Results at Age 8 Years of Early Intervention for Low-Birth-Weight Premature Infants

The Infant Health and Development Program

Cecelia M. McCarton, MD; Jeanne Brooks-Gunn, PhD; Ina F. Wallace, PhD; Charles R. Bauer, MD; Forrest C. Bennett, MD; Judy C. Bernbaum, MD; R. Sue Broyles, MD; Patrick H. Casey, MD; Marie C. McCormick, MD, ScD; David T. Scott, PhD; Jon Tyson, MD; James Tonascia, PhD; Curtis L. Meinert, PhD; for the Infant Health and Development Program Research Group

Objective.—To reevaluate at age 8 years children who had participated during the first 3 years of life in a randomized clinical trial of special services for low-birth-weight (LBW) premature infants.

Design.—Follow-up of a randomized controlled trial of premature infants (≤37 weeks' gestation), stratified by 2 LBW groups (lighter [≤2000 g] and heavier [2001-2500 g]) and divided into intervention (n=377) and follow-up only (n=608) groups.

Setting.—Eight sites serving diverse populations.

Participants.—At age 8 years, 874 children were assessed: 336 in the intervention group and 538 in the follow-up only group.

Intervention.—The 3-year intervention consisted of home visits (birth to 3 years), child development center services (ages 1 to 3 years), and parent group meetings (ages 1 to 3 years).

Primary Outcome Measures.—Cognitive functioning (Weschler Intelligence Scale for Children–III; Peabody Picture Vocabulary Test–Revised); academic achievement (Woodcock-Johnson Tests of Achievement–Revised); and parental reports of school performance, behavior (Child Behavior Checklist), and health (Child General Health Survey).

Results.—At age 8 years, in the entire cohort and in the lighter LBW stratum, the intervention and follow-up only groups were similar on all primary outcome measures. Differences favoring the intervention group were found within the heavier LBW group: full-scale IQ score (4.4 points higher, P=.007), verbal IQ score (4.2 points higher, P=.01), performance IQ score (3.9 points higher, P=.02), mathematics achievement score (4.8 points higher, P=.04), and receptive vocabulary score (6.7 points higher, P=.001). On a physical functioning subscale, the whole intervention group received less favorable ratings, while the lighter LBW intervention group had lower maternal ratings assessing social limitations caused by behavior.

Conclusion.—Although at age 8 years there were modest intervention-related differences in the cognitive and academic skills of heavier LBW premature children, attenuation of the large favorable effects seen at 3 years was observed in both the heavier and lighter LBW groups. This indicates a need to develop additional intervention strategies for LBW premature children that can provide sustained benefits.

JAMA. 1997;277:126-132

From the Albert Einstein College of Medicine, Bronx, NY (Dr McCarton); Center for Children and Families, Teachers College, Columbia University, New York, NY (Dr Brooks-Gunn); Research Triangle Institute, Research Triangle Park, NC (Dr Wallace); Department of Pediatrics, University of Miami School of Medicine, Miami, Fla (Dr Bauer); Department of Pediatrics, University of Washington, Seattle (Drs Bennett and Scott); Children's Hospital of Philadelphia, University of Pennsylvania (Dr Bernbaum); Department of Pediatrics, University of Arkansas for Medical Sciences, Little Rock (Dr Casey); Department of Maternal and Child Health, Children's Hospital, Boston, Mass (Dr McCormick); Department of Pediatrics, University of Texas Southwestern Medical Center, Dallas (Drs Tyson and Broyles); and Department of Biostatistics and Epidemiology, Johns Hopkins University, Baltimore, Md (Drs Tonascia and Meinert).

For a complete list of participants in the Infant Health and Development Program, see *JAMA.* 1990;263:3035-3070, and *JAMA.* 1994;272:1257-1262.

Reprints: Cecelia McCarton, MD, Albert Einstein College of Medicine, Rose F. Kennedy Center, Room 820, 1410 Pelham Parkway South, Bronx, NY 10461.

ON GLOBAL MEASURES of intelligence and other neurodevelopmental assessments, low-birth-weight (LBW) premature children perform more poorly than normal-birth-weight children.[1-3] In addition, by school age, LBW children are at increased risk of learning disabilities,[4,5] academic difficulties,[6-8] and behavior problems,[4,6,7,9] even in the absence of global intellectual deficits. Few research efforts have been devoted to determining whether early intervention could attenuate the long-term problems seen in LBW preterm infants.[10,11]

The Infant Health and Development Program (IHDP) was designed as a multisite randomized clinical trial evaluating in the first 3 years of life the efficacy of center-based educational intervention, home-based family support services, and pediatric follow-up in reducing cognitive, behavioral, and health problems among LBW premature infants. The cognitive and behavioral outcomes of the IHDP trial at ages 3 and 5 years have been reported,[12-15] as have the health outcomes at age 3 years.[12,16] At 3 years, the children in the intervention group had significantly higher intelligence test scores and receptive vocabulary test scores and lower scores on a parental measure of reported behavior problems than the children in the follow-up only group. The rate of maternally reported health conditions over the first 3 years was greater for the children in the intervention group (difference of 0.27), although they were not hospitalized to a greater extent than the children in the follow-up only group. The advantages at age 3 years conferred by the intervention were more pronounced in the heavier LBW stratum (2001-2500 g) than in the lighter LBW stratum (≤2000 g) in terms of IQ score (14 vs 7 points), receptive vocabulary score (9 vs 5 points), and behavior problem score (7 vs 2

points). In addition, the rate of maternally reported health problems was greater in the lighter LBW intervention group than in the lighter LBW follow-up only group (difference of 0.30); no differences in the rate of maternally reported health problems between the intervention and follow-up only groups were found in the heavier LBW stratum.

