Developmental Outcomes of Late Preterm Infants From Infancy to Kindergarten

Prachi Shah, MD, MS, Niko Kaciroti, PhD, Blair Richards, MS, Wonjung Oh, PhD, Julie C. Lumeng, MD

OBJECTIVE: To compare developmental outcomes of late preterm infants (34–36 weeks’ gestation) with infants born at early term (37–38 weeks’ gestation) and term (39–41 weeks’ gestation), from infancy through kindergarten.

METHODS: Sample included 1000 late preterm, 1800 early term, and 3200 term infants ascertained from the Early Childhood Longitudinal Study, Birth Cohort. Direct assessments of development were performed at 9 and 24 months by using the Bayley Short Form–Research Edition T-scores and at preschool and kindergarten using the Early Childhood Longitudinal Study, Birth Cohort reading and mathematics scores. Maternal and infant characteristics were obtained from birth certificate data and parent questionnaires. After controlling for covariates, we compared mean developmental outcomes between late preterm and full-term groups in serial cross-sectional analyses at each timepoint using multilinear regression, with pairwise comparisons testing for group differences by gestational age categories.

RESULTS: With covariates controlled at all timepoints, at 9 months late preterm infants demonstrated less optimal developmental outcomes (T = 47.31) compared with infants born early term (T = 49.12) and term (T = 50.09) (P < .0001). This association was not seen at 24 months, (P = .66) but reemerged at preschool. Late preterm infants demonstrated less optimal scores in preschool reading (P = .0006), preschool mathematics (P = .0014), and kindergarten reading (P = .0007) compared with infants born at term gestation.

CONCLUSIONS: Although late preterm infants demonstrate comparable developmental outcomes to full-term infants (early term and full-term gestation) at 24 months, they demonstrate less optimal reading outcomes at preschool and kindergarten timepoints. Ongoing developmental surveillance for late preterm infants is warranted into preschool and kindergarten.

WHAT’S KNOWN ON THIS SUBJECT: Late preterm infants demonstrate less optimal developmental outcomes compared with early term and full-term children at kindergarten and early school age timepoints. Infant and preschool outcomes in late preterm infants remain underresearched.

WHAT THIS STUDY ADDS: Compared with full-term gestational groups, late preterm infants demonstrate similar developmental outcomes at 24 months but less optimal outcomes at preschool and kindergarten. Despite no developmental differences at 24 months, developmental surveillance for late preterms is warranted into preschool and kindergarten.
Late preterm infants (gestational age 34 weeks, 0 days to 36 weeks, 6 days) account for 75% of preterm births, translating to >400,000 late preterm births per year. It has been thought that late preterm infants were similar to full-term infants, with little risk for long-term morbidities, but increasing evidence suggests that late preterm infants have higher rates of learning problems than their full-term counterparts, often needing special education supports at school-age.

Recent studies have examined the possibility of a dose–response association between gestational age and early school performance by comparing early school outcomes of late preterm infants with outcomes of children born at early term (37–38 weeks) and term gestation (39–41 weeks). Compared with children born at early term and term gestation, children born in the late preterm period demonstrate poorer performance on tests of early school readiness, spatial abilities, and verbal reasoning in early childhood, poorer educational achievement at age 5, and poorer school performance at age 7. Despite these findings, little is known about the pattern of development in late preterm infants in the period before school entry. We have identified only 2 studies comparing developmental outcomes of late preterm infants with outcomes of children born early term and term infants in infancy, with disparate results, wherein 1 study found that late preterm infants demonstrated suboptimal development compared with early term and term infants at 8 and 18 months, and the other study found no differences at 2 years.

These contradictory findings suggest that more research is needed to examine the developmental outcomes of late preterm infants in infancy and to determine when developmental risks emerge between infancy and kindergarten to better inform early intervention and developmental monitoring strategies.

We used a nationally representative, population-based sample to examine developmental and early reading and math outcomes of late preterm infants, compared with infants born early term (37–38 weeks) and term (39–41 weeks) gestation from infancy through kindergarten. We hypothesized that compared with infants born at early term and term gestation, late preterm infants would manifest lower developmental scores in early infancy and that the developmental vulnerabilities would persist until kindergarten.

