Nine-Year Follow-up of a Home-Visitation Program: A Randomized Trial

**WHAT’S KNOWN ON THIS SUBJECT:** A number of studies have shown that home-visiting interventions can improve outcomes for children being raised in families that face multiple adversities. It is less clear how well these benefits are sustained over long periods.

**WHAT THIS STUDY ADDS:** The current study shows that the Early Start program of home visitation has benefits in terms of reducing child abuse, increasing parental competence, and improving childhood behavioral adjustment for up to 9 years, suggesting long-term benefits of home visitation.

**abstract**

**OBJECTIVE:** To evaluate the extent to which a home-visitation program (Early Start) had benefits for child abuse, child behavior, and parental- and family-level benefits to the 9-year follow-up.

**METHODS:** A randomized controlled trial in which 220 families receiving Early Start were contrasted with a control series of 223 families not receiving the program. Families were enrolled in the program for up to 5 years. Outcomes were assessed at 6 months, annually from 1 year to 6 years, and at 9 years after trial entry.

**RESULTS:** Comparisons between the Early Start and control series showed that families in the Early Start program showed significant ($P < .05$) benefits in reduced risk of hospital attendance for unintentional injury, lower risk of parent-reported harsh punishment, lower levels of physical punishment, higher parenting competence scores, and more positive child behavioral adjustment scores. Effect sizes (Cohen’s “$d$”) ranged from 0.13 to 0.29 (median = 0.25). There were no significant differences (all $P$ values > .05) between the Early Start and control series on a range of measures of parental behavior and family outcomes, including maternal depression, parental substance use, intimate partner violence, adverse economic outcomes, and life stress.

**CONCLUSIONS:** The Early Start program was associated with small to moderate benefits in a range of areas relating to child abuse, physical punishment, child behavior, and parenting competence. There was little evidence to suggest that the Early Start program had benefits that extended to the level of parents or family overall. *Pediatrics* 2013;131:297-303
There has been increasing evidence that children reared in families facing multiple challenges are at increased risks of a range of adverse outcomes in later life, which has led to interventions to mitigate the risks faced by children born into these environments. One approach to addressing these issues has been intensive home-visiting programs, which provide families with young children with support, mentorship, and assistance with child-rearing and related family issues. These programs include the Nurse Family Partnership, Healthy Families America, Healthy Start, Early Head Start, and Early Start. All of these programs have been evaluated by using randomized control designs but findings from these trials have been mixed, with some programs showing benefits and others failing to show benefits. In a recent review, Howard and Brooks-Gunn found that home-visiting programs had reported benefits for a number of outcomes, including child abuse, child health care, quality of home environment, parenting, parental depression, and childhood cognitive skills.

A previous article reported the outcomes of a randomized trial of the New Zealand Early Start program, which provides intensive home visiting to families facing multiple challenges. Home visiting is provided by trained family support workers with qualifications in nursing, teaching, or allied disciplines. Outcomes at 36-month follow-up showed that children in the Early Start series had higher rates of general practitioner contact (P < .05), higher rates of well-child care (P < .05), lower rates of hospital attendance for unintentional injury (P < .01), lower rates of parentally reported child abuse (P < .01), greater use of preschool education (P < .05), more positive and less punitive parenting (P < .05), and lower rates of childhood behavioral problems (P < .05). No benefits were found for a series of outcomes relating to parental and family circumstances, including maternal depression, family violence, parental substance use, family material conditions, family income, and welfare dependence.

This article reports an analysis of a 9-year follow-up of the Early Start and control families. The aims of this analysis were to examine the extent to which

1. Gains made by the Early Start series in health, child abuse, parenting, and behavioral and related outcomes at 36 months were sustained over the 9-year follow-up period.
2. There was evidence of benefits for parental behavior or family circumstances of Early Start over the 9-year follow-up period.

METHODS

Overview of Research Design

The evaluation of the Early Start program used a randomized control design in which the outcomes of 220 families enrolled in the program were compared with the outcomes of 223 control families. Assessments were made at baseline, 6 months, annually from 1 year to 6 years, and then at 9 years after trial enrollment. Sample size was determined by the availability of funding for services provided to the Early Start series. The study design had the ethical approval of the Canterbury Ethics Committee. No adverse effects of the intervention were observed.

Client Recruitment, Randomization, and Sample Retention

Trial recruitment took place over a 19-month period beginning in January 2000 during which Plunket nurses throughout the Christchurch area screened families by using a procedure described in previous articles. Figure 1 provides a Consolidated Standards of Reporting Trials diagram of the screening, recruitment, and follow-up process. The figure shows that a total of 4523 families were seen by Plunket nurses. Of these families, 588 (13%) were deemed to be eligible for the trial on the basis of the screening criteria, and 443 (75%) agreed to participate.

