Published in final edited form as:
Dev Med Child Neurol. 2009 July ; 51(7): 536–544. doi:10.1111/j.1469-8749.2009.03318.x.

Validation of the Functional Status II questionnaire in the assessment of extremely-low-birthweight infants

DAVID DA COSTA, MD[1], **CARLA M BANN, PHD**[2], **NELLIE I HANSEN, MPH**[2], **SEETHA SHANKARAN, MD**[3], and **VIRGINIA DELANEY-BLACK, MD, MPH**[3] **FOR THE NATIONAL INSTITUTE OF CHILD HEALTH AND HUMAN DEVELOPMENT NEONATAL RESEARCH NETWORK**

[1] Department of Pediatrics, William Beaumont Hospital, Royal Oak, MI, USA

[2] Department of Pediatrics RTI International, NC, USA

[3] Carmen and Ann Adams Department of Pediatrics, Children's Hospital of Michigan, Wayne State University, Detroit, MI, USA

Abstract

AIM—The increased survival of infants born at extremely low birthweight (ELBW) has been associated with significant morbidity, including higher rates of neurodevelopmental disability. However, formalized testing to evaluate these problems is both time-consuming and costly. The revised Functional Status questionnaire (FS-II) was designed to assess caregivers' perceptions of the functional status of children with chronic diseases.

METHOD—We evaluated the reliability and validity of the FS-II for ELBW infants at 18 to 22 months corrected age using data from the US Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) Neonatal Research Network (NRN). Exploratory factor analyses were conducted using data from the network's first follow-up study of 1080 children born in 1993 to 1994 (508 males, 572 females [53%]), and results were confirmed using data from the next network follow-up of 4022 children born in 1995 to 2000 (1864 males, 2158 females [54%]).

RESULTS—Results suggest that a two-factor solution comprising measures of general health and independence is most appropriate for ELBW infants. These factors differed from those found among chronically ill children, and new, more appropriate scales are presented for screening ELBW survivors. Both scales demonstrated good internal consistency: Cronbach's α=0.87 for general health and α=0.75 for independence. Construct validity of the scales was assessed by comparing mean scores on the scales according to scores on the Bayley Scales of Infant Development, second edition (BSID-II), and medical conditions.

INTERPRETATION—As hypothesized, infants with greater functional impairments according to their BSID-II scores or medical conditions had lower scores on the general health and independence scales, supporting the validity of the scales.

Correspondence to Dr Virginia Delaney-Black, Carmen and Ann Adams Department of Pediatrics, The Children's Research Center of Michigan, 3901 Beaubien, Detroit, MI 48201, USA. vdelaney@med.wayne.edu.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

This material is available as part of the online article from http://www.blackwell.synergy.com/doi/10.1111/j.1469-8749.2009.03318.x (this will link you to the article abstract).

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Decreases in perinatal and infant mortality[1] reflect the success of sustaining the lives of infants born increasingly more preterm. However, the increased survival of infants with extremely low birthweight (ELBW; 401–1000g) may be associated with more functional disabilities.[2–4] Overt neurodevelopmental impairments have been identified in up to 50% of surviving ELBW infants,[2,5] and subtle deficits have been noted even more frequently.[6–8] ELBW infants need ongoing developmental assessment well past infancy to identify disabilities, yet formalized testing is both time-consuming and costly, and, therefore, is often available only to the smallest or sickest newborn infants.

ELBW survivors with a disability may receive major health interventions to improve their quality of life and facilitate caregiving. Many of these interventions are resource-intensive and expensive, and are performed only when the child meets certain predefined criteria, criteria which may be less relevant to the child or the caregiver. It is, therefore, necessary to measure such interventions and outcomes from the family's perspective to assess their effectiveness. In this regard, a functional status measure would be useful both as a screening tool and as a measure of intervention outcome.

Many health-related quality-of-life questionnaires have been used in adults. However, as Hack[9] points out, adaptation of quality-of-life instruments for children is hampered by the necessary requirement of adding a proxy respondent. Despite this concern, several investigators have used quality-of-life measures to assess outcomes of even those born extremely preterm.[10–12]

Recognizing the difficulties associated with follow-up of infants born at ELBW, participants at a workshop sponsored by the National Institute of Child Health and Human Development (NICHD) developed consensus statements addressing criteria for follow-up and appropriate assessment methods and ages for evaluation. The consensus authors concluded that alternative, less costly methods and approaches should be developed to supplement current standardized testing.[13] Specifically needed were assessments that are easy to administer, that provide useful information to developmental specialists, parents, and teachers, and that can be incorporated into the routine of both healthcare and educational professionals.

