Baseline assessments have been described in previous reports.[11,12] A variable was created to index women's psychological resources measured at registration. It was based on the mean z scores of their (1) intelligence,[15] (2) mental health,[16] (3) sense of mastery,[17] plus (4) self-efficacy (women's confidence in their ability to behave in accordance with the major behavioral objectives of the program).[14] The psychological resource variable was standardized to a mean of 100 and SD of 10 and then dichotomized at values ≤100 vs >100, creating a median split. Confirmatory factor analysis showed that the 4 components of the psychological resources variable form a single latent variable ($\alpha = .61$). We also created a single index of household poverty that was based on the averaged z scores of the variables household discretionary income, whether the head of the household was employed, and household density (number of people per

room), standardized to a mean of 100 and an SD of 10. It also formed a single latent variable ($\alpha = .54$).

Primary Maternal Outcomes

The following variables were primary maternal life-course outcomes: interval between birth of first and second children; cumulative subsequent births per year after birth of the first child through the first child's ninth birthday; duration of her relationship with current partner; being partnered with, cohabiting with, or being married to the child's biological father; her sense of mastery; duration of use of welfare (AFDC and TANF) and food stamps per year after birth of the first child; the counts of maternal arrests and days jailed; and the count of substances used (≥ 3 drinks of alcohol ≥ 3 times per month in the past year, use of marijuana, and use of cocaine since last interview at child age 6). We originally included the count of maternal behavioral impairments as a result of substance use as a primary outcome but did not include it given the infrequency of positive responses.

Secondary Maternal Life-Course Outcomes

Other outcomes were examined to help elucidate the functional and economic effects of the hypothesized changes in maternal life course, although such effects were not observed in the Elmira trial by child age 15 or at earlier phases of this trial: the counts of subsequent miscarriages, abortions, and low birth weight newborns; reported participation in the workforce; depression[18]; whether they had experienced physical violence from any of their partners since their first child was 6[19]; and the portion of time that their current partners were employed while they were together after birth of the first child.

Primary Child Outcomes

We abstracted children's GPAs in reading, math, and behavior (conduct) from their school records. To characterize failed adjustment to early elementary school, we created variables based on the counts of failed GPAs in reading and math (<1.0 for both subjects) and conduct at the end of the school year for each of the children's first 3 grades. We also abstracted the children's achievement-test scores (primarily the Tennessee Comprehensive Assessment Program Achievement Test),[20] expressed in percentiles derived from national standards.

We assessed teacher report of antisocial behavior (described in "Secondary Child Outcomes") and maternal report of child disruptive behavior disorders (eg, conduct disorder, oppositional defiant disorder, attention-deficit/hyperactivity disorder) and depressive and anxiety disorders (major depression, dysthymia, generalized anxiety disorder, separation anxiety disorder, and social phobia) for the past year using the Computerized Diagnostic Interview Schedule for Children.[21] Given low rates of reported disorders, we used subthreshold diagnoses produced by the Computerized Diagnostic Interview Schedule for Children in addition to diagnoses that met conventional diagnostic criteria. Even with these broader definitions of disorder, the rates were too infrequently occurring for individual disorders to be used as dependent variables, so we created counts of the disorders within 2 broad categories: (1) a count of depressive and anxiety disorders reported in the past year (major depression, dysthymia, generalized anxiety, separation anxiety, and social phobia) with actual values ranging between 0 and 5 and (2) a count of disruptive behavior disorders reported in the past year (either oppositional defiant disorder or conduct disorder, and attention-deficit disorder of any type), with actual values ranging between 0 and 2. Scores that exceeded 1 constitute comorbidity within these broad areas of externalizing and internalizing disorders. Overall, these counts of disorders predicted children's academic grades and teachers' reports of child behavior better than did individual diagnostic categories, a finding consistent with evidence that comorbidity reflects greater severity.[22,23] We applied the strict "D" impairment criteria to these diagnoses to increase the likelihood that the disorders would have functional significance.

Secondary Child Outcomes

We counted the number of times children were retained in grades 1 to 3. We also coded whether they were placed in special education and collected teachers' assessments of children's behavior in the classroom using items from the Social Competence Scale[24] and Social Health Profile[25] from the Fast Track trial and the Teacher Observation of Child Adjustment Revised.[26] The items from these 3 instruments were subjected to principle axis analysis, which produced 3 scales with high levels of internal consistency: (1) antisocial behavior, a primary outcome ($\alpha = .95$), and (2) academically focused behavior ($\alpha = .95$) and peer affiliation ($\alpha = .80$). The scales produced by these analyses were standardized to means of 100 and SDs of 10.

Finally, we systematically assessed children's death by sending every case in which the child was born alive and on which we did not complete a maternal assessment at age 9 to the National Death Index (NDI), administered by the Centers for Disease Control and Prevention. We sent identifying information on the children's names, date of birth, location of birth, race, and (where they were available) social security numbers and matched each case with the NDI data. We coded the age of the child at death (in days) and the *International Classification of Diseases, Ninth Revision* cause of death from NDI.