At the point of referral to the trial, families were randomized to either Early Start or control series based on a computer-generated series of random numbers, and no blinding to assignment was undertaken. Of the 443 families referred to the trial, 220 were assigned to the Early Start series and 223 to the control series. Families in the Early Start series were contacted by family support workers and invited to participate in the service, and were given an opportunity to try out the program for 1 month before actual enrollment. All families assigned to the Early Start service were studied, irrespective of the result of the 1-month assessment. Control series families were contacted by survey interviewers and invited to participate, and paid NZD $50 per interview. Of those invited to participate in Early Start, 206 (96%) agreed to enter the service, whereas 221 (99%) in the control series agreed to participate in the trial. By 9 years, the drop-out rate in the Early Start group was 16% (35/220) compared with 11% (24/223) for the control group. At 9 years, 171 (78%) of the Early Start group was studied compared with 199 (89%) of the control group. These differences were statistically significant (P < .001) and largely attributable to the 14 families who declined to enter Early Start and the trial.

Service Delivery

Early Start, described in detail previously, uses a social learning model approach to home visitation. The critical elements of this model include (1) assessment of family needs, issues,
challenges, strengths, and resources; (2) development of a positive partnership between the family support worker and client; (3) collaborative problem solving to devise solutions to family challenges; (4) the provision of support, mentoring, and advice to assist client families to mobilize their strengths and resources; and (5) involvement with the family throughout the child's preschool years. The Supplemental Information provides details of needs assessment, family support workers, client load, service goals, and service delivery principles.

Research Assessments
Participants in the trial were assessed at baseline, 6 months, annually from 1 year to 6 years, and at 9 years from trial enrollment, by using the following:

1. Client interviews. Client families were assessed on a structured interview administered in the clients’ homes by a trained survey interviewer. Interviews typically lasted between 45 minutes to an hour.

2. Medical record data. Parents were asked to give signed consent for the research group to access the hospital records for the child enrolled in the study. More than 99% of parents provided this consent.

3. Teacher questionnaires. At 5, 6, and 9 years, teachers completed a questionnaire about the child's behavioral adjustment at school.

4. Parenting competence. This was measured by the Parenting Scale, a 30-item scale spanning a number of domains of parenting behaviors, including laxness of discipline, parental overreactions, and parental verbosity. Factor analysis showed that these items formed a unidimensional factor reflecting the extent of parenting competence. The scale reliability ranged from $\alpha = 0.82$ to 0.83.

5. Child behavior (parent and teacher reports). Parent and teacher reports of child behavior problems defined as attendance at an accident and emergency department, or admission to a hospital, was obtained from local (Christchurch and the surrounding region) and national (New Zealand) hospital records subject to the signed consent of the child’s parents. These records provided a complete listing of all hospital attendances for any study participant.

2. Harsh punishment/Parental use of physical punishment. Parents were interviewed on the Parent–Child Conflict Tactics Scale at each assessment from the 1- to 6-year follow-up and again at the 9-year follow-up. Two measures were constructed:

- A measure of whether either parent used harsh or abusive punishment, which included shaking, hitting with an object, hitting with a fist or kicking, grabbing by the neck/choking, hitting as hard as possible, burning/scalding, throwing/knocking down, and slapping on the head.

- A diversity score measure of the reported use of physical punishment, constructed by summing the number of different types of physical punishment methods reported by the parents over the follow-up periods of 1 to 3 years, 4 to 6 years, and 9 years.

Figure 1
Consolidated Standards of Reporting Trials diagram of the screening, recruitment, and follow-up process for the Early Start program.
were obtained at ages 5, 6, and 9 years by using the Strengths and Difficulties Questionnaire (SDQ). The SDQ is a 30-item scale designed to assess a number of child behavior domains, including externalizing behaviors (conduct problems and hyperactivity/inattention) and internalizing behaviors (emotionality and peer difficulties) during the 6 months before assessment. An overall score representing total difficulties was computed by summing the externalizing and internalizing behavior subscales; α coefficients for parent and teacher reports ranged from 0.70 to 0.85.

5. Maternal depression. This was assessed at the 5-, 6-, and 9-year follow-up by using questions from the Composite International Diagnostic Inventory measuring Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition symptoms of major depression. The measure of depressive symptomatology consisted of the total number of symptoms reported by participants in the 12 months before each assessment point.