At age 5 years, 2 years after the intervention ended, we saw an attenuation of IQ effects. Specifically, there were no significant overall differences in IQ score, receptive vocabulary, reported behavior problems, or health measures between the intervention and follow-up only children. However, within the heavier LBW stratum, the intervention group had higher full-scale IQ scores (4 points) and verbal IQ scores (4 points) as well as higher receptive vocabulary scores (6 points) than the follow-up only group.

The present report extends the follow-up to age 8 years. Age 8 years was chosen because it is one age at which increased academic demands may be placed on a child.[17] As such, it is a common time for grade failure and academic problems to emerge. Our hypotheses were that enhancements of global measures of cognitive function, such as IQ score, that were found at 3 years would be attenuated by age 8 years, but that significant differences favoring the intervention group would be found in school performance measures of reading and mathematics achievement and in reduced rates of grade failure.

METHODS

Detailed information on the selection of participants, study design, and components of the intervention program has been published previously[12-18] and will be summarized here briefly. Infants were eligible for the study if they had a birth weight of 2500 g or less, had a gestational age of 37 weeks or less, resided in the catchment area, and did not have a severe medical illness or neurological impairment. Patients were enrolled from October 1984 through August 1985. A total of 985 infants constituted the primary analysis group for the trial. These infants were randomly assigned to the intervention group (n=377) or the follow-up only group (n=608) using a design with 2 birth-weight strata: lighter LBW (≤2000 g [n=623]) and heavier LBW (2001-2500 g [n=362]).[12,15,19] Although the baseline characteristics of the study sample varied greatly across the 8 sites, the randomization procedure resulted in comparable intervention and follow-up only groups at study entry.[12]

Infants in both the intervention and follow-up only groups received the same periodic (through age 8 years) medical, developmental, and social assessments, with referrals made for other services as needed. The early-intervention program began at discharge from the neonatal nursery and continued until October 1988—until each child was at least 36 months of age, corrected for prematurity. The intervention had 3 components[13]: home visits through age 3 years, attendance at specially designed child development centers beginning at age 1 year through age 3 years, and a series of parent group meetings every other month during the second and third years. Home visits were designed to occur weekly in the first year and biweekly in the second and third years. Over the 3 years an average of 66.7 home visits were made. Attendance at the child development centers was to be for full days 50 weeks per year. The children attended the centers an average of 267 days per year (over both years). Parent group meetings were to be held 4 times per year; on average (over 2 years), there were 3.7 parent group meetings per year.

After the intervention ended at age 3 years, the sites attempted to find appropriate community education programs for children in both groups. Resources available at each site differed greatly; however, within the sites there were no intervention differences between the intervention and follow-up only groups in maternal reports of enrollment in education programs at age 4 years.

Of the original 985 infants in the primary analysis group, 874 (89%) were evaluated at age 8 years for a comprehensive assessment of cognitive functioning, academic achievement, behavior, and health. Individual sites differed in the percentage of children they were able to evaluate, ranging from 84% to 97%. The percentages of children within the intervention and follow-up only groups were similar, both overall and within the sites. The mean age of assessment was 8.1 years (range, 7.9-8.9 years).

The cognitive, academic achievement, and behavior evaluations at age 8 years were performed by centrally trained assessors who were masked as to the child's treatment group and history. Health status was assessed by clinic staff who had access to the child's treatment group assignment and history. As with the other evaluations, the assessment at age 8 years was to be conducted at the enrolling IHDP site. Whenever possible, children who had moved away from their enrolling site were assessed at alternative locations, either at a nearby IHDP site or in their own community, by a masked assessor who was sent to evaluate them. Of the 874 children who participated in the masked portion of the assessment, all but 59 (7%) were seen at an IHDP site.

Primary outcome measures were in the domains of cognitive skills, school performance, behavior, and health status. Several measures were administered in each domain. The complete battery is listed in Table 1, but only the primary measures indicated in Table 1 are reported here. The battery of tests given to the child is listed in the order of administration. The battery generally took no more than 3 hours to complete.

