Published in final edited form as:

J Pediatr. 2012 August ; 161(2): 222–8.e3. doi:10.1016/j.jpeds.2012.01.057.

Are Outcomes of Extremely Preterm Infants Improving? Impact of Bayley Assessment on Outcomes

Betty R. Vohr, MD¹, Bonnie E. Stephens, MD¹, Rosemary D. Higgins, MD², Carla M. Bann, PhD³, Susan R. Hintz, MD, MS Epi⁴, Abhik Das, PhD⁵, Jamie E. Newman, PhD, MPH³, Myriam Peralta-Carcelen, MD, MPH⁶, Kimberly Yolton, PhD⁷, Anna M. Dusick, MD, FAAP⁸, Patricia W. Evans, MD⁹, Ricki F. Goldstein, MD¹⁰, Richard A. Ehrenkranz, MD¹¹, Athina Pappas, MD¹², Ira Adams-Chapman, MD¹³, Deanne E. Wilson-Costello, MD¹⁴, Charles R. Bauer, MD¹⁵, Anna Bodnar, MD¹⁶, Roy J. Heyne, MD¹⁷, Yvonne E. Vaucher, MD, MPH¹⁸, Robert G. Dillard, MD¹⁹, Michael J. Acarregui, MD²⁰, Elisabeth C. McGowan, MD²¹, Gary J. Myers, MD²², and Janell Fuller, MD²³ for the *Eunice Kennedy Shriver* National Institute of Child Health and Network Human Development Neonatal Research^{*}

¹Department of Pediatrics, Women & Infants' Hospital, Brown University, Providence, RI

²*Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD

³Statistics and Epidemiology Unit, RTI International, Research Triangle Park, NC

⁴Department of Pediatrics, Stanford University School of Medicine, Palo Alto, CA

⁵Statistics and Epidemiology Unit, RTI International, Rockville, MD

⁶Division of Neonatology, University of Alabama at Birmingham, Birmingham, AL

⁷Department of Pediatrics, University of Cincinnati, Cincinnati, OH

⁸Department of Pediatrics, Indiana University School of Medicine, Indianapolis, IN

⁹Department of Pediatrics, University of Texas Medical School at Houston, Houston, TX

¹⁰Department of Pediatrics, Duke University, Durham, NC

¹¹Department of Pediatrics, Yale University School of Medicine, New Haven, CT

¹²Department of Pediatrics, Wayne State University, Detroit, MI

¹³Department of Pediatrics, Emory University School of Medicine and Children's Healthcare of Atlanta, Atlanta, GA

¹⁴Department of Pediatrics, Rainbow Babies & Children's Hospital, Case Western Reserve University, Cleveland, OH

¹⁵University of Miami Miller School of Medicine, Miami, FL

¹⁶Department of Pediatrics, Division of Neonatology, University of Utah School of Medicine, Salt Lake City, UT

Copyright © 2012 Mosby Inc. All rights reserved.

Reprint requests: Betty R. Vohr, MD, Women & Infant's Hospital, 101 Dudley Street, Providence, RI 02905. bvohr@wihri.org.

^{*}A list of members of the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development Neonatal Research Network is available at www.jpeds.com (Appendix).

The authors declare no conflicts of interest.

We are indebted to our medical and nursing colleagues and the infants and their parents who agreed to participate in this study.

¹⁷Department of Pediatrics, University of Texas Southwestern Medical Center, Dallas, TX

¹⁸Division of Neonatology, University of California San Diego, La Jolla, CA

¹⁹Wake Forest University School of Medicine, Winston-Salem, NC

²⁰Department of Pediatrics, University of Iowa, Iowa City, IA

²¹Department of Pediatrics, Division of Newborn Medicine, Floating Hospital for Children, Tufts Medical Center, Boston, MA

²²University of Rochester School of Medicine and Dentistry, Rochester, NY

²³University of New Mexico Health Sciences Center, Albuquerque, NM

Abstract

Objectives—To compare 18- to 22-month cognitive scores and neurodevelopmental impairment (NDI) in 2 time periods using the National Institute of Child Health and Human Development's Neonatal Research Network assessment of extremely low birth weight infants with the Bayley Scales of Infant Development, Second Edition (Bayley II) in 2006–2007 (period 1) and using the Bayley Scales of Infant and Toddler Development, Third Edition (Bayley III), with separate cognitive and language scores, in 2008–2011 (period 2).

Study design—Scores were compared with bivariate analysis, and regression analyses were run to identify differences in NDI rates.

Results—Mean Bayley III cognitive scores were 11 points higher than mean Bayley II cognitive scores. The NDI rate was reduced by 70% (from 43% in period 1 to 13% in period 2; $P < .0001$). Multivariate analyses revealed that Bayley III contributed to a decreased risk of NDI by 5 definitions: cognitive score <70 and <85 , cognitive or language score <70 ; cognitive or motor score <70 , and cognitive, language, or motor score <70 ($P < .001$).

Conclusion—Whether the Bayley III is overestimating cognitive performance or whether it is a more valid assessment of emerging cognitive skills than the Bayley II is uncertain. Because the Bayley III identifies significantly fewer children with disability, it is recommended that all extremely low birth weight infants be offered early intervention services at the time of discharge from the neonatal intensive care unit, and that Bayley scores be interpreted with caution.

Extremely preterm (PT) infants are at increased risk for cognitive impairments. The most common individual severe impairment identified using the standard composite outcome of neurodevelopmental impairment (NDI) for extremely low birth weight (ELBW) infants at age 18 and 30 months is cognitive impairment, defined as a score >2 SD below the mean (<70).^{1–5} Early developmental/cognitive function of PT children has traditionally been assessed using the Bayley Scales of Infant Development.^{6,7} Updates of developmental tests such as the Bayley Scales are standard because of steadily increasing scores over time (3 points per year), a phenomenon known as the Flynn effect.⁸ There is some evidence of this effect. A report from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development Neonatal Research Network (NRN) on Bayley scores of infants born at gestational age (GA) <32 weeks demonstrated increasing scores between 1993 and 1998.³ A limitation of the Bayley Scales of Infant Development, Second Edition (Bayley II) is its inclusion of only 2 developmental scores, the Mental Developmental Index (MDI), a composite of cognitive, receptive language, and expressive language tasks, and the Psychomotor Developmental Index (PDI), a composite of fine and gross motor skills. This limitation contributed to the development of the Bayley Scales of Infant and Toddler Development, Third Edition (Bayley III),⁷ which contains 3 individual developmental scores: a cognitive composite score (Cog), a language composite score (with receptive and

expressive subscores), and a motor composite score (with gross and fine motor sub-scores), in addition to social-emotional and adaptive behavior domains. This allows the examiner to identify a deficit in a specific developmental domain as well as relative strengths and challenges. It was expected that separating the language scores from the Cog score might result in higher cognitive scores.

The National Institute of Child Health and Human Development's NRN has been reporting on the outcomes of ELBW infants (birth weight <1000 g) with the Bayley II since 1993. The NRN converted to the Bayley III in January 2008. This change in protocol provides an opportunity to compare rates of low (<2 SD) Bayley II MDI scores and low Bayley III Cog scores.