**METHODS**

**Study Design and Sample**

Data were drawn from the Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), a nationally representative, population-based longitudinal study sponsored by the US Department of Education’s National Center for Education Statistics (NCES) in the Institute for Education Science. The ECLS-B is based on a nationally representative probability sample of children born in the United States in 2001 (inclusive). Data were collected from >10,000 children and their parents at 9 months, with subsequent assessments at 24-month, preschool, and kindergarten timepoints. Because the NCES had an analytic interest in low birth weight infants, this population was oversampled. Use of data-weighting procedures adjusted for disproportionate sampling and survey nonresponse, thereby allowing weighted results to generate nationally representative estimates. Data collection consisted of home visits with parent interview and direct child assessments and included information on children’s cognitive, emotional, and physical development across multiple settings. The sample for our study excluded children with congenital and chromosomal abnormalities, included children born at 34 to 41 weeks inclusive, and used data from birth, 9-month, 24-month, preschool, and kindergarten timepoints. Our study was considered exempt by the Institutional Review Board because the research involved the use of a publicly available dataset in which the participants were deidentified, and data could not be linked to the participants.

**Measures**

**Outcomes**

Infant Development. Developmental outcomes at 9 and 24 months were ascertained with the Bayley Short Form—Research Edition (BSF-R) Mental T scores. The BSF-R, which was formulated from the Bayley Scales of Infant Development, Second Edition (BSID-II), assessed children’s performance on tasks requiring memory, problem solving, and language skills and was administered at home by trained NCES staff. The BSF-R estimates the number of items a child would have gotten correct on the full BSID-II through the use of item response theory (IRT) modeling and demonstrates a reliability coefficient of 0.80 compared with the full BSID-II. The BSF-R Mental T scores were norm-referenced by age to the ECLS-B population, adjusted for prematurity, and scaled to have a mean = 50, SD = 10. Preschool and Kindergarten Assessments. Children were assessed at preschool and kindergarten via specialized reading and math assessments developed for the ECLS-B. The reading assessment was formulated from existing instruments including the Peabody Picture Vocabulary Test, Third Edition and Preschool Comprehensive Test of Phonological and Print Processing and measured basic reading skills (letter and word recognition, understanding letter–sound relationships, phonological
awareness, sight word recognition, and understanding words in the context of simple sentences). The reliability of the early reading assessment is described by the IRT reliability coefficient, reported as 0.84 at preschool and 0.92 at kindergarten. Scores provide ability estimates in a particular domain and were reported as normally distributed θ scores, which demonstrated a range of $-2.47$ to $2.60$ (mean $= -0.49$, SD $= 0.74$) at preschool and $-2.11$ to $3.09$ (mean $= 0.33$, SD $= 0.86$) at kindergarten.$^{19}$ The ECLS-B mathematics assessment incorporated questions to test the following content areas: number sense, geometry, counting numerical operations, and pattern recognition. The IRT reliability coefficient for the early mathematics assessment was 0.89 at preschool and 0.92 at kindergarten. The mathematics θ scores demonstrated a range of $-2.84$ to $2.38$ (mean $= -0.47$, SD $= 0.78$) at preschool and $-2.42$ to $3.12$ (mean $= 0.38$, SD $= 0.80$) at kindergarten.$^{19}$

**Predictor**

Gestational age was ascertained from birth certificate data from the ECLS-B restricted use data set,$^{20}$ and infants were categorized as being late preterm (gestational age 34–36 weeks),$^{1}$ early term (37–38 weeks), or full term (39–41 weeks) per conventional gestational age categories.$^{21}$

**Covariates**

Maternal and infant characteristics were included as covariates. The following maternal characteristics were ascertained from the restricted ECLS-B birth certificate data: maternal age, race or ethnicity, marital status (married or unmarried), history of prenatal smoking, duration of breastfeeding, and plurality (singleton, twin, multiple gestation). Also included were measures of maternal education (less than high school, high school graduate, some college, college graduate, more than college) and poverty ($<$185% federal poverty limit, $\geq$185% federal poverty line), which were incorporated into a single composite measure of household socioeconomic status created by ECLS-B at 9 months.$^{15}$ Parenting characteristics from 9 months to preschool were included from coded observations of parent–child interactions. Nine-month parenting was assessed via the 50-item parent scale of the Nursing Child Assessment Teaching Scale.$^{22}$ The Parental Responsiveness subscale ($\alpha = 0.68$) reflects parental sensitivity to child cues (range 0–50).$^{15}$ Parenting at 24 months and preschool was ascertained from coded interactions from the Two-Bags Task, a 10-minute semistructured parent–child activity, which assessed parental sensitivity, parental cognitive stimulation, and parental emotional support on a Likert scale (range 1–7). Higher scores indicated more positive parenting behaviors.$^{18,23}$