Differences between the intervention and follow-up only groups with respect to each baseline characteristic were assessed using t statistics for continuous measures and χ^2 statistics for categorical measures. Separate multiple linear regression models[20] were developed for each outcome measure within each of the 2 birth-weight strata. The same set of 8 explanatory variables measured at enrollment was used in each model: intervention treatment group assignment (intervention vs follow-up only), site (8 categories), birth weight (grams), sex (male vs female), Neonatal Health Index (calculated based on length of stay in the neonatal nursery, adjusted for birth weight, and standardized to a mean of 100, with higher scores indicating better health),[21] maternal education (did not complete high school, completed high school, postsecondary schooling), maternal age (years), and maternal ethnicity (African American, Hispanic, and white/other). Population marginal mean values and differences (intervention vs follow-up only) for each outcome measure were computed from the regression models to determine if the intervention group controlled for potentially confounding effects of 1 or more of the 7 other baseline explanatory variables.[22] The population marginal means estimate the expected outcome measure in either the intervention or follow-up only group with the 7 other explanatory variables equated to their average values within a particular birth-weight stratum. Nominal P values derived from the regression models were used for primary outcome comparisons; no adjustments were made for multiple comparisons or multiple outcomes.[23]

RESULTS

Table 2 compares the intervention and follow-up only groups at age 8 years for baseline data obtained at enrollment. The baseline characteristics of the intervention and follow-up only groups were comparable. Although within each site the baseline characteristics were comparable

Table 1.—Contents of the Age 8 Years Assessment for Children in the Infant Health and Development Program

Measure*	Domains	Respondent
Developmental Test of Visual-Motor Integration (scale, 0-100±15; Beery,[38] 1989)	Cognitive: visual, graphomotor skills	Child
Wechsler Intelligence Scale for Children–III† (scale, 0-100±15; Wechsler,[39] 1991)	Cognitive: verbal IQ, performance IQ, full-scale IQ	Child
Rey-Osterrieth Complex Figure (Waber and Holmes,[40] 1985)	Cognitive: planning, organizing and drawing, complex design	Child
Woodcock-Johnson Tests of Achievement–Revised† (scale, 0-100±15; Woodcock and Johnson,[41] 1990)	School performance: reading and math skills	Child
Matrices (scale, 50±10; Elliott,[42] 1990)	Cognitive: nonverbal reasoning	Child
Peabody Picture Vocabulary Test–Revised† (scale, 0-100±15; Dunn and Dunn,[43] 1981)	Cognitive: receptive vocabulary	Child
Wide Range Assessment of Memory and Learning: Story Memory (scale, 0-10±3; Sheslow and Adams,[44] 1990)	Cognitive: memory for connected discourse	Child
Behavior Rating Profile–2 (Brown and Hammill,[45] 1990)	Behavior	Child
Growth measures	Health	Child
Psychological Examination Behavior Profile (Buka et al,[9] 1992)	Behavior	Child assessor
Child Behavior Checklist† (Achenbach and Edelbrock,[46] 1983)	Behavior	Mother
Grade retention and special education†	School performance	Mother
Health questionnaire	Health	Mother
Child General Health Survey† (Landgraf et al,[47] 1993)	Health: health-related quality of life	Mother

*Child measures are listed in the order of administration. Scale values are given in the form, range±SD.
†Primary outcome measure.

Table 2.—Baseline Characteristics of Children in the Infant Health and Development Program

Characteristic	At Enrollment		At Age 8 y	
	Intervention Group	Follow-up Only Group	Intervention Group	Follow-up Only Group
Total No. (% of randomized infants)	377 (100.0)	608 (100.0)	336 (89.1)	538 (88.5)
Birth weight, g*	1819±439	1781±468	1799±439	1787±466
P	.20		.68	
Birth-weight group, %				
≤2000 g	62.3	63.8	64.6	63.9
2001-2500 g	37.7	36.2	35.4	36.1
P	.64		.85	
Gestational age, wk*	33.0±2.5	33.0±2.8	32.9±2.5	33.0±2.8
P	.82		.52	
Neonatal Health Index*	100.7±15.8	99.6±16.1	100.6±15.9	99.9±15.4
P	.26		.50	
Maternal age at enrollment, y*	24.6±5.9	24.9±6.1	24.5±5.8	25.0±6.1
P	.45		.21	
Maternal education, %				
<High school	43.0	38.2	40.8	36.2
High school	27.6	27.3	29.2	28.4
>High school	29.4	34.5	30.1	35.3
P	.20		.24	
Race, % of infants				
African American	53.3	52.0	53.0	52.8
Hispanic	9.8	11.2	9.5	9.8
White/other	36.9	36.8	37.5	37.4
P	.78		.99	
Male, %	50.4	48.7	49.7	49.4
P	.60		.94	
Clinic, distribution omitted, *P*	.24		.53	

*Values are mean±SD.

for the intervention and follow-up only groups, there were differences between the sites, particularly on demographic characteristics, such as level of maternal education and ethnicity.

Cognitive Outcomes

Overall, we found no statistical differences between the intervention and follow-up only groups (Table 3). However, the heavier LBW intervention group displayed significantly higher scores than the heavier follow-up only group on the Weschler Intelligence Scale for Children–III verbal IQ, performance IQ, and full-scale IQ tests. The heavier LBW intervention group also had significantly higher scores on the Peabody Picture Vocabulary Test–Revised. No differences were found between the lighter LBW intervention and follow-up only groups.