The objectives of the present study were (1) to compare cognitive outcomes in infants born at a GA of 26^{6/7} weeks and a birth weight of 1000 g and assessed during period 1 (2006–2007) using the Bayley II and in infants assessed during period 2 (2008–2011) using the Bayley III; (2) to explore the utility of a threshold of <70 versus <85 on the Bayley III to reflect impairment; and (3) to compare the association of major neonatal morbidities with NDI rates between period 1 and period 2. We hypothesized that cognitive scores would be higher and NDI rates lower in period 2, and that the Bayley III would have comparable performance to the Bayley II in identifying an increased risk of NDI in infants with major neonatal morbidities.

Methods

This was a retrospective cohort study of ELBW infants with a birth weight of 401–1000 g and born at a GA of <27 weeks who were admitted to one of 20 neonatal intensive care units in the NRN and underwent comprehensive neurologic and developmental assessments at 18–22 months corrected age (CA) during calendar years 2006–2011. The infants were categorized into 2 groups, those assessed during period 1 using the Bayley II (n = 1012) and those assessed during period 2 using the Bayley III (n = 1616). In addition, to address potential differences in factors other than Bayley scores that could affect the rate of NDI, we used propensity score matching with SAS software (SAS Institute, Cary, North Carolina)⁹ to select sub-samples of children from the 2 time periods who were similar based on the following characteristics: maternal age, education, race, public insurance coverage, antenatal steroid use, cesarean delivery, birth weight, GA, multiple birth, child's sex, intraventricular hemorrhage (IVH)/periventricular leukomalacia (PVL), necrotizing enterocolitis (NEC), sepsis, broncho-pulmonary dysplasia (BPD), retinopathy of prematurity (ROP), postnatal steroid use, number of days on ventilation, duration of hospital stay, adjusted age at follow-up assessment, cerebral palsy (CP), vision impairment, hearing impairment, and research center. These matched samples included 922 children from each time period for a total sample of 1844.

All maternal and neonatal data, treatments, and clinical outcomes are prospectively collected in the NRN generic database. Centers participating in the NRN received local Institutional Review Board approval for data collection. Trained research coordinators obtained data based on the definitions listed in the NRN Manual of Operations. The effects of 4 major neonatal risk factors (brain injury, defined as IVH grade 3–4 or cystic PVL; BPD; NEC; and sepsis) on NDI rates were examined. BPD was defined as treatment with supplemental oxygen at 36 weeks gestation. IVH grade 3–4 was based on the Papile classification scheme,¹⁰ and represents the maximum grade noted on cranial ultrasound before discharge. Cystic PVL was based on the reported head ultrasound examination. Sepsis was combined early-onset sepsis (<72 hours) and/or late-onset sepsis (>72 hours) based on positive blood cultures, and NEC was defined as modified Bell classification stage II A or higher.

The follow-up evaluation included neurologic, hearing, vision, and developmental assessment. The NRN has protocols in place for annual training of all examiners to ensure reliability of all study assessments.¹¹ The neurologic examination, based on the Amiel-Tison assessment scheme,¹² includes an evaluation of tone, strength, reflexes, angles, and posture. CP was defined as a nonprogressive central nervous system disorder characterized by abnormal muscle tone in at least one extremity and abnormal control of movement and posture. Hearing status was obtained by parental history, and hearing impairment was confirmed by audiologic testing. A history of eye examinations and procedures since initial discharge was obtained, and a standard eye examination was completed. Blindness was defined as bilateral corrected vision of worse than 20/200.

Development was assessed with the Bayley II in period 1 and the Bayley III in period 2. Both tests have a mean score of 100 ± 15 . A score <70 (>2 SD below the mean) indicates significant delay, and a score of <85 (>1 SD below the mean) indicates at least mild to moderate delay. The Bayley III manual documents the establishment of content, criterion-related and construct validity. A standardization sample of 1700 children for the Bayley III was reported to be representative of the 2000 US Census population survey data for parent education, ethnicity, and geographic location. Children with a history of PT birth or cognitive, physical, or behavioral issues composed 10% of the sample. The percentage of PT or low birth weight infants in the standardization sample of 1700 was limited, however, and included 85 infants (5%) aged <37 weeks tested between 2 months and 42 months CA, or 2.3% during each time period. Although the Bayley III encompasses cognitive, language, motor, social-emotional, and adaptive behavior dimensions, the NRN made the decision to not use the social-emotional and adaptive behavior interviews because of time constraints. Children who could not be assessed due to severe developmental delay were assigned Bayley II MDI and PDI scores of 49 (>3 SD below the mean and the lowest score) and a Bayley III Cog score of 54 (>3 SD below the mean and the lowest score).

The NRN has used a standard definition for the composite outcome NDI. The initial definition of NDI (1993–2006) included the presence of any of the following: Bayley MDI <70 , Bayley PDI <70 , moderate to severe CP, bilateral blindness, or bilateral hearing loss requiring amplification. Table I (available at www.jpeds.com) presents these definitions. The primary difference between the 2 time periods was the Bayley assessment. During period 1, the developmental assessment included both MDI and PDI scores. During period 2, the Cog score was included, but the language score was not. Second, the Bayley III motor composite scale initially was not administered during period 2 because of time constraints, and the Palisano Gross Motor Function Classification System (GMFCS), which assesses gross motor function but not fine motor function, was included. In January 2010, the assessment was expanded to include the Bayley III motor composite score; 596 of the 1616 children in period 2 underwent the motor composite assessment. Children with amplification for hearing loss who were able to follow directions and communicate with the examiner were not classified as impaired during period 2 ($n = 29$). For this report, however, we used the period 1 definition of bilateral hearing loss for both cohorts in our analyses.

Statistical Analyses

Bivariate analyses were conducted to compare children in the 2 study periods, with the ² test used to compare categorical variables and ANOVA used to compare continuous variables. Similar analyses were conducted to compare development, motor, and neurosensory outcomes across the 2 study periods. The odds of NDI in each study period was calculated using the NRN's standard definitions and compared between groups using logistic regression models, controlling for maternal age, maternal education, race, public insurance; antenatal steroid use, cesarean delivery, birth weight, GA, multiple birth, sex, IVH grade 3–4/cystic PVL, NEC, sepsis, BPD, ROP, postnatal steroid use, ventilation days,

length of hospital stay, adjusted age at follow-up, and NRN center. To further explore possible sources of discrepancies in NDI rates between the 2 study periods, comparisons were repeated to determine whether modifying the threshold for the Bayley III and/or adding in the language composite would produce NDI rates comparable to those with the Bayley II. Finally, we investigated the ability of the Bayley II and III to detect NDI among children with major neonatal morbidities, including BPD, NEC, IVH grade 3–4/cystic PVL, and culture-positive early- or late-onset sepsis, as a test of construct validity. Logistic regression analyses were performed to evaluate the independent associations between major neonatal morbidities and NDI, controlling for the aforementioned covariates.