Using the criteria from that workshop, the investigators of the NICHD Neonatal Research Network (NRN) evaluated available pediatric screening measures and chose to assess ELBW infants with the revised version of the Functional Status questionnaire (FS-II) developed by Stein and Jessop.[14] The FS-II, designed for children with chronic diseases, assesses parents' perceptions of their child in four domains: physical, psychological, cognitive, and social functioning. It can be administered by a trained layperson in less than 30 minutes, and has been used with a young sample. All of these attributes are appropriate for screening ELBW infants to assess the need for a more formal evaluation or early intervention. However, the FS-II had not been validated in a population of ELBW infants.

The analyses by Stein and Jessop used data from 732 children aged 2 weeks to 16 years who were either chronically ill or well. Principal-component analysis determined the factor structure associated with the questionnaire items for four age groups (<1y, 1y–23 mo, 2–3y, ≥4y). In each age group, a total score was derived from a one-factor solution, and two subscores were derived from a two-factor solution. Higher scores indicate more favorable functional status. In the absence of data for ELBW children, we aimed to evaluate the factor structure of the FS-II data collected from a large cohort of ELBW children at 18 to 22 months corrected age and to compare our results with those of Stein and Jessop. Specifically we hypothesized that the FS-II, as it exists, is a reliable and valid proxy of the health and functional status of ELBW children, and, if it was not, that a revised measure could be developed.

METHOD

Participants

The NICHD NRN maintains a registry of infants weighing between 401g and 1500g born at participating centers. Trained personnel collect maternal and delivery data soon after birth and infant data until discharge, 120 days of age, or death. Surviving infants with a birthweight of 1000g or less are asked to return for a comprehensive follow-up visit at 18 to 22 months corrected age. Data from infants born between January 1, 1993 and September 30, 2000 who participated in the follow-up study were evaluated. Participation in the follow-up study required written or verbal informed consent from the parent or primary caregiver. The institutional review board of each participating center approved the study. Two network follow-up studies were performed, which were a retrospective study of children born before December 31, 1994 and a prospective study of children born after that time who were enrolled before hospital discharge. To allow for cross-validation of results, we used this natural grouping of data to split the sample into the following two groups, based on the timing of network follow-up studies: children born between January 1, 1993 and December 31, 1994 (exploratory cohort) and those born between January 1, 1995 and September 30, 2000 (confirmatory cohort).

Assessment

Neonatal information collected in the registry included birthweight, gestational age, sex, race, respiratory distress syndrome, receipt of surfactants, bronchopulmonary dysplasia (BPD), and intraventricular hemorrhage (IVH). Using growth charts developed by Alexander et al.,[15] infants were classified as small for gestational age at birth, defined by birthweight below the 10th centile for sex and gestational age. Respiratory distress syndrome was defined as the presence of all four of the following: need for oxygen from 6 to 24 hours of life, clinical features of respiratory distress within 24 hours, need for respiratory support up to 24 hours, and abnormal chest radiograph within 24 hours. BPD was defined as the need for supplemental oxygen at 36 weeks postmenstrual age. IVH was defined by cranial ultrasonography using criteria based on Papile et al.[16]

The 18- to 22-month follow-up included physical and neurological examination and neurodevelopmental assessment using the Bayley Scales of Infant Development, second edition (BSID-II).[17] During the parent interview, the 28-item revised FS-II was administered by a trained interviewer. Questions are administered in two parts. Part 1 asks how often the child performs a specified activity or behavior. Part 2 probes all of the part 1 responses that are indicative of poor functioning to determine whether the behavior was due fully, partly, or not at all to chronic illness. For this study, a chronic illness was defined as either a long-term problem that the child had because of preterm birth or a similar-meaning chronic illness secondary to the child's history of preterm birth. Both probes were used to assure parent understanding. All FS-II items were recoded using the standard scoring instructions for the questionnaire.[18] As specified in the instructions, to isolate the impact of the children's chronic conditions on their functional status, if the parent indicated that the child exhibited poor functioning that was not at all because of illness associated with preterm birth (part 2 answers), that item was recoded to indicate good functioning. Items were reverse-coded as needed so that higher values indicated better health (e.g. standard coding for 'eat well' would assign a low numerical value for the most positive response; this was recoded so that the most positive responses now had a higher value). The mode was substituted for missing item responses. Only eight (<1%) of the 1080 children in the exploratory group and 53 (1%) of the 4022 children in the confirmatory group had missing values for any of the items. Of these 61 children, almost all (95%) were missing data for only one or two items.

As a sensitivity check, we ran the analyses excluding any children with missing data (rather than using mode substitution), and the results were nearly identical to those presented here.

