



Published in final edited form as:

J Perinatol. 2015 July ; 35(7): 516–521. doi:10.1038/jp.2014.243.

Do currently recommended Bayley III cutoffs overestimate motor impairment in infants born <27 weeks gestation?

Andrea F. Duncan, MD, MSCR¹, Carla Bann, PhD², Cathy Boatman, MS³, Susan R. Hintz, MD, MS Epi⁴, Yvonne E. Vaucher, MD, MPH⁵, Betty R. Vohr, MD⁶, Kimberly Yolton, PhD⁷, Roy J. Heyne, MD⁸, and the Eunice Kennedy Shriver National Institute of Child Health and Human Development Neonatal Research Network*

¹ Department of Pediatrics, Division of Neonatology, University of Texas Health Science Center at Houston, Houston, TX, USA.

² Statistics and Epidemiology Unit, RTI International, Research Triangle Park, NC, USA.

³ Children's Medical Center, Dallas, TX, USA.

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

⁵ Department of Pediatrics, University of California San Diego, La Jolla, CA, USA.

⁶ Department of Pediatrics, Women & Infants' Hospital, Brown University, Providence, RI, USA.

⁷ Department of Pediatrics, Children's Hospital Medical Center and University of Cincinnati, Cincinnati, OH, USA.

⁸ Department of Pediatrics, University of Texas Southwestern Medical Center at Dallas, Dallas, TX, USA.

Abstract

Objective—To determine whether a Bayley-III Motor Composite score of 85 may overestimate moderate-severe motor impairment by analyzing Bayley-III motor components and developing cut-point scores for each.

Study Design—Retrospective study of 1183 children born <27 weeks gestation at NICHD Neonatal Research Network centers and evaluated at 18-22 months corrected age. Gross Motor Function Classification System determined gross motor impairment. Statistical analyses included linear and logistic regression and sensitivity/specificity.

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:http://www.nature.com/authors/editorial_policies/license.html#terms

Corresponding Author: Andrea F. Duncan, MD Address: Department of Pediatrics Division of Neonatology University of Texas Health Science Center at Houston 6431 Fannin St, MSB 3.242 Phone: 713.500.5455 Fax: 713.500.5794

Andrea.F.Duncan@uth.tmc.edu.

*A list of members of the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development Neonatal Research Network is available in the Supplementary Information.

Conflicts of Interest Statement:

The authors have no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Supplementary Information is available at the Journal of Perinatology's website.

The authors declare no conflicts of interest.

Results—Bayley-III Motor Composite scores were strong indicators of gross/fine motor impairment. A Motor Composite cut-point of 73 markedly improved specificity for identifying gross and/or fine motor impairment (94% compared with a specificity of 76% for the proposed new cut point of 85). A Fine Motor Scaled Score <3 differentiated mild from moderate-severe fine motor impairment.

Conclusions—This study indicates that a Bayley-III Motor Composite Score of 85 may overestimate impairment. Further studies are needed employing term controls and longer follow-up.

Introduction

Motor function is an important component of neurodevelopmental assessment and is included in assessment tools such as the Bayley Scales of Infant (and Toddler) Development (Bayley).^{1,2} Motor function is also used in classification of neurodevelopmental outcomes for pre-school children in research studies, such as those reported by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development Neonatal Research Network (NRN).

Prior to 2006, the 2nd Edition of the Bayley (Bayley-II) Psychomotor Developmental Index (PDI)¹ was the standard tool for determining motor outcomes among high-risk preschool children, and a PDI <70 (>2 SD below the test mean) indicated moderate-severe motor impairment for NRN research. The Gross Motor Function Classification System (GMFCS)³ was also used by the NRN to assess the level of motor impairment in children found to have cerebral palsy (CP). A GMFCS level 2 was used as the cutoff for moderate-severe CP. The NRN definition of Neurodevelopmental Impairment (NDI) included a PDI<70 and/or CP with a GMFCS 2 along with one or more of the following: a Mental Developmental Index (the Bayley-II measure of cognition and language; MDI) <70, deafness, or blindness.⁴