6. Parental substance use. This was assessed via 3 measures: (1) parental cigarette smoking (for the month before assessment) at the 5-, 6-, and 9-year follow-up; (2) parental alcohol problems, based on questions from the Composite International Diagnostic Inventory relating to Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition alcohol abuse/dependence symptom criteria, in the 12 months before the 5-, 6-, and 9-year follow-up; and (3) parental use of cannabis or other illicit drugs, since the previous assessment, at the 5-, 6-, and 9-year follow-up.

7. Family violence. Intimate partner violence was assessed for the 12 months before the 5-, 6-, and 9-year assessments by using the Revised Conflict Tactics Scale. This instrument measures the domains of minor psychological aggression, severe psychological aggression, minor physical assault, and severe physical assault.

8. Family economic circumstances. Family economic circumstances at the 5-, 6-, and 9-year follow-up were assessed via 3 measures: (1) whether either parent was in receipt of any welfare benefit at each point of assessment; (2) the total amount of debt reported by families at the time of each assessment; and (3) the number of hardship items endorsed by parents by using the Economic Living Standards Index.

9. Family life stress. The family’s exposure to adverse life events in the previous 12 months was obtained at the 5-, 6-, and 9-year follow-up by using a 47-item questionnaire based on the Holmes and Rahe Social Readjustment Rating Scale.

Statistical Analyses

The results in Tables 1 and 2 compare the Early Start and control series on outcomes assessed to 9 years, and summarize more extensive repeated measures data that are reported in the Supplemental Information. To test for differences between the Early Start and control groups, a unified analysis methodology was used by fitting the model in equation 1 to each outcome measure:

\[ f(Y_t) = B0 + B1X + B2t + B3(X_t) + U \]

where \( Y_t \) is the measure of the outcome observed at time \( t \), \( X \) is a dichotomous variable representing treatment status (Early Start/control), and \( X_t \) is an interaction term. This model addresses 3 questions:

1. Treatment effects. The parameter \( B1 \) estimates the overall treatment difference in the outcome \( Y \) over the observation periods.

2. Time effects. The parameter \( B2 \) represents changes in the expected value of \( Y \) over time.

3. Time by treatment interaction. The parameter \( B3 \) represents the extent to which any effects of treatment on the outcome \( Y \) varied with time.

This model was fitted by using population-averaged Generalized Estimating Equation methods. Effect sizes were obtained by computing Cohen’s d for the results pooled across the assessment periods. All tests of significance were 2-tailed.

To test for the multivariate effects of the intervention in Tables 1 and 2, multivariate tests of significance were conducted. The methods used are described in detail in the Supplemental Information.

### TABLE 1 Child Outcomes up to the 9-Year Follow-up

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Controls</th>
<th>Early Start</th>
<th>Cohen’s d (95% Confidence Interval)</th>
<th>( P )</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Attending hospital for unintentional injury (0 to 9 y)</td>
<td>42.1</td>
<td>28.3</td>
<td>0.29 (0.09–0.49)</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>% Parent-reported harsh punishment (0 to 9 y)</td>
<td>20.1</td>
<td>9.8</td>
<td>0.29 (0.09–0.49)</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Mean (SD) physical punishment score (0 to 9 y)</td>
<td>1.44 (1.30)</td>
<td>1.29 (1.13)</td>
<td>0.13 (0.01–0.24)</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>Mean (SD) parenting competence score (5, 6, 9 y)</td>
<td>9.88 (1.05)</td>
<td>10.13 (0.92)</td>
<td>0.25 (0.13–0.37)</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Mean (SD) total parent-reported SDQ score (5, 6, 9 y)</td>
<td>10.08 (1.06)</td>
<td>9.91 (0.91)</td>
<td>0.17 (0.06–0.29)</td>
<td>&lt;.05</td>
</tr>
</tbody>
</table>

* Test of significance obtained from Generalized Estimating Equation model.
Table 2 Parent and Family Outcomes From 5-Year to 9-Year Follow-up