Results

Maternal and infant characteristics are presented in Table II (available at www.jpeds.com). There were slight differences between the study periods for the samples before matching. Study period 2 included slightly more educated mothers, 3% more children who received antenatal steroids, and 5% fewer children with ROP or sepsis. In unadjusted analyses, Bayley III Cog scores were significantly higher than Bayley II MDI scores, and fewer children were classified as impaired, regardless of whether a threshold of 70 or 85 was used (Table III). The Bayley III motor composite score was 6 points higher than the Bayley II PDI. Significantly fewer children in period 2 had moderate/severe CP and bilateral hearing loss in period 2, but there were no between-group differences in the rate of vision impairment. The combined rate of CP, blindness, and hearing loss decreased from 9% in period 1 to 7% in period 2.

With the use of propensity score matching for the sample selection, there were no significant differences between the matched samples in maternal and infant characteristics (Table II) or CA at follow-up, CP, GMFCS, and vision and hearing impairment (Table III). However, even after matching the samples on these characteristics, significant differences in Bayley scores remained (Table III).

Using the standard NRN definitions, 43% of children in study period 1 were classified as NDI, compared with 13% in period 2 (Table IV). When the threshold for the Cog score was raised from 70 to 80 for the period 2 definition, the percentage of NDI increased from 13% to 22%. Maintaining a threshold of 70 and combining the children with a language score or Cog score <70 resulted in a 23% NDI rate in period 2; the combination of Cog or language score with a threshold of 80 resulted in an NDI rate of 47%. The differences in NDI rate between the 2 study periods were similar in the matched samples (Table IV). Applying the NRN standard definitions, the NDI rate was 43% in period 1 and 14% in period 2.

Another factor that needs to be addressed is the NRN's use of a GMFCS score ≥ 2 as an independent indicator of motor impairment for a limited time during period 2. The rate was 7%-8%. During the time when both the GMFCS and Bayley III motor composite were administered, 7% of children had a GMFCS score ≥ 2 and 14% had a Bayley III motor composite score <70. This finding demonstrates the importance of administering the complete Bayley motor test. Despite the use of the complete Bayley III motor assessment, the percentage that met the NRN criteria for NDI was still only half of that with a PDI <70 and reinforces the fact that the GMFCS is meant to identify gross motor function only. This was confirmed in our analyses of matched cohorts, with 8% with a GMFCS score ≥ 2 , 15% with a Bayley III motor composite score <70, and 25% with a Bayley II PDI <70. This finding suggests that the Bayley III motor score is 40% less likely to identify disability than the Bayley II PDI and 47% more likely to identify disability than the GMFCS.

Table IV also presents the unadjusted and adjusted ORs of being diagnosed with NDI during period 2 compared with period 1. After controlling for demographic data, medical risk factors, and research center, there were significant differences in NDI rates across the 2 study periods. For all but 3 of the definitions (Cog/language <80 and <85; Cog/motor/language <80), children in period 2 had significantly lower odds of NDI. These findings are consistent with data from the analysis of matched samples.

We performed regression analyses to assess the relationship between neonatal morbidities and the standard NRN definition of NDI during the 2 time periods. During both study periods and for both the original and matched sample comparisons, in unadjusted analyses a threshold of 70 clearly differentiated the increased risk of NDI for all 4 morbidities. NDI rates for any of the 4 morbidities were significantly lower during period 2. For example, 54% of infants with BPD had NDI during period 1, compared with 17% during period 2. In addition, 30% of infants without BPD had NDI during period 1, compared with only 8% during period 2. In adjusted analyses, only infants with IVH grade 3–4 or cystic PVL were considered at significantly increased risk for NDI during both period 1 (OR, 2.36; 95% CI, 1.57–3.55) and period 2 (OR, 5.06; 95% CI, 3.04–8.43) (Table V).

Discussion

Our findings support our hypothesis that Bayley III Cog scores are higher than Bayley II MDI scores. This is not surprising, given that the Bayley III manual states that the Cog scores are 7 points higher than the Bayley II MDI; in our cohort, they were 11 points higher. This difference of 4 points is small but represents one-quarter SD and thus is potentially clinically significant.

The reasons for this difference are not clear. The Bayley III separation of language and Cog scores was developed to minimize effects of language delay on the cognitive assessment. Intelligence tests invariably separate verbal IQ from performance IQ. Thus, this separation may result in a cognitive index that is more closely associated with school-age cognitive performance and facilitates the identification of children needing speech therapy. Evidence will be reviewed to examine the validity of the Bayley III compared with the Bayley II.

Outcome studies with the Bayley II consistently identified high rates of cognitive impairments in this age group.^{1–5,11} In contrast, studies of PT children at school age are more likely to report mean Cog scores within 1 SD of the normed mean.^{13–15} Evaluating the stability of developmental outcomes between age 18–24 months and school age is limited by intervening events, changes in the classification of cognitive disability,¹⁶ and methodological issues, including the fact that intelligence tests are more psychometrically sophisticated.^{17–19} Hack et al¹³ examined the predictive validity of the Bayley II in ELBW infants and found a drop in the rate of cognitive impairment from 39% at 20 months to 16% at 8 years. Roberts et al¹⁶ compared very PT children at age 2 years using the Bayley II MDI and at age 8 years using the Wechsler Intelligence Scale for Children IV IQ²⁰ and identified poor stability of disability classification. Those studies suggest that the Bayley II may overidentify moderate to severe cognitive impairment.

The opposing argument is that the lack of association between early Bayley II scores and IQ scores at older ages is secondary to the fact that PT children have real cognitive deficits at early ages that recover over time. If this is the case, then the Bayley III may be underidentifying transient impairments that would benefit from early intervention. Anderson et al²¹ reported that the Bayley III underestimates the delay in both PT and term children at age 2 years. Their mean Bayley III Cog scores of both PT infants (96.9 ± 14) and term infants (108.9 ± 14) were higher than expected. Moderate to severe cognitive delay on the

Bayley III was identified in only 3% of their PT cohort, compared with 8% of the NRN cohort. Only 7% of their cohort had a language score >2 SD below the mean, compared with 20% of the NRN cohort. Studies of early outcomes of ELBW infants from Australia consistently report higher cognitive scores than US studies.^{22,23} A study of PT children (GA <30 weeks) at age 2 years found that 19%-20% of Bayley III language scores and 10%-12% of Cog scores were >2 SD below the mean, consistent with the findings in our cohort.²⁴

The Bayley II and III tests have some unique characteristics that may be contributing to the changes in scores. Of the 30 language items included in the Bayley II MDI, 24 (80%) are also used in the Bayley III language score. The content of some items has changed, however; for example, question 127 of Bayley II states that the child must use a 3-word sentence, whereas in the Bayley III, this is changed to 2 multiple-word utterances. Four (13%) Bayley II language items that were a part of the MDI are now included in the Bayley III Cog scale, and 2 items (6%) were dropped. In addition, the Bayley II uses item sets with established start and stop points, which may establish an artificial ceiling. The Bayley III also has a start point based on age, but the examiner continues with the test items until the child misses 5 in a row. This allows a bright child to achieve a higher level. It is apparent that the changes in the test design, test items, and method of test administration that may contribute to higher scores. It also remains a concern that updating a cognitive test usually resets to lower rather than higher cognitive scores, and that the mean scores for the PT group in the standardization sample are high for both language (97.0 ± 17) and motor (96.4 ± 15).