Statistical Models and Methods of Analysis

Data analyses were conducted and reported on all women who were randomly assigned insofar as outcome

data were available. The primary statistical model consisted of a 2-level treatment factor (treatment 2 vs 4), a 2-level factor reflecting mothers' psychological resources (above versus below the sample median), the interaction between these classification factors, and 2 covariates (household poverty and maternal childrearing attitudes associated with child maltreatment) measured at intake to adjust for treatment nonequivalence ($P < .10$) among participants assessed at the 9-year follow-up. Given its powerful influence on children's functioning, gender of the child was included in the analyses of child outcomes, after we examined whether child gender interacted with treatment. Estimates and tests were adjusted for all covariates. Homogeneity of regressions was tested for all covariates.[27]

For quantitative outcomes on which we had multiple assessments for each mother (eg, number of months of using welfare per year after birth of the first child) or each child, we analyzed the data using mixed models that included, in addition to the variables from the core model (treatment, psychological resources, household poverty, maternal childrearing attitudes plus child gender for child outcomes), children (or mothers) as levels of a random factor, a fixed repeated measures classification factor for time of assessment, and all interactions of time with the other fixed classification factors. School performance outcomes (GPAs and achievement-test scores) were available for 2 subject areas (math and reading) for each of 3 grades.[1–3] For these outcomes, grade level was the repeated measure over time, and the model included a second fixed repeated measures factor for subject area. (Conduct grades were analyzed in an analogous model excluding the 2-level repeated measure for subject area.) The schools attended by the children at grade 3 were included in the model as levels of an additional random classification factor, which takes into account variance associated with schools. Using repeated measures in the analyses minimizes problems with sample attrition by allowing use of all assessments.

An error structure was assumed with different variances at each time (and subject area) and general covariances between time (and subject areas) for a given child. These were assumed to be the same for all children, and covariances between children were assumed to be negligible. Schools were assumed to have common variances and 0 covariances.

The key tests focused on the treatment effect averaged over all other fixed classification variables, including those within subjects, and the same treatment effect restricted to low psychological resources. The figures show the least square means over time, which is also averaged over other fixed classification effects. In all models, we examined the interactions of treatment with time and the 3-way interactions of treatment, time, and psychological resources; none was significant. For maternal repeated outcomes, we report results averaged over the entire period for which we have data as well as the interval between 6 and 9 years of the first child's life, because the 6- to 9-year period reflects the time covered by this phase of follow-up. We show results for the entire time period for which we have data, because this is the first time in this trial that we have examined the full longitudinal effects of the program on outcomes. Where necessary, we weighted contrasts to account for varying time intervals.

All dependent variables were examined to determine their distributional characteristics. Quantitative dependent variables were analyzed in the general linear model, and dichotomous outcomes, such as rates of cohabitation, were analyzed in the logistic-linear model. We created a variable that characterized mother's degree of involvement with the father of the child according to 3 dichotomous conditions (married to, living with, or partnered with) and analyzed it using ordered logistic regression.

We analyzed low-frequency count outcomes (eg, counts of subsequent abortions, low birth weight newborns, depressive and anxiety disorders) in generalized linear models with negative binomial error and log link assumptions. To reduce problems with overspecification of sparse outcomes, we specified models that were pared-down versions of the primary model by including only terms that were significant or trends. For all maternal low-frequency count outcomes except the rates of subsequent low birth weight newborns, we included only the treatment classification factor, with no covariates; we analyzed the count of subsequent low birth weight newborns in a model that included treatment, psychological resources, the treatment × psychological resource interaction, and the household poverty covariate. For all child low-frequency count outcomes except mortality, the model consisted of treatment, psychological resources, and child gender (no interactions); the child mortality dichotomous outcome was tested in a simple treatment model with no psychological resource factor or adjustments for covariates.

To facilitate ascertainment of whether program effects on child outcomes were indeed greater for children who were born to mothers with low psychological resources, Table 2 presents child quantitative and dichotomous outcomes for both treatment main effects and effects for the group defined by mothers' having low psychological resources. Low-frequency count outcomes typically were too sparse to allow stable estimates of treatment effect within levels of psychological resources.

The P values in the tables are from likelihood ratio tests; Wald tests can be problematic in some situations.[28] The confidence intervals (CIs) shown are Wald CIs as a matter of convenience because likelihood ratio tests do not lend themselves as straightforwardly to the construction of CIs.

TABLE 2 Adjusted Estimate of Program Effects on Children's Academic Performance, Behavior, and Mental Health Through Child Age 9

Outcomes	Treatment Group		Treatment Comparison	
			Comparison vs Nurse	
	Comparison, Least-Square Mean (SE)	Nurse-Visited, Least-Square Mean (SE)	P	ES (CI)[a]
Quantitative sample[b]				
GPA (reading and math), grades 1–3[c,d]				
Whole	2.59 (0.04)	2.69 (0.06)	.200	0.09 (−0.05 to 0.22)
Low-resource	2.44 (0.06)	2.68 (0.09)	.016	0.22 (0.04 to 0.41)
Achievement tests (reading and math), grades 1–3[c,d]				
Whole	41.63 (1.34)	44.61 (1.86)	.174	0.11 (−0.05 to 0.26)
Low-resource	35.72 (1.78)	44.89 (2.53)	.002	0.33 (0.12 to 0.54)
Conduct grades, grades 1–3[c,d]				
Whole	2.68 (0.04)	2.71 (0.07)	.673	0.03 (−0.11 to 0.17)
Low-resource	2.65 (0.06)	2.68 (0.09)	.749	0.03 (−0.16 to 0.22)
Antisocial behavior, grade 3[a,c]				
Whole	100.08 (0.51)	99.77 (0.77)	.742	−0.03 (−0.21 to 0.15)
Low-resource	100.17 (0.71)	100.18 (1.06)	.994	0.00 (−0.25 to 0.25)
Academically focused behavior, grade 3[e]				
Whole	100.08 (0.51)	100.10 (0.77)	.981	0.00 (−0.18 to 0.18)
Low-resource	98.70 (0.70)	99.59 (1.05)	.471	0.09 (−0.15 to 0.33)
Peer affiliation, grade 3[e]				
Whole	99.92 (0.51)	100.35 (0.77)	.643	0.04 (−0.14 to 0.23)
Low-resource	99.37 0.70	99.56 1.06	.882	0.02 (−0.23 to 0.26)