Children living outside of the home or with friends of the family at the time of assessment were excluded from analysis, as it was likely that the individual providing the responses would have less familiarity with the child. An a priori decision was made that the sex of the primary caregiver might also influence the caregiver response. Thus data provided by male caregivers were excluded. Participants missing more than 50% of the responses on the FS-II were also excluded.

Statistical analysis

Characteristics of the exploratory and confirmatory cohorts were compared using the Fisher's exact or χ^2 test for categorical variables and the Wilcoxon test for continuous variables.

The psychometric properties of the FS-II were evaluated. To investigate the scale's factorial validity, a two-step procedure was used to identify the most appropriate factor structure for the FS-II. Factor analyses were performed initially for the exploratory cohort. As items were categorical rather than continuous, polychoric correlations were computed, and unweighted least-squares factor analysis was conducted. Polychoric correlations are used to assess the relationship between two categorical variables, assuming that they reflect continuous underlying variables. Promax, an oblique factor rotation method was used, because it was expected that the factors would be correlated. Squared multiple correlations were used as the prior communality estimates. The analyses were conducted using SAS version 8 (SAS Inc., Cary, NC, USA).

Factor structures from the exploratory factor analyses were tested using confirmatory factor analysis models conducted for the confirmatory cohort. Goodness-of-fit statistics for each model were examined to determine the most appropriate solution. We examined values for the comparative fit index (CFI), Tucker–Lewis index, and the root mean square error of approximation (RMSEA). If the model fits well, the CFI and Tucker–Lewis index should have values of 0.90 or greater, and the RMSEA should be 0.05 or less. Correlated errors were permitted between pairs of items with very similar content. Because the items are categorical, the maximum likelihood–robust means and variance estimator was used, providing parameter estimates that are robust to non-normality. We hypothesized that, consistent with the results from Stein and Jessop,[14] the FS-II would group into two factors.

As caregiver perceptions may vary depending on the child's sex, we tested the consistency of the factor structure across sex. First, we fit a two-factor multigroup model that constrained the factor loadings to be equal for both sexes; then we freed one loading at a time and computed the change in model χ^2 to determine whether the loadings for the individual items varied significantly across sex. The χ^2 difference tests were computed using the method described by Muthén and Muthén.[19]

Using the final factor structure, scores were computed for each study participant. Cronbach's alphas were computed to assess internal consistency of the scales. Scale construct validity was examined by comparing mean scores for groups of participants according to their BSID-II scores and whether they were blind (legally blind as reported by parents) or hearing impaired (required hearing aids in both ears), or had cerebral palsy or IVH grade 3 or 4. In addition, scores were compared for groups based on a composite measure of neurodevelopmental impairment, defined as one or more of the following: mental or psychomotor developmental index less than 70, cerebral palsy, blind in both eyes, or hearing aids in both ears. We expected lower FS-II scores for children with IVH grade 3

or 4 or neurodevelopmental impairment. The statistical significance of the mean comparisons was assessed using Wilcoxon tests.

Finally, we computed means and standard deviations of the two scale scores according to demographical characteristics. Wilcoxon tests were conducted to test for significant overall differences in scores, and Tukey's studentized range to test for comparisons between specific groups. A *p*-value of <0.05 was considered statistically significant.

RESULTS

A total of 11 240 ELBW infants born between January 1993 and December 2000 were admitted to NICHD NRN centers. Of the 7131 who survived, 5192 (73%) had parents who completed the FS-II at 18 to 22 months. Ninety children (2%) were excluded because their caregiver did not attend the follow-up session, they were not living at home, or they had a male primary caregiver. The exploratory cohort had 1080 participants; 4022 were in the confirmatory cohort, based on the timing of the NICHD NRN studies (total *n*=5102).

Neonatal and sociodemographical characteristics between groups are compared in Tables I and II. In the confirmatory cohort, there were more white children, greater use of surfactants, and more children diagnosed with BPD. However, fewer newborn infants in the confirmatory cohort were diagnosed with respiratory distress syndrome or severe IVH, and the median time to discharge was 3 days shorter. In the confirmatory cohort at follow-up, the primary caregiver was more likely to be the mother, to have at least a high-school diploma, and to have a higher household income and private medical insurance. These differences most likely reflect differences in NICHD NRN enrollment sites as well as changes in medical practice between time periods.