As hypothesized, the NDI rate was reduced by 70% (from 43% in period 1 to 13% in period 2; $P < .001$) using the new definition of Bayley III Cog <70. The odds for NDI using this definition in period 2 were approximately one-seventh of those in period 1 after adjusting for important perinatal and neonatal confounders. Language and fine motor delays were not included in the period 2 NDI definition. Although a PDI of >2 SD below the mean was included in the definition in period 1, during period 2, a GMFCS score 2 was included to account for significant gross motor delay. Incorporating the Bayley III motor subtest, effective January 1, 2010, had minor effects on the NDI rate, increasing it from 13% to 18%. In the matched comparisons, Bayley III motor scores were 5 points higher than PDI scores. Other changes affected NDI during period 2, however; mothers were more likely to receive antenatal steroids and to have a high school diploma, and infants were less likely to have ROP and sepsis.

We performed additional analyses to determine the impact of various components of NDI. The addition of a low language score increased the NDI rate to 23%. A Cog index of <85 was most comparable with the period 1 NDI rate (32% vs 43%; $P < .001$), an 11% difference. Of interest, the rates of components of NDI other than Bayley Cog score also decreased in period 2, including moderate to severe CP (from 8% to 6%), low motor score (from 26% to 14%), and bilateral hearing loss (from 3% to 2%). These data suggest that in addition to the significant changes in Bayley scores, there is also some small, but actual improvement in other outcomes as well. Other investigators have reported a decreased rate of CP in ELBW infants.^{25,26} Our findings suggest that there are limitations to the use of NDI over time, and that neurosensory outcomes should be reported separately from cognitive and language outcomes rather than as a combined outcome. We might have been overidentifying major cognitive impairment and underidentifying minor cognitive impairments with a Bayley II threshold of 70. Further study is needed to examine whether children with lesser degrees of cognitive impairment (70–85) on the Bayley III at 18 months CA are at increased risk for school-age learning disorders. A comparison of major neurosensory impairments (CP, bilateral blindness, bilateral hearing impairment) for the 2 time periods in matched samples found that the rate remained stable in adjusted analyses, at

9% for period 1 and 8% for period 2. Identification of a specified outcome that matches the study or clinical objective is the goal.

The establishment of a threshold both for outcomes of a trial and eligibility for services has clinical significance. Because Bayley III Cog scores are higher, eligibility for early intervention may be affected in some states. Thus, it is recommended that all ELBW infants be offered early intervention services. A 2011 Pediatric Academic Societies presentation by investigators from the EPICure study addressed the issue of threshold. These investigators developed an equation to determine equivalency from their cohort of extremely PT children at age 29–41 months, and concluded that a Bayley II MDI score of <70 was equivalent to a Bayley III Cog score of <80 in predicting NDI.²⁷ Our data support an equivalent cutpoint of 80–85.

Although rates of the 4 major neonatal medical morbidities remained unchanged between the 2 study periods, the proportion of children with and without a risk factor who had NDI dropped significantly. After adjusting for multiple confounders, during both study periods, only IVH grade 3–4/PVL was significantly associated with increased risk of NDI. The magnitude of the association with NDI was almost doubled in period 2, suggesting that by being more discriminatory, the Bayley III may be bringing expected associations between major neurologic morbidities and NDI into sharper relief. In other words, children with the lowest Bayley III scores are more likely to have a severe disability than those with the lowest Bayley II scores. It also suggests that the Bayley III is not effective in identifying mild developmental delay, and thus may reflect limited sensitivity at this age.

Differences in Bayley scores or the rate of NDI between the 2 study periods also may be related to changes in risk factors over time. The use of propensity score matching minimized possible differences in the risk factors studied between the 2 periods; however, it is possible that NDI rates in the 2 periods may be affected by other unmeasured factors or secular trends toward better cognitive outcomes over time. Other limitations of the present study are the lack of a full-term comparison group and only a single assessment of cognitive function. In addition, the NRN cohort is a unique population of ELBW infants from major neonatal intensive care units in the United States, and the findings can be compared only with similar cohorts of infants.

Whether the Bayley III is overestimating cognitive performance or whether it is a more valid assessment of emerging cognitive skills than Bayley II is unclear. Our findings indicate possible weaknesses in the reference standard test most commonly used in clinical trials and suggest that Bayley III scores should be interpreted with caution until these issues are clarified. Finally, because the determinants of cognitive outcomes at school age are multifactorial, the predictive value of tests administered at age 18–24 months will never approach 100%.

Acknowledgments

Supported by grants from the National Institutes of Health and the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development. Data collected at participating sites of the National Institute of Child Health and Human Development's Neonatal Research Network were transmitted to RTI International, the data coordinating center for the network, which stored, managed, and analyzed the data for this study.

Glossary

Bayley II	Bayley Scales of Infant Development, Second Edition
Bayley III	Bayley Scales of Infant and Toddler Development, Third Edition

BPD	Bronchopulmonary dysplasia
CA	Corrected age
Cog	Cognitive composite
CP	Cerebral palsy
ELBW	Extremely low birth weight
GA	Gestational age
GMFCS	Gross Motor Function Classification System
IVH	Intraventricular hemorrhage
MDI	Mental Developmental Index
NDI	Neurodevelopmental impairment
NEC	Necrotizing enterocolitis
NRN	Neonatal Research Network
PDI	Psychomotor Developmental Index
PT	Preterm
PVL	Periventricular leukomalacia
ROP	Retinopathy of prematurity

References

1. Hintz SR, Kendrick DE, Vohr BR, Poole WK, Higgins RD. Changes in neurodevelopmental outcomes at 18 to 22 months' corrected age among infants of less than 25 weeks' gestational age born in 1993–1999. *Pediatrics*. 2005; 115:1645–51. [PubMed: 15930228]
2. Laptook AR, O'Shea TM, Shankaran S, Bhaskar B. Adverse neurodevelopmental outcomes among extremely low birth weight infants with a normal head ultrasound: prevalence and antecedents. *Pediatrics*. 2005; 115:673–80. [PubMed: 15741371]
3. Vohr BR, Wright LL, Poole WK, Mc Donald SA. Neurodevelopmental outcomes of extremely low birth weight infants <32 weeks' gestation between 1993 and 1998. *Pediatrics*. 2005; 116:635–43. [PubMed: 16143580]
4. Vohr BR, Wright LL, Hack M, Aylward G, Hirtz D. Follow-up care of high-risk infants. *Pediatrics*. 2004; 114(Suppl):1377–97.
5. Schmidt B, Roberts RS, Davis P, Doyle LW, Barrington KJ, Ohlsson A, et al. Long-term effects of caffeine therapy for apnea of prematurity. *N Engl J Med*. 2007; 357:1893–902. [PubMed: 17989382]
6. Bayley, N. Bayley Scales of Infant Development. 2. San Antonio, TX: The Psychological Corporation; 1993.
7. Bayley, N. Bayley Scales of Infant Development. 3. San Antonio, TX: The Psychological Corporation; 2006.
8. Flynn J. Searching for justice: the discovery of IQ gains over time. *Am Psychol*. 1999; 54:5–20.
9. Parson, LS. Performing a 1:N case-control match on propensity score. Proceedings of the 29th Annual SAS User's Group International Conference; Cary, NC. May 9–12, 2004; Available from: <http://www2.sas.com/proceedings/sugi29/165-29.pdf>
10. Papile LA, Munsick-Bruno G, Schaefer A. Relationship of cerebral intra-ventricular hemorrhage and early childhood neurologic handicaps. *J Pediatr*. 1983; 103:273–7. [PubMed: 6875724]
11. Vohr BR, Wright LL, Dusick AM, Perritt R, Poole WK, Tyson JE, et al. Center differences and outcomes of extremely low birth weight infants. *Pediatrics*. 2004; 113:781–9. [PubMed: 15060228]