Low-frequency count[f]	Incidence	Incidence	P	IR (CI)[a]
Count of conduct failures, grades 1–3, whole[c,g]	0.10	0.06	.091	0.56 (−1.26 to 0.11)
Count of depressive and anxiety disorders, whole[c]	0.19	0.12	.116	0.64 (−0.99 to 0.11)
Count of disruptive behavior disorders (with impairment), whole[c]	0.31	0.36	.417	1.15 (−0.19 to 0.47)

Dichotomous[a]	%	%	P	OR (CI)[a]
Any academic failures, grades 1–3, whole[c,h]	5.1	7.0	.372	1.40 (0.67 to 2.92)
Ever retained, grades 1–3, whole	12.4	16.0	.247	1.35 (0.82 to 2.21)
Ever placed in special education, grades 1–3, whole	2.3	2.2	.972	0.98 (0.36 to 2.65)

OR indicates odds ratio.

[a] CIs are estimated from Wald tests (the standard SAS output for CIs), whereas the P values are based on likelihood ratio tests.

[b] Model for quantitative and dichotomous outcomes included treatment condition, maternal psychological resources, and child gender as classification factors and household poverty and childrearing beliefs associated with maltreatment as covariates.

[c] Primary outcome.

[d] Outcome examined with repeated measures analysis.

[e] Scale derived from principle components analysis of teacher report of child behavior.

[f] Model for count of these outcomes included treatment condition, the psychological resources classification factor, and gender of child.

[g] Count of average GAP in conduct <1.0.

[h] Whether child failed both reading and math (GPA <1.0) in any grade.

RESULTS

Baseline Equivalence of Treatment Groups

As shown in Table 3, the treatment groups were similar on background characteristics for participants on whom 9-year follow-up assessments were conducted, with the following exceptions: at intake, nurse-visited women (treatment 4) lived in households with greater levels of poverty and worse scores on childrearing attitudes associated with child maltreatment than did women in the comparison group. These differences suggest that the nurse-visited group at child age 9 had a higher portion of families who had been at risk at enrollment, although, as shown in Table 1, the proportion of families on whom assessments were conducted was high and nearly equivalent across treatment conditions.

Maternal Life Course

Table 4 shows that during the 9-year period after birth of the first child, among women with at least 1 subsequent child, nurse-visited women had longer intervals between the births of first and second children (40.73 vs 34.09 months; effect size [ES] = 0.29; P = .002) and had fewer cumulative subsequent births per year than did their control-group counterparts (0.81 vs 0.93; ES = −0.14; P = .045). Figure 1 shows the cumulative number of subsequent births per year by treatment assignment and women's psychological resources; it emphasizes that the treatment main effect on number of cumulative subsequent births was limited to women with initially high psychological resources (0.69 vs 0.91; ES = −0.26; P = .010; data not shown in Table 4),

TABLE 3 Background Characteristics of Participants on Whom 9-Year Assessments Were Completed

Background Variable	Treatment Group	
	Comparison ($n = 436$)	Nurse-Visited ($n = 191$)
Married, %		
Whole	1.4	0.5
Low-resource	0.4	1.0
Maternal race, nonblack, %		
Whole	5.7	7.9
Low-resource	4.8	6.8
Head of household employed, %		
Whole	56.8	50.0
Low-resource	52.4	50.0
Drank alcohol last 14 d, %		
Whole	4.4	4.2
Low-resource	5.7	5.8
Smoked cigarettes last 3 d, %		
Whole	8.3	9.4
Low-resource	8.3	10.7
Used marijuana last 14 d, %		
Whole	1.6	1.0
Low-resource	1.7	1.9
Any drug use (screen), %		
Whole	4.2	3.3
Low-resource	7.4	6.0
Any sexually transmitted disease, prerandomization, %		
Whole	33.6	37.2
Low-resource	32.8	40.8
Maternal age, mean (SD), y		
Whole	18.03 (3.19)	18.02 (3.30)
Low-resource	18.10 (3.28)	18.13 (3.86)
Gestational age at randomization, mean (SD), wk		
Whole	16.53 (5.74)	16.56 (5.58)
Low-resource	16.35 (5.83)	16.80 (5.50)
Psychological resources index, mean (SD)[a,b]		
Whole	99.84 (9.95)	99.62 (10.81)
Low-resource	92.27 (5.77)	91.73 (6.77)
Highest grade completed, mother, mean (SD)		
Whole	10.24 (1.87)	10.06 (2.00)
Low-resource	9.91 (1.91)	9.54 (2.02)
Household poverty index, mean (SD)[b,c]		
Whole	99.59 (10.05)	102.02 (9.96)
Low-resource	101.91 (10.08)	103.67 (9.59)
Neighborhood adversity index, mean (SD)[b,d]		
Whole	3.22 (2.02)	3.35 (2.33)
Low-resource	3.36 (1.90)	3.23 (2.34)
Conflict with mother, mean (SD)[b,e]		
Whole	99.80 (10.35)	100.46 (9.17)
Low-resource	101.79 (12.57)	101.31 (10.15)
Conflict with partner, mean (SD)[b,e]		
Whole	99.73 (10.13)	100.62 (9.69)
Low-resource	100.93 (11.33)	102.36 (11.38)
Attitudes toward child rearing predictive of child abuse, mean (SD)[b,e]		
Whole	99.55 (9.52)	101.04 (10.97)
Low-resource	102.47 (9.05)	104.82 (9.42)

[a] Average z scores of women's sense of mastery/self-efficacy, mental health, and intellectual functioning.