We also examined whether there were sustained effects on IQ scores of the intervention as a function of the mother's level of education. In the entire group as well as the 2 LBW strata, there were significant IQ differences between the children as a function of the mother's level of education. Children of mothers with post–high school education had higher IQ scores than children of mothers who had completed high school, while children of mothers with less than 12 years of education had the lowest IQ scores. The differences between the intervention and follow-up only groups, however, were consistent across the 3 maternal education groups. Within the heavier LBW stratum, there was a 4-point difference between the intervention and follow-up only groups for each of the maternal education subgroups—the same intervention-related difference as within the entire heavier LBW stratum. However, there were no IQ differences between the intervention and follow-up only groups for any maternal education subgroup in the group as a whole or in the lighter LBW stratum.

School Performance Outcomes

Although there were no overall differences between the intervention and follow-up only groups on either composite of the Woodcock-Johnson Tests of Achievement–Revised (Table 4), the heavier LBW intervention group had significantly higher Woodcock-Johnson broad mathematics scores than the heavier LBW follow-up only group. No intervention differences were found in either the reading or mathematics scores within the lighter LBW stratum.

Parents were asked whether children had repeated a grade or were classified as needing special education. Results indicate that the percentages of grade repetition and of children classified for special education were similar in the overall intervention and follow-up only groups and within the 2 birth-weight strata.

Behavior Outcomes

Results for the primary behavior measure are shown in Table 5. When rated

Table 3.—Cognitive Measures at Age 8 Years for Childen in the Infant Health and Development Program*

	Birth Weight, g											
	≤2000								2001-2500			
	Intervention Group	Follow-up Only Group	Difference, Intervention vs Follow-up Only		Intervention Group	Follow-up Only Group	Difference, Intervention vs Follow-up Only		Intervention Group	Follow-up Only Group	Difference, Intervention vs Follow-up Only	
			Mean Score	P			Mean Score	P			Mean Score	P
No. of children	336	538	...	...	217	344	...	...	119	194	...	...
Wechsler Intelligence Scale for Children–III												
Verbal IQ	93.4	92.4	0.9	.36	91.3	91.3	0.1	.97	98.5	94.3	4.2	.01
Performance IQ	89.6	90.0	−0.3	.75	87.2	89.3	−2.1	.12	95.1	91.2	3.9	.02
Full-scale IQ	90.8	90.5	0.3	.77	88.3	89.5	−1.1	.39	96.5	92.1	4.4	.007
Peabody Picture Vocabulary Test–Revised	84.9	84.9	0.1	.96	81.6	84.4	−2.8	.10	92.4	85.7	6.7	.001

*All scores are means. Population marginal means and differences were derived from multiple linear regression models that adjusted for site, race/ethnicity, maternal education, maternal age, and Neonatal Health Index. In addition, the birth-weight group was included in the regression models for the overall sample.

Table 4.—School Performance at Age 8 Years for Children in the Infant Health and Development Program

	Birth Weight, g											
	≤2000								2001-2500			
	Intervention Group	Follow-up Only Group	Difference, Intervention vs Follow-up Only		Intervention Group	Follow-up Only Group	Difference, Intervention vs Follow-up Only		Intervention Group	Follow-up Only Group	Difference, Intervention vs Follow-up Only	
			Mean Score	P			Mean Score	P			Mean Score	P
Woodcock-Johnson Test of Achievement–Revised*												
No. of children	336	532	...	...	217	340	...	...	119	192	...	...
Mean score												
Broad reading	98.2	96.7	1.6	.24	96.4	95.4	1.0	.56	102.2	98.8	3.4	.10
Broad math	95.9	95.3	0.6	.70	93.0	93.6	−0.6	.74	102.6	97.7	4.9	.04
Maternal Interviews												
No. of interviews	338	536	...	...	218	344	...	...	120	192	...	...
Mother rating, %												
Repeated grade	14.2	15.1	...	.72	13.4	15.1	...	.56	15.8	15.1	...	.86
Classified as needing special education	17.2	20.0	...	.30	20.3	20.7	...	.90	11.7	18.8	...	.09

*All scores are means. Population marginal means and differences were derived from multiple linear regression models that adjusted for site, race/ethnicity, maternal education, maternal age, and Neonatal Health Index. In addition, the birth-weight group was included in the regression models for the overall sample.

by parents on the Child Behavior Checklist, the intervention group had scores equivalent to those of the follow-up only group, and the 2 birth-weight strata were also comparable. Moreover, the percentages of children with high scores indicative of serious behavior problems were similar in the intervention and follow-up only groups and within the 2 birth-weight strata, with between 8% and 12% rated as above the cutpoint.

Health Outcomes

Responses to questions regarding illnesses and hospitalizations revealed that overall the groups had similar numbers of hospitalizations, surgical procedures, and school absences. As seen in Table 5, the intervention and follow-up only groups received similar ratings on most scales of the Child General Health Survey. However, in the sample as a whole, the intervention group received lower ratings on the Physical Functioning Scale, indicating that these children were perceived as being more limited than those in the follow-up only group in the performance of physical activities, such as playing sports, walking up stairs, bending, lifting, or caring for themselves. While the heavier LBW intervention group was not distinguished from the heavier LBW follow-up only group on any measure of health, the lighter LBW intervention group received lower ratings than the lighter LBW follow-up only group on the scale assessing role/social limitations due to behavior. This indicates that their parents found them, because of behavior problems, to be more limited in the kind, amount, and degree of schoolwork and/or social activities they could perform.