Exploratory factor analysis was conducted using FS-II scores obtained from the exploratory cohort. Based on the eigenvalues, scree plot, and pattern of loadings, the results suggest a two-factor solution; eigenvalues for the first two factors are 12.3 and 2.5 (Table SI, supporting information published online, which contains factor loadings of the two-factor solution). Consistent with Stein and Jessop's results for chronically ill children at 12 to 23 months,[14] factor 1 appears to measure general health. However, some items included in the general health factor differed between studies. In contrast to our results, items FS1 (eats well) and FS25 (no trouble with task) loaded on the general health factor in Stein and Jessop's factor analysis, whereas items FS13 (interested in environment), FS23 (temper tantrums), and FS27 (timid) did not. Differences were also found between the studies on items included in factor 2. Stein and Jessop found that items FS6 (sick and tired), FS7 (occupies self), FS9 (irritable), FS12 (seems difficult), FS13 (interested in environment), and FS30 (concentrates) but not item FS5 (communicates) loaded on the second factor. Although Stein and Jessop labeled this factor 'responsiveness', it appears that in our ELBW cohort the second factor may be measuring independence. Our factor 2 includes measures of a child's ability to function without assistance, such as items FS22 (gets around house) and FS28 (eats independently), whereas some items indicative of responsiveness, such as FS9 (irritable) and FS13 (interested in environment), did not load on this factor.

In addition to the two-factor solution, we also fit a one-factor model. Using the ELBW cohort data, all items loaded on this overall factor except item FS15 (special equipment), which had a loading of 0.39. By contrast, items FS5 (communicates), FS22 (gets around house), FS23 (temper tantrums), FS27 (timid), and FS28 (eats independently) did not load on the total factor in Stein and Jessop's analysis.[14]

We tested these one- and two-factor solutions by conducting a confirmatory factor analysis using data collected from the 4022 infants included in the later cohort. The two-factor

solution provided a better fit, with higher comparative fit and Tucker–Lewis indices (CFI: 0.93 vs 0.87; Tucker–Lewis index: 0.97 vs 0.95), and lower RMSEA (0.047 vs 0.068). These findings indicate that the FS-II items may be most appropriately divided into two factors, one representing general health and the other representing independence, and confirm the results from our exploratory factor analysis.

A multigroup factor analysis of the confirmatory cohort data suggested that the factor structure is generally consistent across male and female infants. The items split into the same two factors for both sexes, and the χ^2 difference tests suggested that loadings for only two of the items (FS27=timid and FS23=temper tantrums) varied significantly across the two sexes, with a marginally significant difference on a third item (FS13=interested in environment). Results by sex are shown in Table SII (supporting information, published online).

Data from the exploratory and confirmatory cohorts were combined and two new scales corresponding to the two factors were calculated using the following equations:

$$\text{general health} = ([\text{sum of items } \{2-4, 6, 8-14, 21, 23, 24, 27, 29, 32, 33\}]/36) \times 100$$

$$\text{independence} = ([\text{sum of items } \{5, 22, 25, 26, 28, 34\}]/12) \times 100$$

Both scales demonstrated good internal consistency (Cronbach's α=0.87 for general health, α=0.75 for independence). To assess the construct validity of the scales, mean scores on the two scales were compared according to BSID-II scores and medical conditions (Table III). Infants with greater functional impairments based on their BSID-II scores or medical conditions had significantly lower scores on the general health and independence scales ($p<0.001$). As may be expected, the differences appear to be particularly pronounced for the independence scale.

To provide comparison data for future users of these scales, mean scale scores and percentages of children scoring one or two standard deviations below the mean are shown for children in this high-risk cohort by neonatal characteristics in Table IV. On average, scores were higher for females than for males and for those with greater birthweights and born at later gestational ages. Significant differences were also observed by race, with black children having lower scores than white children.

DISCUSSION

Functional assessment in children is described by McCabe and Granger as 'an effort to systematically describe and measure a child's abilities and limitations when performing the activities of daily living.'[20] Well-characterized FS-IIs have not been widely available for infants born at ELBW. In older children, at least two problems have limited the use of functional status or quality-of-life questionnaires. First, the respondent is typically the parent rather than the child. Second, functional status measures for children must account for the normal, nonlinear evolution of developmental skills, particularly for very young children. In adults, functional status or quality-of-life assessment responses are influenced by cultural, social, and educational backgrounds of the individual; similar effects may be seen in children. In addition for child measures, the child's medical history, use of outpatient services, and expectations for the child all play a significant role in shaping parental responses. Responses may also be altered by the respondent's knowledge of normal child development, previous child-rearing experiences, and exposure to the developmental

assessments and educational enrichment programs that the child attends. Despite these limitations, in older children functional status or quality-of-life questionnaires have been found to be accurate in determining overall developmental delay.[21]

The FS-II was originally developed and revised by Stein and Jessop[14] for infants and children with chronic illnesses. It has age-appropriate sets of questions and is designed to measure the impact of chronic physical disorders on the functioning of children across the entire pediatric age group. It has also been used by the Infant Health and Development Program to assess improvements in health status as a result of early intervention programs for low-birthweight infants.[22] Additionally, the parent version, but not the child version, of a Dutch adaptation of the FS-II has been reliably used to measure health status of children with asthma.[23]

This study was conducted using two similar samples of ELBW survivors assessed at 18 to 22 months corrected age. Our combined sample size of 5102 children is, to our knowledge, the largest in the literature assessing pediatric health-related quality-of-life measures. We found that both scales (general health and independence) derived using the combined data have acceptable internal consistency reliability with Cronbach's alphas of 0.87 and 0.75 respectively.