[b] Standardized to sample mean = 100, SD = 10.

[c] Average z scores of household discretionary income, housing density, and whether head of household was employed.

[d] Average of variables calculated in SD units above the national means of components that comprise a standard neighborhood disorganization scale (eg, percentage of block group below the federal poverty level; percentage of families headed by single women; percentage of families receiving public assistance).[37]

[e] Locally developed scale that assesses the degree to which individual provides emotional and material support to mother.

TABLE 4 Adjusted Estimate of Program Effects on Maternal Life-Course Outcomes 9 Years After Delivery of First Child

Outcomes	Treatment Group		Treatment Comparisons	
	Comparison, Least-Square Mean (SE)	Nurse-Visited, Least-Square Mean (SE)	Comparison vs Nurse	
			P	ES (CI)[a]
Quantitative[b]				
No. of months between birth of first and second child[c,d]	34.09 (1.19)	40.73 (1.81)	.002	0.29 (0.11 to 0.48)
Cumulative subsequent live births per year (0–9 y)[c,d]	0.93 (0.03)	0.81 (0.05)	.045	−0.14 (−0.28 to −0.00)
Cumulative subsequent live births per year (6–9 y)[c,d]	1.53 (0.05)	1.40 (0.08)	.165	−0.16 (−0.39 to 0.07)
No. of months with current partner (at 6 and 9 y)[c,d]	44.48 (1.48)	51.89 (2.25)	.006	0.23 (0.07 to 0.39)
No. of months with current partner (at 9 y)[c,d]	52.40 (2.09)	61.59 (3.18)	.016	0.28 (0.05 to 0.51)
No. of months on AFDC/TANF per year (0–9 y)[c,d]	5.92 (0.15)	5.21 (0.22)	.008	−0.14 (−0.25 to −0.04)
No. of months on TANF (6–9 y)[c,d]	4.01 (0.22)	3.39 (0.33)	.117	−0.12 (−0.28 to 0.03)
No. of months on food stamps per year (0–9 y)[c,d]	7.80 (0.14)	6.98 (0.21)	.001	−0.17 (−0.28 to −0.07)
No. of months on food stamps per year (6–9 y)[c,d]	5.92 (0.24)	4.89 (0.36)	.017	−0.21 (−0.39 to −0.04)
Maternal mastery (6 mo–9 y)[c,d]	99.50 (0.30)	101.03 (0.45)	.005	0.15 (0.05 to 0.26)
Maternal mastery (9 y)[c,d]	99.75 (0.44)	100.79 (0.67)	.196	0.10 (−0.05 to 0.26)
No. of months on Medicaid per year (0–9 y)[d]	10.07 (0.13)	9.71 (0.19)	.119	−0.09 (−0.20 to 0.02)
No. of months on Medicaid per year (6–9 y)[d]	8.74 (0.23)	8.79 (0.34)	.889	0.01 (−0.19 to 0.22)
No. of months employed per year (2–9 y)[d]	3.86 (0.12)	3.84 (0.19)	.930	−0.01 (−0.13 to 0.12)
No. of months employed per year (6–9 y)[d]	7.39 (0.19)	6.86 (0.30)	.132	−0.15 (−0.34 to 0.05)
No. of months with employed partner (at 6 and 9 y)[d]	38.43 (1.36)	46.04 (2.05)	.002	0.25 (0.09 to 0.41)
No. of months with employed partner (at 9 y)[d]	46.01 (1.94)	54.95 (2.95)	.012	0.30 (0.07 to 0.53)
Maternal depression, Center for Epidemiologic Studies Depression Scale (9 y)	1.72 (0.03)	1.71 (0.04)	.874	−0.01 (−0.18 to 0.15)
Low-frequency count[e]	Incidence	Incidence	P	IR (CI)[a]
No. of maternal arrests (6–9 y)[c]	0.30	0.41	.163	1.35 (−0.12 to 0.71)
No. of substances used (moderate/heavy alcohol, marijuana, cocaine) since last interview (9 y)[c]	0.17	0.10	.075	0.62 (−1.03 to 0.06)
No. of subsequent miscarriages (0–9 y)	0.24	0.24	.9998	1.00 (−0.40 to 0.40)
No. of subsequent abortions (0–9 y)	0.20	0.14	.152	0.67 (−0.94 to 0.15)
No. of subsequent low birth weight newborns (0–9 y)	0.27	0.18	.073	0.66 (−0.89 to 0.05)
Dichotomous[b]	%	%	P	OR (CI)[a]
Relationship with Father of Child Scale[c,d]			.091	1.58 (0.93 to 2.67)
Married to father of child	4.9	7.4		
Lives with father of child (not married)	1.6	2.4		
Partnered with father of child (not married or cohabiting)	2.8	4.0		
Married (9 y)[c]	19.3	22.5	.226	1.30 (0.85 to 1.99)
Mother jailed (6–9 y)[c]	2.5	3.7	.456	1.46 (0.55 to 3.85)
Partnered (9 y)	35.1	38.2	.329	1.20 (0.84 to 1.71)
Subsequent still birth (0–9 y)	2.1	1.0	.359	0.51 (0.11 to 2.38)
Any domestic violence (6–9 y)	23.7	20.6	.373	0.81 (0.52 to 1.29)

[a] CIs are estimated from Wald tests (the standard SAS output for CIs), whereas the P values are based on likelihood ratio tests.