COMMENT

The present report details the long-term follow-up of a multisite randomized clinical trial of early intervention from birth to 3 years of age for LBW children. While the intervention improved cognitive test scores and reduced behavior problems at age 3 years, 5 years after the intervention ended, at age 8 years, overall IQ score, cognitive skills, school achievement, behavior, and health indices were similar in the 2 groups of children. Both the intervention and follow-up only groups had IQ and vocabulary scores substantially below expectations based on standardized norms. Moreover, there was no reduction in grade repetition or use of special educational services. On the other hand, there were statistically significant, albeit clinically modest, differences between the heavier LBW intervention and follow-up only groups on several measures of cognition and school achievement. As at age 5 years, there were no differences found on any

Table 5.—Behavior and Health Measures at Age 8 Years for Children in the Infant Health and Development Program*

	Birth Weight, g											
	≤2000								2001-2500			
	Intervention Group	Follow-up Only Group	Difference, Intervention vs Follow-up Only		Intervention Group	Follow-up Only Group	Difference, Intervention vs Follow-up Only		Intervention Group	Follow-up Only Group	Difference, Intervention vs Follow-up Only	
			Mean Score	P			Mean Score	P			Mean Score	P
Child Behavior Checklist*												
No. of maternal responses	334	536	...	...	216	342	...	...	118	194	...	...
Mean score	32.1	31.9	0.2	.87	33.0	31.9	1.1	.52	30.0	31.3	−1.4	.53
% above cutpoint†	11.1	9.5	...	.45	12.0	11.1	...	.74	9.3	6.7	...	.40
Child General Health Survey[47]*												
No. of maternal responses	338	539	...	...	218	346	...	...	120	193	...	...
Mean score												
General health	82.0	81.8	0.2	.90	81.6	81.0	0.6	.75	82.8	83.4	−0.6	.79
Physical Functioning Scale	95.7	97.5	−1.8	.03	94.6	96.8	−2.2	.07	97.7	98.8	−1.1	.20
Role/Social Limitations Scale												
Due to behavior	86.2	88.6	−2.4	.14	84.4	89.0	−4.7	.02	90.1	88.0	2.1	.46
Due to emotional problems	92.8	93.5	−0.6	.62	92.5	94.1	−1.6	.32	93.5	92.7	0.9	.70
Due to physical health	93.7	95.6	−1.9	.09	93.1	95.2	−2.1	.16	94.9	96.3	−1.4	.40
Behavior Problems Scale	74.0	73.9	0.1	.90	73.3	74.4	−1.1	.32	75.6	73.3	2.3	.12
Mental Health Scale	79.9	79.8	0.1	.87	79.6	79.5	0.1	.91	80.3	80.4	−0.1	.95
Bodily Pain Scale	88.1	88.5	−0.4	.75	88.5	88.5	−0.1	.97	87.5	88.6	−1.1	.58
Self-esteem Scale	85.7	85.7	0.1	.95	85.5	86.0	−0.5	.66	86.1	85.4	0.7	.64
General Health Perceptions	77.7	77.0	0.7	.50	76.4	76.0	0.4	.78	80.4	78.8	1.6	.31

*All scores are means. Population marginal means and differences were derived from multiple linear regression models that adjusted for site, race/ethnicity, maternal education, maternal age, and Neonatal Health Index. In addition, the birth-weight group was included in the regression models for the overall sample.
†The cutpoints for the total score were 54 or greater for girls and 61 or greater for boys.

primary outcome measure at age 8 years between the lighter LBW intervention and follow-up only groups. The diminution of intervention effects after termination of this program is consistent with the findings of numerous studies of normal-birth-weight disadvantaged children.[24-28]

Cognitive Outcomes

The success of the intervention for the heavier LBW children can be gauged by their higher test scores in several cognitive skill areas (ie, intelligence, mathematics, and receptive vocabulary) as well as on some measures of nonreported secondary outcomes (ie, nonverbal reasoning and aspects of organizing the reproduction of a complex, visually presented design). These findings indicate that the intervention had modest success across a range of intellectual functions for this subgroup of infants.

The observed mean difference in full-scale, verbal, and performance IQ scores of 4 points favoring the heavier LBW intervention group is consistent with findings from the Abecedarian project, a 5-year preschool educational intervention for normal-birth-weight socially disadvantaged children that was the basis of the IHDP educational intervention component. At age 8 years, a difference of 2.5 points was found for the Abecedarian intervention group.[29] In addition, the mean difference in mathematics test scores of 5 points is similar to the 9-point difference found in the Abecedarian project. This similarity in intervention effects is notable given that the IHDP target population consisted of infants at biological risk and that the intervention stopped at age 3 years.

In contrast to expectations, no differences in cognitive or behavior measures were detected at 5 years[15] or 8 years of age in the lighter LBW stratum. An intervention lasting only through age 3 years may have been insufficient to sustain long-term effects because the biomedical problems were greater in this group of children. In both the intervention and follow-up only groups, the lighter LBW children had higher rates of neurological abnormalities than the heavier LBW children[30] and had lower IQ, receptive vocabulary, reading, and mathematics scores.