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Controls</th>
<th>Early Start</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (SD) depression symptoms</td>
<td>1.39 (2.97)</td>
<td>1.55 (3.12)</td>
<td>&gt; .20</td>
</tr>
<tr>
<td>% Reporting cigarette smoking</td>
<td>63.7</td>
<td>65.9</td>
<td>&gt; .40</td>
</tr>
<tr>
<td>% Reporting alcohol problems</td>
<td>7.7</td>
<td>8.5</td>
<td>&gt; .80</td>
</tr>
<tr>
<td>% Used cannabis/other illicit drugs</td>
<td>18.5</td>
<td>24.1</td>
<td>&lt; .10</td>
</tr>
<tr>
<td>Mean (SD) IPV victimization score</td>
<td>9.98 (0.84)</td>
<td>10.02 (1.15)</td>
<td>&gt; .60</td>
</tr>
<tr>
<td>Mean (SD) IPV perpetration score</td>
<td>9.99 (0.96)</td>
<td>10.01 (1.04)</td>
<td>&gt; .80</td>
</tr>
<tr>
<td>% Welfare dependent</td>
<td>56.8</td>
<td>59.5</td>
<td>&gt; .50</td>
</tr>
<tr>
<td>Mean (SD) debt (in New Zealand dollars)</td>
<td>4492 (8524)</td>
<td>5248 (9140)</td>
<td>&gt; .20</td>
</tr>
<tr>
<td>Mean (SD) no. of hardship factors</td>
<td>3.46 (3.56)</td>
<td>3.52 (3.44)</td>
<td>&gt; .80</td>
</tr>
<tr>
<td>Mean (SD) adverse life events score</td>
<td>9.97 (0.88)</td>
<td>10.03 (1.02)</td>
<td>&gt; .40</td>
</tr>
</tbody>
</table>

IPV, intimate partner violence.
* Test of significance obtained from Generalized Estimating Equation model.

RESULTS

Child Outcomes up to the 9-Year Follow-up

Table 1 compares the Early Start and control series on measures assessed over the 9-year follow-up. In the interests of brevity, this table reports the results to the 9-year follow-up. The Supplemental Information provides a more extensive account of the data on which Table 1 is based. The table shows that by the 9-year follow-up, children in the Early Start series had:

1. Overall rates of hospital admission (28.3%) that were 33% lower than rates for the control series (42.1%; P < .05; d = 0.29).
2. Overall rates of parentally reported harsh punishment (9.8%) that were more than 50% lower than those for the control series (20.1%; P < .05; d = 0.29).
3. Pooled mean physical punishment scores (1.29) that were 0.13 SDs lower than those of the control series (1.44; P < .05; d = 0.13).
4. Pooled mean parenting competence scores (10.13) that were 0.25 SDs higher than those of the control series (9.88; P < .01; d = 0.25).
5. Pooled mean parentally reported total SDQ behavior scores (9.91) that were 0.17 SDs lower than the control series (10.08; P < .05; d = 0.17).

As shown in the Supplemental Information, for all comparisons there were no statistically significant treatment x time of measurement interactions. A multivariate regression analysis of the 5 outcomes showed that there was a statistically significant overall effect, X^2(5) = 22.56, P < .001. The analysis in Table 1 was extended to consider teacher-reported measures of behavior over the period from 5 to 9 years. This analysis showed that there were no statistically significant differences between the Early Start and control groups on teacher reports of overall SDQ scores.

Parental Behavior and Family Circumstances

Table 2 compares the Early Start and control groups on a series of measures of parent behavior and family circumstances, including maternal depression, substance use, family violence, economic circumstances, and adverse life events assessed over the period from 5 to 9 years. In all cases, there were no statistically significant differences (all P values > .05) between the Early Start and control groups. A multivariate regression analysis of the 5 outcomes showed that the joint effect of the 10 outcomes was statistically non-significant, X^2(10) = 6.30, P > .70.

Analysis of Losses to Follow-up

As shown in Fig 1, losses to follow-up for the Early Start series (22%) were substantially and significantly higher (P < .01) than losses to follow-up for the control series (11%). The principal reason for this was that 14 families randomized to Early Start declined to enter the service or participate in the evaluation.

To examine the effects of sample loss on trial validity, 2 analyses were conducted:

1. Comparison on baseline measures. The Early Start and control group members studied at the 9-year follow-up were compared on 35 measures assessed at the baseline assessment. These measures spanned sociodemographic and ethnic background, economic circumstances, maternal childhood disadvantage, maternal adjustment, and pregnancy and childbirth. This analysis was extended to examine differences between those who dropped out before the 9-year follow-up and those who did not drop out, in both the Early Start and control groups.

2. Imputation of missing data. In addition, the data were reanalyzed by using a full intention-to-treat design by imputing data for those participants not observed throughout the 9-year follow-up. This analysis produced findings that were consistent with the analysis of the observed data in Table 1 (see Supplemental Information).

DISCUSSION

The findings of this 9-year assessment led to 3 general conclusions. First, significant benefits were observed for a number of child outcomes, including childhood hospital attendance for...
unintentional injury (P < .05), parentally reported harsh discipline (P < .05), punitive parenting (P < .05), parental competence (P < .01), and parentally reported child behavioral problems (P < .05). In all cases, tests of treatment × time of measurement interaction suggested that these benefits were sustained over the 9-year follow-up. Effect sizes (Cohen’s d) ranged from 0.13 to 0.29 with a median of 0.25.