[b] Model for quantitative and dichotomous outcomes included treatment condition and maternal psychological resources as classification factors and household poverty and childrearing beliefs associated with maltreatment as covariates.

[c] Primary outcomes.

[d] Outcome examined with repeated measures analysis.

[e] Model for low-frequency count outcomes included treatment condition only, except for the count of subsequent low birth weight newborns, which included the treatment and maternal psychological resources classification factors and the treatment × psychological resource interaction, plus household poverty as a covariate.

averaging across the entire period after birth of the first child. Consistent with the reduction in cumulative subsequent births, nurse-visited women, as a trend, had fewer subsequent low birth weight newborns (0.18 vs 0.27; incidence ratio [IR]: 0.66; P = .073).

Averaging across the 6- and 9-year follow-up periods (the only times that we asked about duration of partner relations), nurse-visited mothers had longer relationships with their current partners (51.89 vs 44.48 months; ES = 0.23; P = .006), an effect illustrated in Fig 2. The program effect was particularly pronounced at child age 9 (61.59 vs 52.40 months; ES = 0.28; P = .016). In correspondence with their longer partnered relationships, nurse-visited women were associated with employed partners to a greater degree than were women in the control group (46.04 vs 38.43 months; ES = 0.25; P = .002). As a trend, at the first child's ninth year, nurse-visited women were more likely to be married to, living with, or partnered with their first child's biological father (odds ratio: 1.58; P = .091).

From birth through child age 9, nurse-visited women used welfare (AFDC/TANF) and food stamps
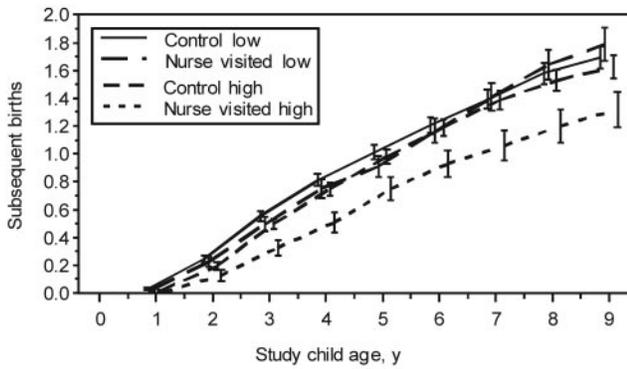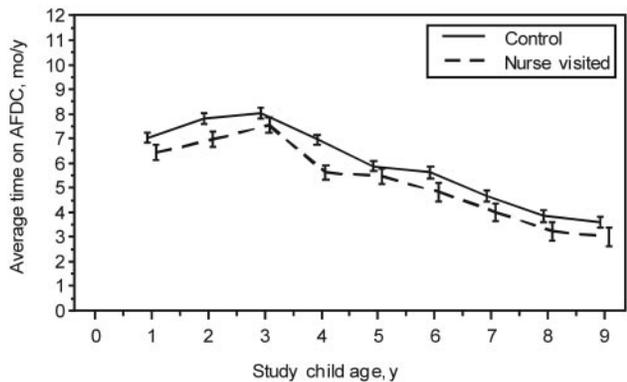
FIGURE 1
Program impact on number of subsequent children who were born to mothers after birth of first child by maternal psychological resources.
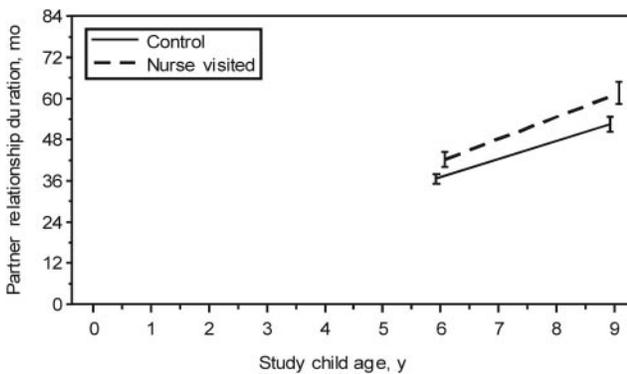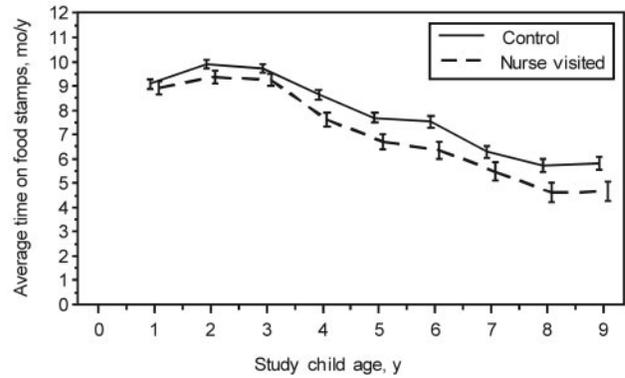

FIGURE 2
Program impact on duration of partner relations at 6 and 9 years after birth of first child.

for fewer months per year (5.21 vs 5.92 [ES = −0.14; P = .008] and 6.98 vs 7.80 months per year [ES = −0.17; P = .001], respectively). Figures 3 and 4 present the pattern of program effects on welfare and food stamp use during the 9-year period after birth of the first child. For the 6- to 9-year interval, the program effect on food stamps was significant (4.89 vs 5.92 months per year; ES = −0.21; P = .017), but the

FIGURE 3
Program impact on use of AFDC/TANF (mean number of months per year) after birth of first child.