Infant characteristics and neonatal risks, also ascertained from birth certificate data, included infant gender; birth weight; fetal growth, characterized as small for gestational age ($<$10%), appropriate for gestational age (10%–90%), and large for gestational age ($>$90%), defined per national standards$^{24}$; 5-minute Apgar scores (dichotomized as $>$7 vs $\leq$7); and days of hospitalization (0–1 days, 2–7 days, 8–14 days, 15–30 days, $>$30 days). Enrollment in early intervention, ascertained from 9-month and 24-month parent questionnaires, and age and month of preschool and kindergarten assessment were also included as covariates.$^{19}$

**Statistical Analyses**

Maternal and child characteristics were examined via descriptive statistics to identify outliers. Because of the presence of outliers regarding age of developmental assessment at the different timepoints, the sample was restricted to infants assessed between 8 and 11 months (9-month timepoint), 23 to 25 months (24-month timepoint), 42 to 59 months (preschool timepoint), and 60 to 72 months (kindergarten timepoint). Child characteristics were compared by gestational age group via 1-way analysis of variance for continuous variables and logistic regression for binomial and categorical variables.

In a series of cross-sectional analyses at each timepoint, multilinear regression (performed with the SURVEYREG procedure in SAS [SAS Institute, Inc, Cary, NC]) was used to examine the association between mean developmental outcomes at 9 and 24 months and mean reading and math outcomes at preschool and kindergarten timepoints with gestational age categories (late preterm, early term, or term). Initial models examined the contribution of gestational age groups and covariates on outcomes at 9 months, 24 months, preschool, and kindergarten timepoints. The covariates with a $P$ value $<$0.05 (maternal age, maternal race or ethnicity, socioeconomic status at 9 months, parenting, infant gender, birth weight, and receipt of early intervention services) were included in the final model for adjusted analyses. Analyses at preschool and kindergarten also included age at assessment and month of school. Post hoc tests examined pairwise differences between gestational age categories at a significance level of $P < .05$.

All analyses were conducted in SAS 9.4.$^{25}$ Because of the complex sample design, sample weights and the Jackknife method were used to account for stratification, clustering, and unit nonresponse. Results are reported for the weighted analyses, and in accord with the NCES requirements for ECLS-B data usage, reported numbers were rounded to the nearest 50.
RESULTS

Characteristics of the Sample

The unweighted sample included a total of 6000 infants: 1000 late preterm (gestational age 34–36 weeks), 1800 early term (37–38 weeks), and 3200 term (39–41 weeks) infants. The mean age for mothers was 27.5 years (SD = 5.0), most (68.9%) were married, most (79.2%) had a high school education or greater, and approximately half of families (53%) were at or above the poverty threshold, defined as income >185% of the federal poverty line. Maternal characteristics are shown in Table 1. Infant characteristics did not vary by gestational age group, with the exception of lower birth weight, lower 5-minute Apgar scores, more days of hospitalization, and higher percentage of early intervention services at 9 months for infants born late preterm. Child characteristics are shown in Table 2.

Group Differences: Developmental Outcomes by Gestational Age Categories

9-Month Cognitive Outcomes

Adjusted analysis showed that infants born late preterm demonstrated lower cognitive scores compared with other gestational age groups as measured by the BSF-R Mental T scores at 9 months (F = 13.97; P < .0001) (Table 3). Post hoc pairwise comparisons revealed that late preterm infants demonstrated lower cognitive scores compared with early term (P = .0006) and term infants (P < .0001). This association was also seen in infants with higher gestational age: Early term infants demonstrated lower cognitive outcomes compared with term infants (P = .02) (Fig 1A).

24-Month Cognitive Outcomes

Contrary to 9-month developmental outcomes, after we controlled for covariates, there were no differences in 24-month cognitive outcomes between late preterm infants and full-term infants (F = 0.41; P = .66) (Table 3, Fig 1B).

Preschool Outcomes

After we controlled for covariates, infants born late preterm demonstrated lower reading scores at preschool compared with other gestational age groups as measured by the ECLS-B Preschool Reading Assessment θ scores (F = 9.44; P = .0002) (Table 3). Post hoc pairwise comparisons revealed that late preterm infants demonstrated lower reading scores compared with full-term infants (P = .0006) but not early term infants (P = .14) (Fig 1C). There were also differences in preschool math outcomes among gestational age groups (F = 7.07; P = .0014) (Table 3), with late preterm infants demonstrating lower math scores compared with term (P = .0014) but not early term (P = .07) infants (Fig 1D).