On the basis of the factor analyses of both cohorts from the NICHD NRN, we recommend that, for the evaluation of ELBW infants, clinicians use two scales of revised FS-II items representing general health and independence. Given that the factor structure was generally consistent across the two sexes, with only three of the 26 items exhibiting sex differences, we also recommend using the same scales for both sexes.

In addition to the FS-II, other measures of functional ability and health-related quality of life in children are now available, including the Functional Independence Measure for Children (WeeFIM), which was developed in 1993[10,24] when the NICHD NRN was making its choice of measures. Other measures including the standard version of the Vineland Adaptive Behavior Scales and the Pediatric Evaluation of Disability Inventory[25] are available. However, the former consists of 301 items and takes 45 to 60 minutes to administer, and the latter, which is meant to be a discriminative measure of functional limitation in children, requires 45 minutes and administration by trained personnel.

In summary, based on exploratory and subsequent confirmatory factor analyses, we validated the revised FS-II in the follow-up of high-risk ELBW infants. We propose new general health and independence scales and have provided the equations for determining the scale scores. As a next step we propose to assess how well these scores predict neurodevelopmental status in this cohort of ELBW infants. If predictive, the revised FS-II could serve as an easy-to-administer screening tool in similar cohorts of ELBW infants for identifying children needing formalized developmental testing.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors wish to thank the children and families and their colleagues who made this study possible. The study was supported by grants U10 HD21397, U10 HD34216, U10 HD27853, M01 RR 08084, U10 HD27871, M01 RR 06022, U10 HD21364, U10 HD21415, U10 HD40689, U10 HD27856, M01 RR 00750, U10 HD27904, U10 HD27881, M01 RR 00997, U01 HD36790, U10 HD21385, U10 HD34167, M01 RR 02635, M01 RR 02172, M01 RR 01032, U10 HD27880, M01 RR 00070, U10 HD27851, M01 RR 00039, U10 HD21373, and U01 HD19897 from the US National Institutes of Health.

LIST OF ABBREVIATIONS

BPD	Bronchopulmonary dysplasia
BSID-II	Bayley Scales of Infant Development, second edition
CFI	Comparative fit index
ELBW	Extremely low birthweight
FS-II	Functional Status Questionnaire (revised)
IVH	Intraventricular hemorrhage
NICHD	National Institute of Child Health and Human Development
NRN	Neonatal Research Network
RMSEA	Root mean square error of approximation

References

1. Hintz SR, Poole WK, Wright LL, et al. Changes in mortality and morbidities among infants born at less than 25 weeks during the post-surfactant era. Arch Dis Child Fetal Neonatal Ed 2005;90:F128–33. [PubMed: 15724036]
2. Hintz SR, Kendrick DE, Vohr BR, Poole WK, Higgins RD. for the National Institute of Child Health and Human Development Neonatal Research Network. 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]
3. Bhushan V, Paneth N, Kieley JL. Impact of improved survival of very low birth weight infants on recent secular trends in the prevalence of cerebral palsy. Pediatrics 1993;91:1094–100. [PubMed: 8502508]
4. Rijken M, Stoelhorst GMSJ, Martens SE, et al. Mortality and neurologic, mental, and psychomotor development at 2 years in infants born less than 27 weeks' gestation: the Leiden follow-up project on prematurity. Pediatrics 2003;112:351–58. [PubMed: 12897286]
5. Vohr BR, Wright LL, Dusick AM, et al. Neurodevelopmental and functional outcomes of extremely low birth weight infants in the National Institutes of Child Health and Human Development Neonatal Research Network 1993–1994. Pediatrics 2000;105:1216–26. [PubMed: 10835060]
6. Klebanov PK, Brooks-Gunn J, McCormick MC. School achievement and failure in very low birth weight children. J Dev Behav Pediatr 1994;15:248–56. [PubMed: 7798370]
7. Hack M, Taylor HG, Klein N, Eiben R, Schatschneider C, Mercuri-Minich N. School-age outcomes in children with birth weights under 750g. N Engl J Med 1994;331:753–59. [PubMed: 7520533]
8. Hack M, Taylor HG, Klein N, Mercuri-Minich N. Functional limitations and special health care needs of 10- to 14-year-old children weighing less than 750 grams at birth. Pediatrics 2000;106:554–60. [PubMed: 10969102]
9. Hack M. Consideration of the use of health status, functional outcome, and quality-of-life to monitor neonatal intensive care practice. Pediatrics 1999;103 (Suppl E):319–28. [PubMed: 9917474]
10. Msall ME, DiGaudio K, Duffy LC, La-Forest S, Braun S, Granger CV. WeeFIM. Normative sample of an instrument for tracking functional independence in children. Clin Pediatr (Phila) 1994;33:431–38. [PubMed: 7955782]
11. Saigal S, Stoskopf BL, Rosenbaum PL, Hoult LA, Furlong WJ, Feeny DH. Development of a multi-attribute preschool health status classification system. (Poster presented at the American Pediatric Society and Society for Pediatric Research meeting, 1998). Pediatr Res 1998;43(Suppl): 228A.
12. Scholle SH, Whiteside L, Kelleher K, Bradley R, Casey P. Health status of preterm low-birth-weight infants. Comparison of maternal reports. Arch Pediatr Adolesc Med 1995;149:1351–57. [PubMed: 7489073]