FIGURE 4
Program impact on use of food stamps (mean number of months per year) after birth of first child.

effect on AFDC/TANF was not (3.39 vs 4.01 months per year; ES = −0.12; P = .117).

When examined during the entire 9-year period, nurse-visited women expressed greater mastery over the challenges in their lives (101.03 vs 99.50; ES = 0.15; P = .005). Figure 5 shows that this effect was concentrated during the period while the program was operating (through child age 2). By age 9, the treatment–control difference was no longer significant.

Nurse-visited mothers, as a trend, used fewer substances (the count of moderate-to-heavy alcohol use, marijuana, and cocaine: 0.10 vs 0.17; IR: 0.62, P = .075).

There were no statistically significant program effects on women's subsequent miscarriages, abortions, or stillbirths; arrests or being jailed; use of Medicaid; depression; employment; or marriage or being in a partnered relationship.

Child Outcomes
As shown in Table 2, nurse-visited children who were born to mothers with low psychological resources, compared with their control-group counterparts, had better GPAs averaged across reading and math (2.68 vs 2.44;
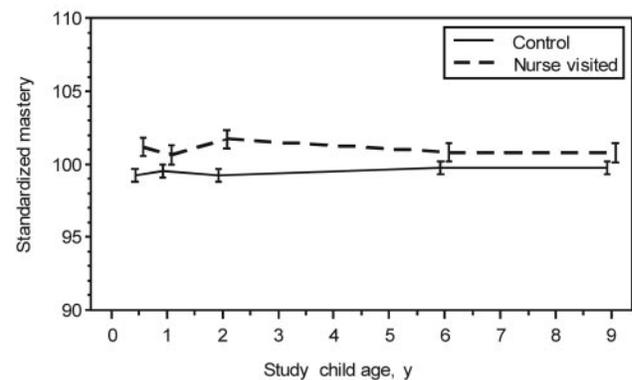
FIGURE 5
Program impact on maternal mastery over time.

ES = 0.22; P = .016) and had better math and reading achievement-test scores in grades 1 to 3 (44.89 vs 35.72; ES = 0.33; P = .002). Program effects sizes on GPAs and achievement-test scores were not as large for reading (ES = 0.17 and 0.23, respectively) as they were for math (ES = 0.28 and 0.43, respectively), but they were sufficiently large for reading to produce effects that were significant averaging across both subject matters. Program effects on children's GPAs during the first 3 years of elementary school are shown in Fig 6, which emphasizes that the group that performed most poorly was control-group children who were born to low-resource mothers; nurse-visited children who were born to low-resource mothers had GPAs that were equivalent to those who were born to high-resource mothers. Overall, nurse-visited children, as a trend, had fewer failures in conduct in the first 3 years of elementary school (0.06 vs 0.10; IR: 0.56; P = .091).

There were no statistically significant program effects on placements in special education or mothers' reports of their children's disruptive behavior disorders or third-grade teachers' reports of children's behavioral or academic adaptation to the classroom.

Infant and Childhood Death

Table 5 shows the pattern of infant and childhood death among firstborn children from birth through child age 9. Control-group children were 4.46 times more likely to die in this age range than were nurse-visited children (20.08 per 1000 vs 4.50 per 1000; odds ratio: 0.22; CI: 0.03–1.74; P = .080). Three of the 10 deaths in the control group were attributed to complications of preterm delivery, 3 to sudden infant death syndrome, and 3 to injury. The 1 nurse-visited death was attributed to a chromosomal abnormality.

DISCUSSION

Through the first child's ninth birthday, the program continued to increase the interval between the births of
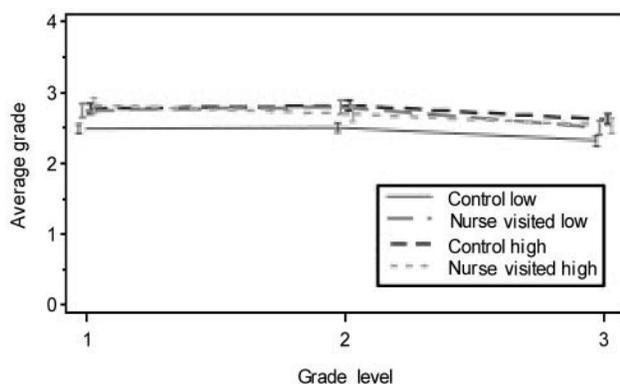
FIGURE 6
Program impact on children's GPAs in reading and math by grade level and maternal psychological resources.

TABLE 5 Causes of Infant and Child Deaths (ICD-9) Among Firstborn Children Through Age 9

Cause of Death (ICD-9 Code)	Age at Death, d
Comparison group (n = 498)	
Extreme prematurity (7650)	3
SIDS (7980)	20
SIDS (7980)	35
Ill-defined intestinal infections (90)	36
SIDS (7980)	49
Multiple congenital anomalies (7597)	152
Chronic respiratory disease arising in perinatal period (7707)	549
Homicide assault by firearm (9654)	1569
Motor vehicle accident (8129)	2100
Accident caused by firearm (9229)	2114
Nurse-visited group (n = 222)	
Chromosomal abnormalities (7589)	24

ICD-9 indicates *International Classification of Diseases, Ninth Revision*; SIDS, sudden infant death syndrome.

first and second children, reduced the cumulative number of subsequent live births per year, increased the stability of mothers' relationships with their partners, and reduced women's use of welfare and food stamps. The impact of the program on cumulative subsequent live births was limited to mothers with initially higher levels of psychological resources, and its impact on use of welfare was concentrated in the 0- to 6-year interval after birth of the first child. Through the first 3 years of elementary school, the program improved the academic achievement of children who were born to mothers with low psychological resources, and as a trend, it reduced the rate of infant and childhood mortality among firstborn children during the 9-year period after birth of the first child.