12. Amiel-Tison, C. Neuromotor status. In: Taeusch, HW.; Yogman, MW., editors. Follow-up management of the high-risk infant. Boston: Little, Brown; 1987.
13. Hack M, Taylor HG, Drotar D, Schluchter M, Cartar L, Wilson-Costello D, et al. Poor predictive validity of the Bayley Scales of Infant Development for cognitive function of extremely low birth weight children at school age. *Pediatrics*. 2005; 116:333–41. [PubMed: 16061586]
14. Luu TM, Ment LR, Schneider KC, Katz KH, Allan WC, Vohr BR. Lasting effects of preterm birth and neonatal brain hemorrhage at 12 years of age. *Pediatrics*. 2009; 123:1037–44. [PubMed: 19255037]
15. Taylor HG, Klein N, Minich NM, Hack M. Middle-school-age outcomes in children with very low birthweight. *Child Dev*. 2000; 71:1495–511. [PubMed: 11194251]
16. Roberts G, Anderson PJ, Doyle LW. The stability of the diagnosis of developmental disability between ages 2 and 8 in a geographic cohort of very preterm children born in 1997. *Arch Dis Child*. 2010; 95:786–90. [PubMed: 19828882]
17. Aylward GP. Methodological issues in outcome studies of at-risk infants. *J Pediatr Psychol*. 2002; 27:37–45. [PubMed: 11726678]
18. Aylward GP. Developmental screening and assessment: what are we thinking? *J Dev Behav Pediatr*. 2009; 30:169–73. [PubMed: 19363370]
19. Aylward GP, Aylward BS. The changing yardstick in measurement of cognitive abilities in infancy. *J Dev Behav Pediatr*. 2011; 32:465–8. [PubMed: 21555956]
20. Wechsler, D. The Wechsler Intelligence Scale for Children. 4. San Antonio, TX: The Psychological Corporation; 2003.
21. Anderson PJ, De Luca CR, Hutchinson E, Roberts G, Doyle LW. Underestimation of developmental delay by the new Bayley III Scale. *Arch Pediatr Adolesc Med*. 2010; 164:352–6. [PubMed: 20368488]
22. Doyle LW. Evaluation of neonatal intensive care for extremely low birth weight infants in Victoria over two decades, I: effectiveness. *Pediatrics*. 2004; 113(3 Pt 1):505–9. [PubMed: 14993541]
23. Kitchen WH, Doyle LW, Ford GW, Murton LJ, Keith CG, Rickards AL, et al. Changing two-year outcome of infants weighing 500 to 999 grams at birth: a hospital study. *J Pediatr*. 1991; 118:938–43. [PubMed: 1828267]
24. Spittle AJ, Anderson PJ, Lee KL, Ferretti C, Eeles A, Orton J, et al. Preventive care at home for very preterm infants improves infant and caregiver outcomes at 2 years. *Pediatrics*. 2010; 126:e171–8. [PubMed: 20547650]
25. Robertson CMT, Watt M-J. Changes in the prevalence of cerebral palsy for children born very prematurely within a population-based program over 30 years. *JAMA*. 2007; 297:2733–40. [PubMed: 17595274]
26. Hack M, Costello DW. Trends in the rates of cerebral palsy associated with neonatal intensive care of preterm children. *Clin Obstet Gynecol*. 2008; 51:763–74. [PubMed: 18981801]
27. Moore T, Johnson S, Haider S, Hennessy E, Marlow N. Relationship between test scores using the second and third editions of the Bayley Scales in extremely premature children. *J Pediatr*. 2011 Epub ahead of print.

Appendix

The following investigators, in addition to those listed as authors, participated in this study:

NRN Steering Committee Chair: Alan H. Jobe, MD, PhD, University of Cincinnati; Michael S. Kaplan, MD, University of Chicago, Pritzker School of Medicine

Alpert Medical School of Brown University and Women & Infant's Hospital of Rhode Island (U10 HD27904): William Oh, MD, Abbot R. Lupton, MD, Angelita M. Hensman, RN, BSN, Barbara Alksnis, PNP, Dawn Andrews, RN, Kristen Angela, RN, Bill Cashore, MD, Kim Francis, RN, Regina A. Gargus, MD, FAAP, Dan Gingras, RRT, Shabnam Lainwala, MD, Theresa M. Leach, MEd, CAES, Martha R. Leonard, BA, BS James R. Moore, MD, Lucy Noel, Rachel V. Walden, MD, Victoria E. Watson, MS, CAS

Case Western Reserve University, Rainbow Babies & Children's Hospital (U10 HD21364, M01 RR80): Michele C. Walsh, MD, MS, Avroy A. Fanaroff, MD, Nancy S. Newman, BA, RN, Bonnie S. Siner, RN, Harriet G. Friedman, MA

Cincinnati Children's Hospital Medical Center, University Hospital and Good Samaritan Hospital (U10 HD27853, M01 RR8084): Kurt Schibler, MD, Edward F. Donovan, MD, Jean J. Steichen, MD, Kate Bridges, MD, Barbara Alexander, RN, Cathy Grisby, BSN, CCRC, Holly L. Mincey, RN, BSN, Jody Hessling, RN, Teresa L. Gratton, PA, Marcia Worley Mersmann, RN, CCRC

Duke University School of Medicine, University Hospital, Alamance Regional Medical Center, and Durham Regional Hospital (U10 HD40492, M01 RR30): Ronald N. Goldberg, MD, C. Michael Cotten, MD, MHS, William F. Malcolm, MD, Patricia Ashley, MD, Kathy J. Auten, MSHS, Kimberley A. Fisher, PhD, FNP-BC, IBCLC, Katherine A. Foy, RN, Sandra Grimes, RN, BSN, Kathryn E. Gustafson, PhD, Melody B. Lohmeyer, RN, MSN

Emory University, Children's Healthcare of Atlanta, Grady Memorial Hospital, and Emory University Hospital Midtown (U10 HD27851, M01 RR39): Barbara J. Stoll, MD, David P. Carlton, MD, Ellen C. Hale, RN, BS, CCRC, Sheena Carter, PhD, Maureen Mulligan LaRossa, RN

Eunice Kennedy Shriver National Institute of Child Health and Human Development: Rosemary D. Higgins, MD, Stephanie Wilson Archer, MA