Kindergarten Outcomes

Similar to preschool reading outcomes, after we controlled for covariates, children born late preterm demonstrated lower reading scores at kindergarten compared with other gestational age groups (F = 6.33; P = .003) (Table 3). Post hoc pairwise comparisons revealed that late preterm infants demonstrated

<table>
<thead>
<tr>
<th>TABLE 1 Sample Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal Characteristics</td>
</tr>
<tr>
<td>Age (y)</td>
</tr>
<tr>
<td>Race or ethnicity, %</td>
</tr>
<tr>
<td>White, non-Hispanic</td>
</tr>
<tr>
<td>Black, non-Hispanic</td>
</tr>
<tr>
<td>Hispanic</td>
</tr>
<tr>
<td>Asian</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>Marital status, %</td>
</tr>
<tr>
<td>Married</td>
</tr>
<tr>
<td>Unmarried</td>
</tr>
<tr>
<td>History of prenatal smoking, %</td>
</tr>
<tr>
<td>No</td>
</tr>
<tr>
<td>Yes</td>
</tr>
<tr>
<td>Duration of breastfeeding, %</td>
</tr>
<tr>
<td>&lt;1 mo</td>
</tr>
<tr>
<td>2–3 mo</td>
</tr>
<tr>
<td>4–6 mo</td>
</tr>
<tr>
<td>&gt;6 mo</td>
</tr>
<tr>
<td>Plurality, %</td>
</tr>
<tr>
<td>Singleton</td>
</tr>
<tr>
<td>Twin</td>
</tr>
<tr>
<td>Multiple (≥3)</td>
</tr>
<tr>
<td>Socioeconomic indicators calculated from measures of education and income at 9 mo</td>
</tr>
<tr>
<td>Maternal education, %</td>
</tr>
<tr>
<td>Less than high school</td>
</tr>
<tr>
<td>High school degree</td>
</tr>
<tr>
<td>Some college</td>
</tr>
<tr>
<td>Bachelor’s degree</td>
</tr>
<tr>
<td>Graduate school</td>
</tr>
<tr>
<td>Below poverty threshold (&lt;185% federal poverty line), %</td>
</tr>
<tr>
<td>No</td>
</tr>
<tr>
<td>Yes</td>
</tr>
<tr>
<td>At or above poverty threshold (≥185% federal poverty line), %</td>
</tr>
</tbody>
</table>

Parenting behavior, 9 mo

Parental responsiveness 34.5, 7.1
Parenting behavior, 24 mo

Parental sensitivity 4.78, 1.2
Parental cognitive stimulation 4.13, 1.5
Parenting behavior, preschool

Parental emotional support 4.5, 1.2
Parental cognitive stimulation 4.2, 1.0
lower kindergarten reading scores compared with early term (P = .02) and full-term (P = .0007) infants (Fig 1E). There were no differences in kindergarten math outcomes among the 3 gestational age groups (F = 0.22; P = .81) (Table 3, Fig 1F).

**DISCUSSION**

This is the first longitudinal study to compare developmental outcomes of late preterm infants with term-born gestational groups (early term and term) from infancy to kindergarten in a nationally representative, population-based sample from the United States. We found that the pattern of suboptimal development of late preterm infants appears to vary from 9 months through kindergarten. Similar to other longitudinal studies in infancy comparing outcomes of late preterm with early term and term infants, we found that late preterm infants demonstrate lower developmental outcomes at 9 months compared with term-born infants (early term and term groups). This association was not seen at 24 months, which is consistent with results from some studies but differs from other studies. At preschool and kindergarten timepoints, similar to previous studies of late preterm infants at preschool and early school age, we found that late preterm infants demonstrate less optimal early reading skills in preschool and kindergarten and less optimal...
FIGURE 1
A. Mean cognitive scores at 9 months by gestational age categories. B. Mean cognitive scores at 24 months by gestational age categories. C. Mean preschool reading scores by gestational age categories. D. Mean preschool math scores by gestational age categories. E. Mean kindergarten reading scores by gestational age categories. F. Mean kindergarten math scores by gestational age categories. Superscripts a, b, and c that differ from one another denote significant differences (P < .05) between late preterm, early term, and term infants at each timepoint.
mathematics skills at preschool compared with full-term groups.

Regarding our 24-month findings, although previous research with the ECLS-B has demonstrated lower developmental outcomes of late preterm infants compared with full-term infants at 24 months, raising the possibility that the subtle developmental delays of late preterm infants in infancy may become less apparent in the toddler period, which may have implications for developmental surveillance in primary care.

Despite no developmental differences observed at 24 months, as in previous research, we found that late preterm infants demonstrate less optimal developmental outcomes at preschool. This suggests that the preschool-age health supervision visit may be an especially important visit for pediatric providers to monitor for emerging developmental delays in children who were born late preterm and to facilitate appropriate management for developmental issues. Children who were born late preterm who demonstrate delays on pediatric screening should be referred for a comprehensive developmental assessment to provide a diagnostic evaluation for developmental delay and to facilitate referrals for appropriate interventions. In addition, these children should be referred for a school-based psychoeducational evaluation to determine the need for school-based supports and services.