Although the statistical significance of the treatment difference in infant and childhood mortality is only marginal, the causes of death are noteworthy. The 1 death in the nurse-visited group was attributed to a chromosomal anomaly. Nine of the 10 deaths in the control group were either associated with preterm delivery or attributable to sudden infant death syndrome or injury. This raises the possibility that when focused on highly impoverished populations such as that sampled in this trial, the program may be able to prevent a range of adverse child outcomes, including death.

The impact of the program on duration of partner relationships and, as a trend, involvement of the child's biological father in the family is consistent with corresponding effects observed on stability of partner relationships in the Elmira program at child age 3[29] and marriage at child age 15.[30] This consistency of effects lends validity to each of these findings.

The impact of the program on fertility-related outcomes among mothers with higher psychological resources is consistent with findings reported previously,[8,11,12] which we have attributed to mothers' developing

the wherewithal to envision and secure employment[11] and manage simultaneously the demands of being an employee and providing competent care for their children.[8] One crucial factor contributing to economic self-sufficiency is pregnancy planning. The lower resource mothers who were visited by nurses, we hypothesize, had fewer personal resources to enable them to manage both roles well and therefore chose to focus their limited resources on the care of their children rather than attempting to make it in the world of work.

Without help, low-resource mothers are at greater risk for having difficulty caring competently for their children, who in turn are at risk for a host of problems. We believe that nurse-visited low-resource mothers chose to focus their resources on the care and protection of their children, and this explains why they were particularly successful, compared with their control-group counterparts, in managing the care of their firstborn children, as reflected in their children's having fewer injuries through age 2 and better cognition, arithmetic achievement, adjustment at age 6, and academic achievement in grades 1 to 3.[7,8,11,12]

It is possible that the reduction in use of AFDC/TANF and food stamps observed during the 9-year period after birth of the first child for the entire sample may be explained at least in part by the nurse-visited women's increased involvement with the first child's biological father and the stability of partnered relationships, given that their partners were frequently employed and most likely brought additional financial resources to the household.

In 1996, the US welfare reform act went into effect,[31] limiting women's lifetime use of public assistance (TANF). Although this may partially explain the gradually diminished impact of the program on use of welfare over time, Tennessee has had a waiver that exempts it from invoking all of the TANF restrictions on use of welfare.[32] Moreover, it is important to note that Medicaid in Tennessee covered pregnant women up to 185% of poverty during the most recent phase of follow-up and paid for 37% of all births in 2000.[33] These high rates of coverage probably account for the failure of the program to reduce use of Medicaid overall in this highly disadvantaged population.

At the 6-year follow-up of this sample, we reported that nurse-visited mothers had placed their children in some form of structured child care or preschool before kindergarten.[8] To determine the degree to which the program effects reported here on academic achievement were accounted for by this increased use of child care and preschool, we repeated the analysis of academic achievement outcomes controlling for enrollment in preschool programs. The impact of the program on the achievement scores and GPAs of children who were born to low-resource mothers was virtually unchanged by this statistical control. Therefore, the program impact on children's achievement observed in this trial is independent of the nurse-visited children's higher enrollment in preschool programs.

These findings are encouraging but must be interpreted in light of their limitations. The first is that we did not have funds to conduct direct assessments of the children at this phase of follow-up and therefore had to rely on maternal report and children's school records for information on child functioning. Assessing child welfare records might have increased our insight into program impact, but these records are particularly unreliable in Tennessee before a legal settlement that was reached in 2001 to improve the state's child welfare infrastructure, including the establishment and maintenance of a valid computerized information system.[34] Moreover, official records pick up only a very small fraction of actual maltreatment.[35]

The second limitation is that some of the outcomes are not independent of one another. We have reported interrelated outcomes to provide the reader with a full understanding of the range of program effects.

Finally, the program impact on childhood mortality does not reach conventional levels of statistical significance. The program–control difference in mortality, however, is attributable to potentially preventable causes, and the finding is consistent with earlier program effects on injuries and qualities of parental care,[12] as well as a recent evaluation of a statewide replication of the NFP in Oklahoma, which found significantly fewer neonatal deaths and preterm deliveries on the part of intervention infants who were born to unmarried mothers without obstetric complications compared with propensity-matched control subjects.[36]

REFERENCES

1. US Department of Health and Human Services. *Mental Health: A Report of the Surgeon General*. Rockville, MD: US Dept of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services, National Institutes of Health; 1999

2. Chaffin M. Is it time to rethink Healthy Start/Healthy Families? *Child Abuse Negl.* 2004;28:589–595

3. New Freedom Commission on Mental Health. *Achieving the Promise: Transforming Mental Health Care in America. Final Report.* Rockville, MD: Department of Health and Human Services; 2003. DHHS publication SMA-03-3832

4. National Commission to Prevent Infant Mortality. *Home Visiting: Opening Doors for America's Pregnant Women and Children.* Washington, DC: National Commission to Prevent Infant Mortality; 1989

5. Gomby DS. *Home Visitation in* 2005: Outcomes for Children and Parents. Washington, DC: Committee for Economic Development; 2005. Invest in Kids Working Paper No. 7