Floating Hospital for Children at Tufts Medical Center (U10 HD53119, M01 RR54): Ivan D. Frantz III, MD, John M. Fiascone, MD, Brenda L. MacKinnon, RNC, Ellen Nylen, RN, BSN, Anne Furey, MPH

Indiana University, University Hospital, Methodist Hospital, Riley Hospital for Children, and Wishard Health Services (U10 HD27856, M01 RR750): Brenda B. Poindexter, MD, MS, James A. Lemons, MD, Dianne E. Herron, RN, Leslie Dawn Wilson, BSN, CCRC, Jessica Bissey, PsyD, HSPP, Ann B. Cook, MS, Faithe Hamer, BS, Carolyn Lytle, MD, MPH, Heike M. Minnich, PsyD, HSPP

RTI International (U10 HD36790): W. Kenneth Poole, PhD, Dennis Wallace, PhD, Jeanette O'Donnell Auman, BS, Margaret Cunningham, BS, Amanda R. Irene, BS, Jamie E. Newman, PhD, MPH, Carolyn M. Petrie Huitema, MS, James W. Pickett II, BS, Scott E. Schaefer, MS, Kristin M. Zaterka-Baxter, RN, BSN

Stanford University, Dominican Hospital, El Camino Hospital, and Lucile Packard Children's Hospital (U10 HD27880, M01 RR70): Krisa P. Van Meurs, MD, David K. Stevenson, MD, M. Bethany Ball, BS, CCRC, Marian M. Adams, MD, Joan M. Baran, PhD, Ginger K. Brudos, PhD, Maria Elena DeAnda, PhD, Anne M. DeBattista, RN, PNP, Jean G. Kohn, MD, MPH, Renee P. Pyle, PhD

University of Alabama at Birmingham Health System and Children's Hospital of Alabama (U10 HD34216, M01 RR32): Waldemar A. Carlo, MD, Namasivayam Ambalavanan, MD, Amanda Soong, MD, Monica V. Collins, RN, BSN, MEd, Shirley S. Cosby, RN, BSN, Vivien A. Phillips, RN, BSN, Kirstin J. Bailey, PhD, Fred J. Biasini, PhD, Stephanie A. Chopko, PhD, Kristen C. Johnston, MSN, CRNP, Kathleen G. Nelson, MD, Cryshelle S. Patterson, PhD, Richard V. Rector, PhD, Sally Whitley, MA, OTR-L, FAOTA

University of California San Diego Medical Center and Sharp Mary Birch Hospital for Women and Newborns (U10 HD40461): Neil N. Finer, MD, Maynard R. Rasmussen, MD,

Paul R. Wozniak, MD, Martha G. Fuller, RN, MSN, Kathy Arnell, RNC, Renee Bridge, RN, Clarence Demetrio, RN, Donna Posin, OTR/L, MPA, Wade Rich, BSHS, RRT

University of Iowa, Children's Hospital (GCRC M01 RR59, U10 HD53109): Edward F. Bell, MD, Karen J. Johnson, RN, BSN, Diane L. Eastman, RN, CPNP, MA, Nancy J. Krutzfield, RN, MA

University of Miami, Holtz Children's Hospital (U10 HD21397): Shahnaz Duara, MD, Ruth Everett-Thomas, RN, MSN, Maria Calejo, MS, Alexis N. Diaz, BA, Silvia M. Frade Eguaras, BA, Andrea Garcia, MA, Kasey Hamlin-Smith, PhD, Michelle Harwood, PhD, Sylvia Hiriart-Fajardo, MD, Elaine O. Mathews, RN, Helina Pierre, BA, Arielle Riguard, MD, Alexandra Stroerger, BA

University of New Mexico Health Sciences Center (U10 HD53089, M01 RR997): Kristi L. Watterberg, MD, Robin K. Ohls, MD, Conra Backstrom Lacy, RN, Jean Lowe, PhD, Rebecca Montman, BSN

University of Rochester Medical Center, Golisano Children's Hospital (U10 HD40521, M01 RR44, UL1 RR024160): Dale L. Phelps, MD, Linda J. Reubens, RN, CCRC, Erica Burnell, RN, Julie Babish Johnson, MSW, Cassandra A. Horihan, MS, Diane Hust, MS, RN, CS, Rosemary L. Jensen, Emily Kushner, MA, Joan Merzbach, LMSW, Lauren Zwetsch, PNP, Kelley Yost, PhD

University of Texas Southwestern Medical Center at Dallas, Parkland Health and Hospital System, and Children's Medical Center Dallas (U10 HD40689, M01 RR633): Pablo J. Sánchez, MD, Charles R. Rosenfeld, MD, Walid A. Salhab, MD, Alicia Guzman, Gaynelle Hensley, RN, Nancy A. Miller, RN, Elizabeth Heyne, PA-C, Linda A. Madden, BSN, RN, CPNP, Sally Adams, PNP, Janet S. Morgan, RN, Catherine Twell Boatman, MS, Melisa H. Leps, RN, Lizette E. Torres, RN

University of Texas Health Science Center at Houston Medical School, Children's Memorial Hermann Hospital, and Lyndon Baines Johnson General Hospital/Harris County Hospital District (U10 HD21373): Jon E. Tyson, MD, MPH, Kathleen A. Kennedy, MD, MPH, Esther G. Akpa, RN, BSN, Nora I. Alaniz, BS, Susan Dieterich, PhD, Charles Green, PhD, Beverly Harris, RN, BSN, Margarita Jiminez, MD, MPH, Anna E. Lis, RN, BSN, Sarah Martin, RN, BSN, Georgia E. McDavid, RN, Brenda H. Morris, MD, M. Layne Poundstone, RN, BSN, Stacey Reddoch, BA, Saba Siddiki, MD, Maegan C. Simmons, RN, Patti L. Pierce Tate, RCP, Sharon L. Wright, MT

University of Utah Medical Center, Intermountain Medical Center, LDS Hospital, and Primary Children's Medical Center (U10 HD53124, M01 RR64): Roger G. Faix, MD, Bradley A. Yoder, MD, Karen A. Osborne, RN, BSN, CCRC, Jennifer J. Jensen, RN, BSN, Cynthia Spencer, RNC, Kimberlee Weaver-Lewis, RN, BSN, R. Karena Strong, RN, BSN, Mike Steffens, PhD, Jill Burnett, RN, Shawna Baker, RN

Wake Forest University, Baptist Medical Center, Brenner Children's Hospital, and Forsyth Medical Center (U10 HD40498, M01 RR7122): T. Michael O'Shea, MD, MPH, Nancy J. Peters, RN, CCRP, Carroll Peterson, MA, Ellen L. Waldrep, MS, Lisa K. Washburn, MD, Cherrie D. Welch, MD, MPH, Melissa Whalen Morris, MA, Gail Wiley, Hounshell, PhD

Wayne State University, Hutzel Women's Hospital, and Children's Hospital of Michigan (U10 HD21385): Seetha Shankaran, MD, Rebecca Bara, RN, BSN, Geraldine Muran, RN, BSN, Laura Goldston, MA

Yale University, Yale-New Haven Children's Hospital, and Bridgeport Hospital (U10 HD27871, M01 RR125, UL1 RR24139): Christine G. Butler, MD, Harris Jacobs, MD, Patricia Cervone, RN, Patricia Gettner, RN, Sheila Greisman, RN, Monica Konstantino, RN, BSN, JoAnn Poulsen, RN, Janet Taft, RN, BSN, Joanne Williams, RN, BSN, Nancy Close, PhD, Walter Gilliam, PhD, Elaine Romano, MSN

National Institute of Child Health and Human Development NRN standard definitions of NDI

Table 1

NDI component	Period 1	Period 2
Neurologic	Moderate to severe CP with GMFCS 2	Moderate to severe CP with GMFCS 2
Development	Bayley II MDI <70 or PDI <70	Bayley III cognitive score <70; GMFCS 2
Vision	<20–200 bilateral	<20–200 bilateral
Hearing	Bilateral amplification for permanent hearing loss	Permanent hearing loss that does not permit the child to understand the directions of the examiner and communicate despite amplification with a cochlear implant or hearing aids.