Based on the results at 3 years of age,[31] we expected greater effects for children of less socioeconomically advantaged backgrounds. Using maternal education as a proxy, we found that there was no special benefit for more socioeconomically vulnerable children at age 8 years. However, it should be noted that in the earlier study maternal education was dichotomized, while in the current study it was trichotomized.

School Performance Outcomes

Although it was hypothesized that the effects of early intervention would be most evident in the prevention of school failure, no differences were found in the percentage of children who repeated a grade or who were placed in special education. The grade retention rates of 14% and 15% in the intervention and follow-up only groups, respectively, were about the same as the rates (16% after 3 years of school) found in the Abecedarian program.[32] The kindergarten and first-grade retention rates for LBW children (12.3%) reported in a large epidemiologic study[33] are comparable to those found in the present study but were lower (7.6%) for the overall distribution of children. However, the rates in the present study are lower than that reported for 1 large LBW sample, in which 24% of LBW children who were 9 years old had repeated a grade.[34] The lower rates of grade retention in the present sample of LBW children may reflect differences in school districts' criteria for grade re-

tention and/or policy changes limiting the number of children held back. It may also be that differences in grade repetition may not emerge until the children are somewhat older, when demands for academic performance are even greater, such as during the later years of elementary school. Support for this hypothesis is garnered from the follow-up of children in the Perry Preschool Program,[35] which was a randomized controlled intervention program during the 1960s of half-day preschool and weekly home visits for 3- and 4-year-olds from socioeconomically disadvantaged backgrounds. Data from the follow-up indicated that differences in rates of grade retention and special education placement were not detected between the intervention and control groups until the children had completed third grade.

Behavioral Outcomes

No intervention differences in parental ratings of behavior were found at age 8 years either in the group as a whole or within either birth-weight stratum. The attenuation of differences in this domain is consistent with the findings at age 5 years. However, the percentage of children with high scores in the heavier (7%) and lighter (12%) LBW strata is greater than the 2% expected based on the standardization sample of the Child Behavior Checklist. A higher-than-average rate of behavioral difficulties is consistent with the findings from other follow-up studies of LBW premature children.[9,36]

Health Outcomes

No differences in any measure of illness or hospitalization were found at age 8 years. These results parallel the findings at age 5 years and indicate that the modest elevation in minor illnesses found in the intervention group at age 3 years was limited to the first year in which children were in congregate child care. However, a few differences were found in the Child General Health Survey, which measures children's health-related quality of life. In the group as a whole, the intervention group received lower ratings than the follow-up only group on a measure of physical limitations in behavior. Although no intervention differences in health behaviors were found in the heavier LBW stratum, within the lighter LBW stratum the intervention group received a lower rating than the follow-up only group on the scale assessing role/social limitations due to behavior. The reasons for the few intervention-related differences in health behaviors remain speculative. One possible explanation is that a great deal of training and observation was done with the mothers in the intervention group, making these mothers more accurate observers and reporters of their children's health-related behaviors.

CONCLUSIONS

After an 8-year study of this nature one must ask what has been learned. In particular, what does the average difference of 4 IQ points found in the heavier LBW children mean? Individually, 4 IQ points would not produce a functionally detectable difference between a child in the intervention group and one in the follow-up only group. However, on a group basis, a 4-point IQ difference between the intervention and follow-up only children might reduce the percentages of children classified as intellectually deficient and of borderline intelligence. For instance, approximately 11% of children in the heavier LBW follow-up only group had IQ scores in the borderline range (ie, 70-80). If these children had received the intervention and had gained the same 4-point increase in their IQ score, only 8% of them would have scored between 70 and 80, a percentage closer to that expected in the population at large (7% of 8-year-olds).

At the same time, it is reasonable to question the economic price involved in sustaining a 4-point IQ difference. Unfortunately, no true cost-benefit test is possible. A major limitation of the IHDP is that there were no prospectively obtained cost data for implementing the program across 8 sites, although program expenses were assessed at the Miami site. The cost of delivering the 3 programmatic components was estimated at $15 146 per year per child.[37] The investigators suggested that this high cost could have been reduced to $8806 per year per child if the centers were located in the community rather than at a central location, cutting down transportation costs, and if the teacher-to-child ratio at the child care center were increased from 2:6 to 2:8.

These costs would be offset if grade retention was decreased and the need for special education was reduced. Although we did not detect such differences, it may have been too soon to accurately assess school failure. The majority of children were in first and second grades. The results of the Perry Preschool Program[35] suggest that eventually more children may be retained or placed in special classes. Only future evaluations of the IHDP cohort will be able to assess whether economic benefits will justify the large costs per child.

These results have several implications for future early-intervention program initiatives. First, our findings suggest that programmatic efforts need to distinguish between heavier and lighter LBW infants. The program did help children with birth weights of 2000 g or less at age 3 years, but, clearly, the effects were smaller and not sustained. The lighter LBW group may contain a higher proportion of neurologically impaired children who could not benefit from the intervention. Second, assessments need to be refined enough to detect qualitative differences between heavier and lighter LBW infants. For example, maternal reports of attention deficit or distractibility may reflect environmental conditions for the heavier LBW child but may be related to neurological impairment in the lighter LBW child. The heavier LBW child may respond to changes in environment that the mother can control, but the lighter LBW child may need a more structured and professionally designed situation. Third, the lack of sustained effect may reflect the need for different or continued support of lighter LBW children. It is the task of future research to characterize the populations needing ongoing support; to explore the type, intensity, and duration of interventions needed to produce sustained effects; and to develop policies to implement such interventions. Finally, the corresponding low mean IQ scores and high rates of school failure of LBW children underscore the importance of public health efforts to reduce premature births.