6. Olds DL, Robinson J, O'Brien R, et al. Home visiting by paraprofessionals and by nurses: a randomized, controlled trial. *Pediatrics.* 2002;110:486–496

7. Olds DL. Prenatal and infancy home visiting by nurses: from randomized trials to community replication. *Prev Sci.* 2002;3:153–172

8. Olds DL, Kitzman H, Cole R, et al. Effects of nurse home-visiting on maternal life course and child development: age 6 follow-up results of a randomized trial. *Pediatrics.* 2004;114:1550–1559

9. Olds DL, Henderson CR Jr, Cole R, et al. Long-term effects of nurse home visitation on children's criminal and antisocial behavior: 15-year follow-up of a randomized controlled trial. *JAMA.* 1998;280:1238–1244

10. Olds DL, Eckenrode J, Henderson CR Jr, et al. Long-term effects of home visitation on maternal life course and child abuse and neglect: fifteen-year follow-up of a randomized trial. *JAMA.* 1997;278:637–643

11. Kitzman H, Olds DL, Sidora K, et al. Enduring effects of nurse home visitation on maternal life course: a 3-year follow-up of a randomized trial. *JAMA.* 2000;283:1983–1989

12. 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.* 1997;278:644–652

13. Soares JF, Wu CF. Some restricted randomization rules in sequential designs. *Common Stat Theory Methods.* 1983;12:2017–2034

14. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. *Psychol Rev.* 1977;84:191–215

15. Shipley W. A self-administered scale for measuring intellectual impairment and deterioration. *J Psychol.* 1940;9:371–377

16. Ware JE, Veit CT, Donald CA. *Refinements in the Measurement of Mental Health for Adults in the Health Insurance Study*. Santa Monica, CA: Rand Corp; 1985

17. Pearlin LI, Schooler C. The structure of coping. *J Health Soc Behav.* 1978;19:2–21

18. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. *Appl Psychol Meas.* 1977;1:385–401

19. Straus MA. Measuring intrafamily conflict and violence: the Conflict Tactics (CT) Scales. *J Marriage Fam.* 1979;41:75–88

20. *State of Tennessee Department of Education Educators' Guide to Understanding TCAP Achievement Test Results*. Monterey, CA: CTB/McGraw-Hill LLC; 2004

21. Shaffer D, Fisher P, Dulcan MK, et al. The NIMH Diagnostic Interview Schedule for Children Version 2.3 (DISC-2.3): description, acceptability, prevalence rates, and performance in the MECA Study. Methods for the Epidemiology of Child and Adolescent Mental Disorders Study. *J Am Acad Child Adolesc Psychiatry.* 1996;35:865–877

22. Kendall PC, Kortlander E, Chansky TE, Brady EU. Comorbidity of anxiety and depression in youth: treatment implications. *J Consult Clin Psychol.* 1992;60:869–880

23. Kessler RC, McGonagle KA, Zhao S, et al. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. *Arch Gen Psychiatry.* 1994;51:8–19

24. Social Competence Scale–Teacher Version. Conduct Problems Prevention Research Group. Available at: www.pubpol.duke.edu/centers/child/fasttrack/techrept/s/sct. Accessed January 13, 2007

25. Social Health Profile. Conduct Problems Prevention Research Group. Available at: www.fasttrackproject.org/techrept/s/shs. Accessed January 13, 2007

26. Werthamer-Larsson L, Kellam S, Wheeler L. Effect of first-grade classroom environment on shy behavior, aggressive behavior, and concentration problems. *Am J Community Psychol.* 1991;19:585–602

27. Henderson CR Jr. Analysis of covariance in the mixed model: higher-level, nonhomogeneous, and random regressions. *Biometrics.* 1982;38:623–640

28. Fears R, Benichou J, Gail M. A reminder of the fallibility of the Wald statistic. *Am Stat.* 1996;50:226–227

29. Olds DL, Lombardi J, Birmingham MT, Henderson CR Jr. *Prenatal/Early Infancy Project: A Follow-up Evaluation at the Third and Fourth Years of Life—Final Report to the Robert Wood Johnson Foundation (Grant 3729)*. Rochester, NY: University of Rochester; 1986

30. Pettitt LM, Olds DL. Father involvement in a program of prenatal and infancy home visiting by nurses. Presented at: Conference IV of the Administration for Children & Families and Office of Planning, Research, & Evaluation Welfare Reform; May 2001; Washington, DC

31. Personal Responsibility, Work Opportunity, and Medicaid Restructuring Act of 1996, H.R. Res. 3734, PL 104–193 (1996)

32. Families First/TANF. Tennessee Department of Human Services. Available at: www.state.tn.us/humanserv/adfam/afs_tanf.htm. Accessed January 9, 2007

33. NGA Center for Best Practices. *MCH Update 2002: State Health Coverage for Low-Income Pregnant Women, Children, and Parents*. Washington, DC: National Governors Association; 2003

34. Civil Action No. 03-00-0445 Brian A. v. Sundquist. Judge Campbell: United States District Court for the Middle District of Tennessee, Nashville Division

35. Theodore AD, Chang JJ, Runyan DK, Hunter WM, Bangdiwala SI, Agans R. Epidemiologic features of the physical and sexual maltreatment of children in the Carolinas. *Pediatrics*. 2005; 115(3). Available at: www.pediatrics.org/cgi/content/full/115/3/e331

36. 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? *Paediatr Perinat Epidemiol.* 2005;19:194–205

37. Lauritsen JL. *How Families and Communities Influence Youth Victimization*. Washington, DC: Office of Justice Programs, US Department of Justice; 2003