13. National Institute of Child Health and Human Development, et al. Follow-up care of high-risk infants. Pediatrics 2004;114(Suppl):1377–97.
14. Stein REK, Jessop DJ. Functional status II (R). A measure of child health status. Med Care 1990;28:1041–55. [PubMed: 2250491]
15. Alexander GR, Himes JH, Kaufman RB, Mor J, Kogan M. A United States national reference for fetal growth. Obstet Gynecol 1996;87:163–68. [PubMed: 8559516]
16. Papile LA, Burstein J, Burstein R, Koffler H. Incidence and evolution of subependymal and intraventricular hemorrhage: a study of infants with birth weights less than 1,500 gm. J Pediatr 1978;92:529–34. [PubMed: 305471]
17. Bayley, N. Bayley Scales of Infant Development. 2. San Antonio, Texas: Psychological Corp; 1993.
18. Stein, REK.; Jessop, JD. Manual for the Functional Status II (R) measure. Bronx, New York: Albert Einstein College of Medicine; 1991.
19. Muthén, LK.; Muthén, B. Mplus users guide. 3. Los Angeles, CA: Muthén &Muthén; 2004.
20. McCabe MA, Granger CV. Content validity of a pediatric functional independence measure. Appl Nurs Res 1990;3:120–22. [PubMed: 2400209]
21. Bruijnzeels MA, van der Wouden JC, Foets M, Prins A, van den Heuvel WJA. Validity and accuracy of interview and diary data on children's medical utilisation in The Netherlands. J Epidemiol Community Health 1998;52:65–69. [PubMed: 9604044]
22. The Infant Health and Development Program. Enhancing the outcomes of low-birthweight, premature infants. A multisite, randomized trial. JAMA 1990;263:3035–42. [PubMed: 2188023]
23. Post MW, Kuyvenhoven MM, Verheij MJ, de Melker RA, Hoes AW. The Dutch version of Functional Status II(R): a questionnaire measuring the functional health status of children. Ned Tijdschr Geneeskd 1998;142:2675–79. [PubMed: 10065223]
24. Msall ME, Rogers BT, Buck GM, Mallen S, Catanzaro NL, Duffy LC. Functional status of extremely preterm infants at kindergarten entry. Dev Med Child Neurol 1993;35:312–20. [PubMed: 8335146]
25. Hayley, SM.; Costner, WJ.; Ludlow, LH.; Haltiwanger, JT.; Andrellos, PJ. Pediatric Evaluation of Disability Index. Nashville, TN: Ellsworth and Vandermeer Press Ltd; 1997.