This study was supported by grant 91-01142-00 from the Pew Charitable Trusts, Philadelphia, Pa; by grant 5R01 HD 27344 from the National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Md; and by grants from the Maternal and Child Health Bureau (Title V, Social Security Act), Washington, DC; the Department of Health and Human Services, Washington, DC; and the Robert Wood Johnson Foundation, Princeton, NJ.

References

1. Aylward GP, Pfeiffer SI, Wright A, Verhulst SJ. Outcome studies of low birth weight infants published in the last decade: a metaanalysis. *J Pediatr.* 1989;115:515-520.
2. Escobar GJ, Littenberg B, Petitti DB. Outcome among surviving very low birthweight infants: a meta-analysis. *Arch Dis Child.* 1991;66:204-211.
3. McCormick MC. Long-term follow-up of infants discharged from neonatal intensive care units. *JAMA.* 1989;261:1767-1772.
4. Hack M, Breslau N, Aram D, Weissman B, Klein N, Borawski-Clark E. The effect of very low birth weight and social risk on neurocognitive abilities at school age. *Dev Behav Pediatr.* 1992;13:412-420.
5. Sostek AM. Prematurity as well as intraventricular hemorrhage influence developmental outcome at 5 years. In: Friedman SL, Sigman MD, eds. *The Psychological Development of Low-Birthweight Children: Annual Advances in Applied Developmental Psychology.* Norwood, NJ: Ablex Publishing Corp; 1992;6:259-274.
6. Klebanov PK, Brooks-Gunn J, McCormick MC. Classroom behavior of very low birth weight elementary school children. *Pediatrics.* 1994;94:700-708.
7. McCormick MC, Brooks-Gunn J, Workman-Daniels K, Turner J, Peckham GJ. The health and