Children meeting the criteria for any of the 4 components are classified as having an NDI. For comparison purposes, we used the period 1 definition of hearing loss for both cohorts in our analyses.

Table II

Maternal and infant characteristics by study period

Characteristic	Original samples			Matched samples		
	Period 1 (n = 1012)	Period 2 (n = 1616)	P value	Period 1 (n = 922)	Period 2 (n = 922)	P value
Maternal						
Age, years, mean $\pm$ SD	28 $\pm$ 7	28 $\pm$ 6	.771	28 $\pm$ 7	27 $\pm$ 6	.368
Education, n (%)						
Less than high school	202 (20)	288 (18)	.027	196 (21)	192 (21)	.819
High school graduate	489 (48)	913 (56)	.020	481 (52)	485 (53)	.852
Some college or more	250 (25)	409 (25)	.517	245 (27)	245 (27)	1.000
Public insurance, n (%)	661 (65)	1167 (72)	.145	476 (52)	470 (51)	.780
White race, n (%)	529 (52)	864 (53)	.288	643 (70)	646 (70)	.879
Infant						
Male sex, n (%)	507 (50)	805 (50)	.948	456 (49)	445 (48)	.608
Birth weight, g, mean $\pm$ SD	746 $\pm$ 134	741 $\pm$ 129	.332	746 $\pm$ 133	748 $\pm$ 129	.724
GA, weeks, mean $\pm$ SD	25 $\pm$ 1	25 $\pm$ 1	.671	25 $\pm$ 1	25 $\pm$ 1	.924
Multiple birth, n (%)	223 (22)	389 (24)	.213	201 (22)	204 (22)	.866
Cesarean delivery, n (%)	633 (63)	1060 (66)	.090	573 (62)	567 (62)	.774
Antenatal steroids, n (%)	861 (85)	1421 (88)	.009	789 (86)	797 (86)	.591
Postnatal steroids, n (%)	146 (14)	259 (16)	.209	132 (14)	135 (15)	.843
Sepsis, n (%)	487 (48)	693 (43)	.011	450 (49)	450 (49)	1.000
NEC >stage 2, n (%)	94 (9)	162 (10)	.519	84 (9)	93 (10)	.477
IVH grade 3-4/PVL, n (%)	178 (18)	277 (17)	.778	156 (17)	157 (17)	.951
ROP, n (%)	808 (80)	1206 (75)	.005	736 (80)	740 (80)	.816
Days in hospital, mean $\pm$ SD	113 $\pm$ 46	114 $\pm$ 43	.517	113 $\pm$ 46	113 $\pm$ 39	.816
BPD, n (%)	546 (54)	867 (54)	.866	498 (54)	505 (55)	.744
Days on ventilation, mean $\pm$ SD	33 $\pm$ 27	32 $\pm$ 27	.326	32 $\pm$ 28	32 $\pm$ 27	.565

Table III

Child development, motor, and neurosensory outcomes by study period

Variable	Original samples			Matched samples		
	Period 1 (n = 1012)	Period 2 (n = 1616)	P value	Period 1 (n = 922)	Period 2 (n = 922)	P value
CA at follow-up, years, mean $\pm$ SD	19.9 $\pm$ 2.6	20.1 $\pm$ 2.8	.016	20 $\pm$ 2	20 $\pm$ 2	.891
Development						
MDI versus cognitive score, mean $\pm$ SD	77 $\pm$ 19	88 $\pm$ 15	<.001	77 $\pm$ 19	88 $\pm$ 15	<.001
MDI <70 versus cognitive <70, n (%)	375 (37)	159 (10)	<.001	339 (37)	92 (10)	<.001
MDI <70 versus cognitive <75, n (%)	375 (37)	237 (15)	<.001	339 (37)	139 (15)	<.001
MDI <70 versus cognitive <80, n (%)	375 (37)	328 (20)	<.001	339 (37)	194 (21)	<.001
MDI <85 versus cognitive <85, n (%)	643 (64)	491 (30)	<.001	581 (63)	290 (31)	<.001
MDI <55 versus cognitive <55, n (%)	188 (19)	47 (3)	<.001	167 (18)	25 (3)	<.001
MDI <70 versus language/cognitive <70, n (%)	375 (37)	347 (21)	<.001	339 (37)	201 (22)	<.001
MDI <70 versus language/cognitive <80, n (%)	375 (37)	743 (46)	<.001	339 (37)	447 (49)	<.001
MDI <85 versus language/cognitive <85, n (%)	643 (64)	886 (55)	<.001	581 (63)	528 (58)	.027
Language score, mean $\pm$ SD	NA	84 $\pm$ 17	NA	N/A	83 $\pm$ 18	NA
Language score <70, n (%)	NA	312 (19)	NA	N/A	178 (20)	NA
Motor						
Any CP, n (%)	164 (16)	202 (13)	.007	144 (16)	140 (15)	.796
Moderate/severe CP, n (%)	76 (8)	90 (6)	.046	65 (7)	60 (7)	.643
GMFCS 2, n (%)	82 (8)	117 (7)	.426	70 (8)	74 (8)	.718
PDI versus motor, mean $\pm$ SD*	81 $\pm$ 20	87 $\pm$ 17	<.001	81 $\pm$ 19	86 $\pm$ 17	<.001
PDI <70 versus motor <70, n (%)*	268 (27)	84 (14)	<.001	233 (25)	51 (15)	<.001
PDI <55 versus motor <55, n (%)*	185 (18)	41 (3)	<.001	159 (17)	25 (8)	<.001
Neurosensory, n (%)						
Vision impairment	9 (1)	8 (0)	.220	6 (1)	4 (0)	.526
Hearing impairment	28 (3)	26 (2)	.039	23 (2)	23 (2)	1.000
CP, vision or hearing impairment	95 (9)	113 (7)	.027	79 (9)	78 (8)	.934

NA, not available in Time 1.

* The Bayley III motor scale was administered to only a subset of children during period 2; 596 of the 1616 children in the original sample and 332 of the 922 children in the matched sample received the assessment.