THE FOLLOWING INSTITUTIONS AND INDIVIDUALS PARTICIPATED IN THE STUDY

Brown University: William Oh MD (principle investigator [PI]), Betty Vohr MD (follow-up [FU] PI), Angelita Hensman RNC (Network Coordinator [NC]), Lucy Noel RNC (FU Coordinator [FC]); Case Western Reserve University: Avroy A Fanaroff MB BCh (PI), Dee Wilson MD (FU PI), Nancy Newman RN (NC), Bonnie Siner RN (FC); Emory University: Barbara J Stoll MD (PI, FU PI), Ellen Hale RNC BS (NC, FC); Harvard University: Ann R Stark MD (PI, FU PI), Kerri Fournier RN (NC); Indiana University: James A Lemons MD (PI), Anna Dusick MD (FU PI), DeeDee Appel RN (NC), Leslie Richards RN (FC); Stanford University: David K Stevenson MD (PI), Susan Hintz MD (FU PI), Bethany Ball RN BS (NC, FC); University of Alabama: Waldemar A Carlo MD (PI), Kathleen Nelson MD (FU PI), Monica Collins RN (NC); Vivien Phillips (FC); University of Cincinnati: Edward F Donovan MD (PI), Jean Steichen MD (FU PI), Marcia Mersmann RN (NC), Cathy Grisby RN (NC), Tari Gratton RN (FC); University of Miami: Shahnaz Duara MD (PI), Charles Bauer MD (FU PI), Ruth Everett RN (NC), Mary Allison RN (FC); University of New Mexico: Lu-Ann Papile MD (PI, FU PI), Conra Backstrom RN (NC); University of Tennessee: Sheldon B Korones MD (PI), Kimberly Yolton PhD (FU PI), Tina Hudson RN (NC); University of Texas–Dallas: Abbot R Laptook MD (PI), Roy Heyne MD (FU PI), Susie Madison RN (NC), Jackie Hickman RN (FC); University of Texas–Houston: Jon E Tyson MD MPH (PI), Brenda Morris MD (FU PI), Georgia McDavid RN (NC), Shannon Rossi (FC); Wayne State University: Seetha Shankaran MD (PI), Virginia Delaney-Black MD MPH (FU PI), Gerry Muran BSN (NC); Debbie Kennedy RN (FC); Yale University:

Richard A Ehrenkranz MD (PI), Linda Mayes MD (FU PI), Pat Gettner RN (NC), Elaine Romano MSN (FC); National Institute of Child Health and Human Development: Linda L Wright MD, Rosemary D Higgins MD (PIs), Beth M McClure MS (NC); George Washington University (data coordinating center 1993–1997): Joel Verter PhD; Research Triangle Institute (data coordinating center from 1998): W Kenneth Poole PhD (PI, FU PI), Betty Hastings (NC). Steering committee chairman: Alan H Jobe MD PhD.

Table I

Neonatal characteristics of the cohorts

	Children born 1993–1994 (exploratory cohort)			Children born 1995–2000 (confirmatory cohort)			
Characteristic[a]	***n***	**%**	**95% CI**	***n***	**%**	**95% CI**	***p*** [b]
Participants	1080	100		4022	100		
Birthweight, g							
401–750	396	37	34–40	1521	38	36–39	0.5
751–1000	684	63	60–66	2501	62	61–64	
Gestational age, wks							
<25	209	19	17–22	758	19	18–20	1.0
25–28	725	67	64–70	2742	68	67–70	
29+	146	14	12–16	521	13	12–14	
Small for gestational age	199	18	16–21	720	18	17–19	0.7
Sex							
Male	508	47	44–50	1864	46	45–48	0.7
Female	572	53	50–56	2158	54	52–55	
Race/ethnicity							
Black	551	51	48–54	1800	45	43–46	<0.001
White	381	35	32–38	1576	39	38–41	
Hispanic	132	12	10–14	522	13	12–14	
Other	16	1	1–2	120	3	3–4	
Respiratory distress syndrome	744	69	66–72	2535	63	62–65	<0.001
Surfactant replacement	760	70	68–73	3088	77	75–78	<0.001
Bronchopulmonary dysplasia	424	40	37–43	1762	44	42–46	0.01
Grade 3 or 4 intraventricular hemorrhage	174	16	14–19	517	13	12–14	0.01
Duration of initial hospital stay, d (mean)[c]	101		99–104	97[d]		96–99	0.01

[a] Information was missing for gestational age (*n*=1), small for gestational age (*n*=1), child's race (*n*=4), respiratory distress syndrome (*n*=11), surfactant use (*n*=1), bronchopulmonary dysplasia (*n*=24), intraventricular hemorrhage (*n*=36), and length of stay (*n*=4).

[b] *p* values for comparisons between the groups by Fisher's exact or χ^2 test for categorical variables and the Wilcoxon test for length of stay.

[c] Medians for length of initial hospital stay were exploratory 93 days and confirmatory 90 days.

$d_{n=4018}$. CI, confidence intervals.

Table II

Sociodemographic characteristics of the cohorts

	Children born 1993–1994 (exploratory cohort)			Children born 1995–2000 (confirmatory cohort)			
Characteristic[a]	***n***	**%**	**95% CI**	***n***	**%**	**(95% CI)**	***p***[b]
Overall	1080	100		4022	100		
Mother's age							
≤19y	193	18	16–20	695	17	16–18	0.7
20 y+	887	82	80–84	3324	83	82–84	
Primary caregiver at 18–22mo							
Mother	971	90	88–92	3695	92	91–93	0.04
Other female	109	10	8–12	327	8	7–9	
Caregiver's education							
<12y	302	28	25–31	981	24	23–26	0.02
High-school degree+	778	72	69–75	3036	76	74–77	
Total income in child's household							
<US$20 000	574	56	53–59	1659	44	42–45	<0.001
US$20 000 +	444	44	41–47	2129	56	55–58	
Child's medical insurance							
Medicaid	674	63	60–65	2354	59	57–60	<0.001
Private	360	33	31–36	1552	39	37–40	
None	42	4	3–5	102	3	2–3	

[a] Information was missing for mother's age (n=3), caregiver's education (n=5), household income (n=296), and child's medical insurance (n=18).