There are some potential explanations for the pattern of development observed in late preterm infants from 9 months to kindergarten. Compared with infants born full term, the late preterm brain manifests neurodevelopmental immaturity, characterized by lower brain volume and less differentiated patterns of myelination and neural connectivity. These structural differences in the brains of children born late preterm, including lower gray matter volume present in infancy and school age, may be associated with the suboptimal development in early reading and math skills we identified, similar to that which has been observed in children born full term.

The lack of developmental differences in children born late preterm observed at 24 months and the reemergence of suboptimal development at preschool may also be related to limitations in the instruments used at the 24-month timepoint and to subtle differences in brain development that become apparent with advancing age. The mental scale of the BSID-II (after which the BSF-R was formulated) was designed to assess sensory perceptual abilities, early verbal communication, and early number concepts but may in fact be a partial measure of gross motor abilities and sensorimotor functioning and may not be an accurate assessment of cognitive abilities in infancy. In addition, the neurocognitive deficits associated with late preterm birth, including deficits in early reading and math skills, may not be reliably detected in the toddler period.

Our study had several strengths and limitations. The study includes a nationally representative sample, the results of which are generalizable to the population, and assessments of child development from infancy through kindergarten. In addition, in contrast to other studies that have used parent report questionnaires to assess development, the ECLS-B administered direct child assessments to measure development, thus minimizing biases that can be associated with parent report measures. One of the limitations of the study relates to different assessments of child development at infancy and the preschool and kindergarten timepoints. Because the BSF-R assessed different developmental constructs in infancy compared with the measures of early reading and math used at the preschool and kindergarten timepoints, it was deemed that the measures were not equivalent, and a longitudinal...
approach to the data was not possible. As a result, we assessed developmental outcomes in a series of cross-sectional analyses rather than following the same child from infancy through kindergarten longitudinally. In addition, the ECLS-B did not include a measure of IQ at preschool and kindergarten, so the measures of early learning are limited to the early reading and math domains and cannot be generalized to an overall measure of intellectual aptitude. Furthermore, there is no information in the ECLS-B about illnesses and hospitalizations in the neonatal period, or the specific details regarding the medical morbidities associated with late preterm birth that may be associated with suboptimal developmental outcomes.

CONCLUSIONS

We found small but meaningful mean differences in developmental outcomes between late preterm infants and full-term groups, which, when extrapolated to a population level, can have potentially significant public health implications for long-term outcomes. Developmental differences in late preterm infants emerged between 24 months and preschool, suggesting that late preterm infants may benefit from closer developmental monitoring, targeted assessments, and interventions before school entry.

REFERENCES

13. Schonhaut L, Armijo I, Pérez M. Gestational age and developmental risk in moderately and late preterm...
and early term infants. Pediatrics. 2015;135(4). Available at: www.pediatrics.org/cgi/content/full/135/4/e835


17. Bayley N. Bayley Scales of Infant and Toddler Development. 2nd ed. San Antonio, TX: Harcourt Assessment; 1993


24. Engle WA. A recommendation for the definition of “late preterm” (near-term) and the birth weight-gestational age classification system. Semin Perinatol. 2006;30(1):2–7


28. Woynschele MA, McCormick MC, Smith VC. Late preterm infants have worse 24-month neurodevelopmental outcomes than term infants. Pediatrics. 2011;127(3). Available at: www.pediatrics.org/cgi/content/full/127/3/e622


Updated Information & Services
including high resolution figures, can be found at:
http://pediatrics.aappublications.org/content/138/2/e20153496

References
This article cites 40 articles, 12 of which you can access for free at:
http://pediatrics.aappublications.org/content/138/2/e20153496#BIBL

Subspecialty Collections
This article, along with others on similar topics, appears in the following collection(s):
- Developmental/Behavioral Pediatrics
  http://www.aappublications.org/cgi/collection/development:behavioral_issues_sub
- Cognition/Language/Learning Disorders
  http://www.aappublications.org/cgi/collection/cognition:language:learning_disorders_sub
- Fetus/Newborn Infant
  http://www.aappublications.org/cgi/collection/fetus:newborn_infant_sub
- Neonatology
  http://www.aappublications.org/cgi/collection/neonatology_sub

Permissions & Licensing
Information about reproducing this article in parts (figures, tables) or in its entirety can be found online at:
http://www.aappublications.org/site/misc/Permissions.xhtml

Reprints
Information about ordering reprints can be found online:
http://www.aappublications.org/site/misc/reprints.xhtml