Table IV
Child NDI based on NRN standard definitions with modified cutpoints for Bayley scores

Period 1			Period 2		Period 2 versus period 1		
Cutpoint for NDI definition	NDI, n (%)	Cutpoint for NDI definition	NDI, n (%)	Unadjusted OR (95% CI)	Adjusted OR (95% CI)	P value	
Original samples							
MDI/PDI <70	437 (43)	Cognitive <70	205 (13)	0.19 (0.16–0.23)	0.14 (0.10–0.18)	<.001	
MDI/PDI <70	437 (43)	Cognitive <80	360 (22)	0.38 (0.32–0.45)	0.35 (0.28–0.44)	<.001	
MDI/PDI <70	437 (43)	Cognitive <85	512 (32)	0.61 (0.52–0.72)	0.64 (0.52–0.79)	<.001	
MDI/PDI <70	437 (43)	Cognitive/language <70	376 (23)	0.41 (0.34–0.48)	0.38 (0.30–0.48)	<.001	
MDI/PDI <70	437 (43)	Cognitive/language <80	756 (47)	1.18 (1.01–1.38)	1.46 (1.18–1.80)	<.001	
MDI/PDI <70	437 (43)	Cognitive/language <85	891 (55)	1.66 (1.41–1.94)	2.21 (1.79–2.73)	<.001	
MDI/PDI <70	437 (43)	Cognitive/motor <70 *	106 (18)	0.29 (0.22–0.36)	0.23 (0.16–0.33)	<.001	
MDI/PDI <70	437 (43)	Cognitive/motor/language <70 *	162 (27)	0.49 (0.40–0.62)	0.50 (0.37–0.68)	<.001	
MDI/PDI <70	437 (43)	Cognitive/motor/language <80 *	324 (54)	1.57 (1.28–1.93)	2.14 (1.61–2.84)	<.001	
Matched samples							
MDI/PDI <70	394 (43)	Cognitive <70	128 (14)	0.22 (0.17–0.27)	0.16 (0.12–0.21)	<.001	
MDI/PDI <70	394 (43)	Cognitive <80	218 (24)	0.42 (0.34–0.51)	0.37 (0.29–0.48)	<.001	
MDI/PDI <70	394 (43)	Cognitive <85	304 (33)	0.66 (0.55–0.80)	0.67 (0.53–0.85)	<.001	
MDI/PDI <70	394 (43)	Cognitive/language <70	222 (24)	0.43 (0.36–0.53)	0.40 (0.31–0.52)	<.001	
MDI/PDI <70	394 (43)	Cognitive/language <80	455 (50)	1.34 (1.11–1.61)	1.62 (1.29–2.05)	<.001	
MDI/PDI <70	394 (43)	Cognitive/language <85	530 (58)	1.86 (1.55–2.24)	2.40 (1.90–3.04)	<.001	
MDI/PDI <70	394 (43)	Cognitive/motor <70 *	66 (20)	0.33 (0.25–0.45)	0.27 (0.18–0.40)	<.001	
MDI/PDI <70	394 (43)	Cognitive/motor/language <70 *	94 (28)	0.53 (0.41–0.70)	0.53 (0.37–0.76)	<.001	
MDI/PDI <70	394 (43)	Cognitive/motor/language <80 *	186 (56)	1.72 (1.33–2.22)	2.28 (1.63–3.18)	<.001	

The reference category for ORs is period 1. Adjusted ORs control for study period, maternal age, education, race, public insurance, antenatal steroids, cesarean delivery, birth weight, GA, multiple birth, sex, IVH 3–4/PVL, NEC, culture-positive sepsis, supplemental oxygen at 36 weeks, ROP, postnatal steroids, ventilation days, duration of hospital stay, center, and adjusted age at follow-up visit.

*The Bayley III motor scale was administered to only a subset of children in period 2; 596 of the 1616 children in the original sample and 332 of the 922 children in the matched sample underwent the assessment.

Table V
Relationships between neonatal major morbidities and NDI based on NRN standard definition by study period

Study period/risk factor	Original samples			Matched samples		
	NDI, n (%)	Unadjusted OR (95% CI)	Adjusted OR (95% CI)	NDI, n (%)	Unadjusted OR (95% CI)	Adjusted OR (95% CI)
Period 1						
Oxygen at 36 weeks						
Yes	297 (54)	2.78 (2.14–3.60) [‡]	1.42 (0.96–2.10)	270 (54)	2.87 (2.18–3.77) [‡]	1.44 (0.97–2.13)
No	140 (30)	Ref	Ref	124 (29)	Ref	Ref
NEC >stage 2						
Yes	46 (49)	1.29 (0.84–1.98)	0.76 (0.42–1.35)	42 (50)	1.38 (0.88–2.16)	0.80 (0.45–1.42)
No	391 (43)	Ref	Ref	352 (42)	Ref	Ref
IVH grade 3–4/PVL						
Yes	113 (63)	2.73 (1.95–3.81) [‡]	2.70 (1.75–4.16) [‡]	99 (63)	2.77 (1.94–3.96) [‡]	2.42 (1.57–3.72) [‡]
No	324 (39)	Ref	Ref	295 (39)	Ref	Ref
Sepsis						
Yes	235 (48)	1.49 (1.16–1.92) [‡]	0.83 (0.60–1.16)	211 (47)	1.39 (1.07–1.81) [‡]	0.87 (0.62–1.20)
No	202 (38)	Ref	Ref	183 (39)	Ref	Ref
Period 2						
Oxygen at 36 weeks						
Yes	147 (17)	2.46 (1.78–3.40) [‡]	1.03 (0.65–1.64)	92 (18)	2.36 (1.57–3.55) [‡]	1.03 (0.58–1.83)
No	56 (8)	Ref	Ref	36 (9)	Ref	Ref
NEC >stage 2						
Yes	32 (20)	1.82 (1.20–2.76) [‡]	1.26 (0.74–2.13)	23 (25)	2.27 (1.36–3.79) [‡]	1.42 (0.74–2.74)
No	173 (12)	Ref	Ref	105 (13)	Ref	Ref
IVH grade 3–4/PVL						
Yes	82 (30)	4.14 (3.01–5.69) [‡]	3.94 (2.62–5.95) [‡]	57 (36)	5.57 (3.71–8.37) [‡]	5.06 (3.04–8.43) [‡]
No	123 (9)	Ref	Ref	71 (9)	Ref	Ref
Sepsis						
Yes	124 (18)	2.25 (1.67–3.04) [‡]	1.51 (1.03–2.20) [*]	82 (18)	2.06 (1.40, 3.04) [‡]	1.10 (0.67, 1.79)
No	81 (9)	Ref	Ref	46 (10)	Ref	Ref

Ref. reference category.

Percentages are the proportion of children with (or without) the risk factor who have NDI (eg, 54% of children in period 1 with oxygen at 36 weeks had NDI). Adjusted ORs control for maternal age, education, race, public insurance, antenatal steroids, cesarean delivery, birth weight, GA, multiple birth, sex, IVH 3-4/PVL, NEC, culture-positive sepsis, supplemental oxygen at 36 weeks, ROP, postnatal steroids, ventilation days, duration of hospital stay, center, and adjusted age at follow-up visit.

* $P < .05$.

[†] $P < .01$.

[‡] $P < .001$.