[b] p values for comparisons between the groups by Fisher's exact or χ^2 test. CI, confidence intervals.

Table III

Mean scale scores by BSID-II score and medical conditions

	Participants (n)	General health mean (SD)	Independence mean (SD)
BSID-II mental developmental index			
<70	1517	88.2 (13.9)	74.3 (24.5)
≥70	3183	93.4 (10.8)	91.9 (12.1)
BSID-II psychomotor developmental index			
<70	1079	88.6 (13.1)	69.9 (25.6)
≥70	3564	92.7 (11.7)	91.1 (13.0)
Blind in both eyes			
Yes	58	85.7 (11.5)	46.6 (29.0)
No	4897	91.7 (12.2)	86.6 (18.4)
Hearing impaired in both ears			
Yes	84	88.2 (12.4)	66.6 (27.0)
No	4828	91.7 (12.2)	86.6 (18.6)
Cerebral palsy			
Yes	740	88.1 (13.3)	66.5 (27.7)
No	4248	92.2 (11.9)	89.5 (14.7)
Grade 3 or 4 intraventricular hemorrhage			
Yes	691	90.5 (11.9)	78.2 (24.5)
No	4375	91.8 (12.2)	87.4 (17.8)
Neurodevelopmental impairment			
Yes	1961	89.0 (13.4)	76.2 (23.6)
No	2697	93.6 (10.8)	92.8 (11.3)

All group comparisons for both scales were significant with $p<0.001$. Neurodevelopmental Impairment was defined as one or more of the following: mental or psychomotor developmental index less than 70, cerebral palsy, blind (legally blind as reported by parents) in both eyes, or hearing impaired (requiring hearing aids in both ears). BSID-II, Bayley Scales of Infant Development, second edition.

Table IV

Scale score norms by neonatal characteristics

	General health			Independence		
		<1SD below mean	<2SD below mean		<1SD below mean	<2SD below mean
Variable	**Mean (SD)**	***n* (%)**	***n* (%)**	**Mean (SD)**	***n* (%)**	***n* (%)**
All participants	91.6 (12.2)	682 (13)	323 (6)	86.1 (19.1)	856 (17)	260 (5)
Sex						
Male	90.9 (12.3)[c]	351 (15)	161 (7)	84.4 (20.1)[c]	478 (20)	144 (6)
Female	92.2 (12.1)	331 (12)	162 (6)	87.6 (18.1)	378 (14)	116 (4)
Birthweight, g						
401–750	91.2 (12.2)[d]	276 (14)	122 (6)	83.1 (20.6)[c]	389 (20)	127 (7)
751–1000	91.9 (12.2)	406 (13)	201 (6)	87.9 (18.0)	467 (15)	133 (4)
Gestational age,[a] wks						
<25	90.6 (12.2)[c]	152 (16)	61 (6)	81.7 (22.1)[c]	227 (23)	82 (8)
25–28	91.6 (12.2)	461 (13)	228 (7)	86.6 (18.6)	549 (16)	158 (5)
≥29	93.1 (11.9)	69 (10)	34 (5)	90.0 (16.2)	79 (12)	20 (3)
Race[b]						
Black	90.0 (13.0)[c]	412 (18)	193 (8)	84.5 (19.6)[c]	457 (19)	133 (6)
White	93.8 (10.4)	166 (8)	73 (4)	87.7 (18.8)	284 (15)	95 (5)
Hispanic	91.2 (13.1)	85 (13)	46 (7)	86.5 (18.5)	99 (15)	27 (4)
Other	91.3 (12.0)	19 (14)	11 (8)	88.7 (17.7)	16 (12)	5 (4)

[a] Gestational age was missing for one child. Post-hoc comparisons indicated significant differences between children with gestational age <25wks vs ≥29wks and between gestational age 25–28wks and ≥29wks on the general health scale and between all levels for the Independence scale ($p<0.05$).

[b] Race information was missing for four children. Post-hoc comparisons indicated significant differences between white and black children for both scales and differences between white and Hispanic children for the general health scale ($p<0.05$).

[c] $p<0.001$.

[d] $p<0.01$.