developmental status of very low–birth-weight children at school age. *JAMA.* 1992;267:2204-2208.
8. Klein NK, Hack M, Breslau N. Children who were very low birth weight: development and academic achievement at nine years of age. *Dev Behav Pediatr.* 1989;10:32-37.
9. Buka SL, Lipsitt LP, Tsuang MT. Emotional and behavioral development of low-birthweight infants. In: Friedman SL, Sigman MD, eds. *The Psychological Development of Low-Birthweight Children: Annual Advances in Applied Developmental Psychology.* Norwood, NJ: Ablex Publishing Corp; 1992;6:187-214.
10. Bennett FC. The effectiveness of early intervention for infants at increased biologic risk. In: Guralnick MJ, Bennett FC, eds. *The Effectiveness of Early Intervention for At-Risk and Handicapped Children.* Orlando, Fla: Academic Press Inc; 1987: 79-112.
11. McCarton CM, Wallace IF, Bennett FC. Preventive interventions with low birth weight premature infants: an evaluation of their success. *Semin Perinatol.* 1995;19:330-340.
12. The Infant Health and Development Program. Enhancing the outcomes of low-birth-weight, premature infants: a multisite, randomized trial. *JAMA.* 1990;263:3035-3042.
13. Ramey CT, Bryant DM, Wasik BH, Sparling JJ, Fendt KH, LaVange LM. Infant Health and Development Program for low birth weight, premature infants: program elements, family participation, and child intelligence. *Pediatrics.* 1992;89: 454-465.
14. McCormick MC, McCarton C, Tonascia J, Brooks-Gunn J. Early educational intervention for very low birth weight infants: results from the Infant Health and Development Program. *J Pediatrics.* 1993;123:527-533.
15. Brooks-Gunn J, McCarton CM, Casey PH, et al. Early intervention in low-birth-weight premature infants: results through age 5 years from the Infant Health and Development Program. *JAMA.* 1994; 272:1257-1262.
16. McCormick MC, Brooks-Gunn J, Shapiro S, Benasich AA, Black G, Gross RT. Health care use among young children in day care: results in a randomized trial of early intervention. *JAMA.* 1991; 265:2212-2217.
17. Caldwell BM. The challenge of third-grade slump: special report: early childhood education. *Principal.* May 1987;66:10-14.
18. Constantine WL, Haynes CW, Spiker D, Kendall-Tackett K, Constantine NA. Recruitment and retention in a clinical trial for low birth weight, premature infants. *Dev Behav Pediatr.* 1993;14:1-7.
19. Kraemer HC, Fendt KH. Random assignment in clinical trials: issues in planning (Infant Health and Development Program). *J Clin Epidemiol.* 1990; 43:1157-1167.
20. McCullagh P, Nelder JA. *Generalized Linear Models.* 2nd ed. London, England: Chapman & Hall Ltd; 1989.
21. Scott DT, Bauer CR, Kraemer HC, Tyson J. A neonatal health index for preterm infants. *Pediatr Res.* 1989;25:263A. Abstract.
22. Searle SR, Speed FM, Milliken GA. Population marginal means in the linear model: an alternative to least squares means. *Am Stat.* 1980;34:216-221.
23. Rothman KJ. No adjustments are needed for multiple comparisons. *Epidemiology.* 1990;1:43-46.
24. The Consortium for Longitudinal Studies. *As the Twig Is Bent: Lasting Effects of Preschool Programs.* Hillsdale, NJ: Lawrence Erlbaum Associates Inc Publishers; 1983.
25. McKey R, Conndelli L, Gansin H, Barrett B, McConkey C, Plantz M. *The Impact of Head Start on Children, Families, and Communities: Final Report of the Head Start Evaluation, Synthesis and Utilization Project.* Washington, DC: CSR Inc; 1985.
26. Wasik BH, Ramey CT, Bryant DM, Sparling JJ. A longitudinal study of two early intervention strategies: Project CARE. *Child Dev.* 1990;61:1682-1696.
27. Lee VE, Brooks-Gunn J, Schnur E, Liaw F-R. Are Head Start effects sustained? a longitudinal follow-up comparison of disadvantaged children attending Head Start, no preschool, and other preschool programs. *Child Dev.* 1990;61:495-507.
28. Barnett WS. Long-term effects of early childhood programs on cognitive and school outcomes. *Future Child.* Winter 1995;5:25-50.
29. Ramey CT, Campbell FA. Poverty, early childhood education, and academic competence: the Abecedarian experiment. In: Huston AC, ed. *Children in Poverty: Child Development and Public Policy.* Cambridge, England: Cambridge University Press; 1991:190-221.
30. McCarton C, Bennett FC, Tonascia J, et al. Neurologic status at 36 months of age. In: Gross RT, Spiker D, eds. *Infant Health and Development Program.* Palo Alto, Calif: Stanford University Press. In press.
31. Brooks-Gunn J, Gross RT, Kraemer HC, Spiker D, Shapiro S. Enhancing the cognitive outcomes of low birth weight, premature infants: for whom is the intervention most effective? *Pediatrics.* 1992; 89:1209-1215.
32. Horacek HJ, Ramey CT, Campbell FA, Hoffman KP, Fletcher RH. Predicting school failure and assessing early intervention with high-risk children. *Am Acad Child Adolesc Psychiatry.* 1987; 26:758-763.
33. Byrd RS, Weitzman ML. Predictors of early grade retention among children in the United States. *Pediatrics.* 1994;93:481-487.
34. Klebanov PK, Brooks-Gunn J, McCormick MC. School achievement and failure in very low birth weight children. *J Dev Behav Pediatr.* 1994;15:248-256.
35. Weikart DP, Bond DT, McNeil JT. *The Ypsilanti Perry Preschool Project: Preschool Years and Longitudinal Results Through Fourth Grade.* Ypsilanti, Mich: High/Scope Educational Foundation; 1978:3. High/Scope Educational Research Foundation Monographs.
36. McCormick MC, Workman-Daniels K, Brooks-Gunn J. The behavioral and emotional well-being of school-age children with different birth weights. *Pediatrics.* 1996;97:18-25.
37. Fewell RR, Scott K. Cost analysis decisions for IHDP. In: Gross RT, Spiker D, eds. *Infant Health and Development Program.* Palo Alto, Calif: Stanford University Press. In press.
38. Beery KE. *Manual for the Developmental Test of Visual-Motor Integration.* 3rd revision. Cleveland, Ohio: Modern Curriculum Press; 1989.
39. Wechsler D. *Manual for the Wechsler Intelligence Scale for Children–Third Edition.* San Antonio, Tex: Psychological Corp; 1991.
40. Waber DP, Holmes JM. Assessing children's copy productions of the Rey-Osterrieth Complex Figure. *J Clin Exp Neuropsychol.* 1985;7:264-280.
41. Woodcock RW, Johnson MB. *Manual for the Woodcock-Johnson Tests of Achievement–Revised.* Allen, Tex: RCL Enterprises; 1990.
42. Elliott CD. *Manual for the Differential Ability Scales.* San Antonio, Tex: Psychological Corp; 1990.
43. Dunn LM, Dunn LM. *Manual for Peabody Picture Vocabulary Test–Revised.* Circle Pines, Minn: American Guidance Service; 1981.
44. Sheslow D, Adams W. *Manual for the Wide Range Assessment of Memory and Learning.* Wilmington, Del: Jastak; 1990.
45. Brown L, Hammill DD. *Manual for the Behavior Rating Profile.* 2nd ed. Austin, Tex: Pro-Ed; 1990.
46. Achenbach TM, Edelbrock C. *Manual for the Child Behavior Checklist and Revised Child Behavior Profile.* Burlington, Vt: Queen City Printers; 1983.
47. Landgraf JM, Ware JE Jr, Schor E, Davies AR, R-Roh K. Comparison of health status profiles in children with medical conditions: preliminary psychometric and clinical results for the Children's Health and Quality of Life Project. Presented at the 10th annual meeting of the Association for Health Services Research; June 1993; Washington, DC.