Second, there was no evidence to suggest that Early Start had benefits for a range of parental behavior and family circumstances, including maternal depression, parental substance use, family violence, family income, family debt, and welfare dependence.

Third, the rate of sample attrition in the Early Start series was twice that of the control series. Tests of sample selection bias were conducted by (1) comparing the Early Start and control groups, and those observed and not observed at the 9-year follow-up on a series of 35 measures obtained at baseline; and (2) the use of imputation methods to produce a full intention-to-treat analysis. Both approaches suggested that it was unlikely that differential sample losses posed a substantial threat to study validity.

These findings are consistent with the 36-month assessment and show that the child-related benefits of Early Start were sustained over the 9-year follow-up. As was the case with the 36-month follow-up, effect sizes were in the range of small to moderate. An important issue concerns the apparent failure of Early Start to produce change in parental behavior and family circumstances. The reasons for this finding are not clear; however, discussions with Early Start staff suggest 3 reasons that conspired to reduce the effectiveness of the program in producing such changes. First, as reflected in the literature, parents were often reluctant and resistant to seek treatment and advice for issues such as depression, substance abuse, and contraception. Second, the uptake of services was limited by the availability of services, in particular drug and alcohol services. Finally, parents were often noncompliant with the advice with which they were provided, particularly with budget advice. These 3 factors would appear to explain the reasons for the lack of success of Early Start and other home-visiting programs to produce significant parental and family-level change.

CONCLUSIONS

The results of this 9-year trial show that programs like Early Start can produce modest changes in outcomes relating to childhood well-being, including unintentional injury, harsh discipline, parenting competence, and child behavior; however, these benefits do not generalize to family-level change. In part, these results are likely to reflect the fact that Early Start, as well as other home-visiting programs with similar features, provide advice and mentorship but do not provide direct therapeutic support. These considerations suggest that future directions in this area should involve the closer integration of home-visiting services like Early Start into organizations providing health, educational, and behavioral support. Under ideal circumstances, home-visiting teams should be part of integrated organizations of medical practitioners, educationalists, and psychologists that provide integrated systems of care and support for families facing multiple challenges.

REFERENCES

12. Gilman SE, Kawachi I, Fitzmaurice GM, Buka SL. Socioeconomic status in childhood and...


Nine-Year Follow-up of a Home-Visitation Program: A Randomized Trial
David M. Fergusson, Joseph M. Boden and L. John Horwood
*Pediatrics* 2013;131;297; originally published online January 28, 2013;
DOI: 10.1542/peds.2012-1612

<table>
<thead>
<tr>
<th>Updated Information &amp; Services</th>
<th>including high resolution figures, can be found at: /content/131/2/297.full.html</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supplementary Material</td>
<td>Supplementary material can be found at: /content/suppl/2013/01/23/peds.2012-1612.DCSupplemental.html</td>
</tr>
<tr>
<td>References</td>
<td>This article cites 35 articles, 4 of which can be accessed free at: /content/131/2/297.full.html#ref-list-1</td>
</tr>
<tr>
<td>Subspecialty Collections</td>
<td>This article, along with others on similar topics, appears in the following collection(s): <em>Administration/Practice Management</em> /cgi/collection/administration:practice_management_sub</td>
</tr>
<tr>
<td>Permissions &amp; Licensing</td>
<td>Information about reproducing this article in parts (figures, tables) or in its entirety can be found online at: /site/misc/Permissions.xhtml</td>
</tr>
<tr>
<td>Reprints</td>
<td>Information about ordering reprints can be found online: /site/misc/reprints.xhtml</td>
</tr>
</tbody>
</table>

PEDIATRICS is the official journal of the American Academy of Pediatrics. A monthly publication, it has been published continuously since 1948. PEDIATRICS is owned, published, and trademarked by the American Academy of Pediatrics, 141 Northwest Point Boulevard, Elk Grove Village, Illinois, 60007. Copyright © 2013 by the American Academy of Pediatrics. All rights reserved. Print ISSN: 0031-4005. Online ISSN: 1098-4275.
Nine-Year Follow-up of a Home-Visitation Program: A Randomized Trial
David M. Fergusson, Joseph M. Boden and L. John Horwood
Pediatrics 2013;131;297; originally published online January 28, 2013;
DOI: 10.1542/peds.2012-1612

The online version of this article, along with updated information and services, is located on the World Wide Web at:
/content/131/2/297.full.html