In 2006, the 3rd edition of the Bayley (Bayley-III)² was introduced. The Bayley-III replaced the PDI with the Motor Composite score, which maintains a mean of 100 (SD 15), but now includes separate fine and gross motor subscale tests. When the NRN began using the Bayley-III Motor Composite for NRN outcome studies in January 2010, it replaced the PDI <70 criteria for NDI with a Motor Composite score <70. Other research groups did the same, but subsequently found decreased rates of motor impairment, as first reported by Anderson et al.⁵ More recently, the NRN reported that, during the period of 2008-2011, a Bayley-III Motor Composite score <70 was 40% less likely to categorize a child with NDI than a Bayley-II PDI <70 during 2006-2007.⁴ Moreover, the mean Motor Composite score was 5 points higher than the PDI, which is similar to the increases reported by other studies,^{2,6-8} all of which suggested that a Bayley-III Motor Composite threshold of <70 may under-identify impairment. Similarly, the Bayley-III Cognitive Composite score has been found to identify fewer cognitive deficits than the Bayley-II MDI.^{4,7,9,10} Though several studies have proposed alternate Bayley-III *Cognitive* Composite cutoffs for impairment,^{4,9,10} none have explicitly proposed an alternate threshold for the *Motor* Composite or analyzed the relative contributions of the gross and fine motor subtests.

In response to the concern for under-identification of impairment using a Motor Composite score of <70, the NRN has begun to consider utilizing an alternate score to identify impairment. Alternate scores from 75 to 85 have been proposed. However, prior to enacting such change, further study was necessary. The primary objectives of this study were thus to 1) determine whether a Bayley-III Motor Composite score of 85 may overestimate gross and/or fine motor impairment, 2) analyze the relative contributions of the gross and fine motor subtests and 3) examine the impact of various theoretical Motor Composite score cut-points on NDI rates.

Subjects and Methods

The study sample included 1,183 children born <27 weeks gestational age who received 18-22 month follow-up at any of the 16 NRN centers between January 1, 2010 and December 31, 2011 (Figure 1). This study was approved by the Institutional Review Boards at all centers, and informed consent was obtained.

Neurodevelopmental Assessments and Classifications

The Bayley-III was administered per NRN protocol by experienced examiners, who were annually certified by 1 of 4 gold-standard psychologists (as described in previous studies),⁴ and scored as per the Bayley-III technical manual.² This standardized scoring method includes determining raw scores for the fine and gross motor subtests, each of which are then converted into scaled scores ranging from 1 (low skill) to 19 (high), with a mean of 10 (SD 3). The fine and gross motor scaled scores are then added together and converted into a Motor Composite score (mean of 100 (SD 15)), whereby a score <70 is 2 SD below the mean. In order to determine whether a Motor Composite score of 85 overestimates motor impairment, benchmark criteria for gross and fine motor impairment were necessary.

Gross Motor

A standard neurological exam was performed by certified examiners, trained to reliability per NRN protocol^{4,11} and gross motor function was determined using standard GMFCS levels defined for children <2 years of age.⁴ The neurologic examination includes assessment of tone, strength, reflexes, joint angles, and posture.⁴ The GMFCS levels reflect progressively more severe limitations in movement and truncal control, and have been defined by Palisano³ as follows: Level 1 cannot walk independently, but can sit with hands free, creep/crawl, pull to stand, cruise; level 2 cannot sit without self-support, gets to sitting, can creep/crawl; level 3 requires external support for lower trunk to sit; and levels 4-5 have progressively more severe limitations. The NRN has qualitatively defined severity of motor impairment based on these GMFCS levels as follows: Level 1 mild impairment and Levels 2 as moderate to severe impairment.⁴ Although the GMFCS was originally designed to classify motor function in children with CP³, the scale has been shown to reliably rate motor performance in children with Down syndrome¹², and has also been used as an outcome measure (independent of a diagnosis of CP) in a variety of neurodevelopmental follow-up studies, including several from the NRN¹³⁻¹⁵ and others¹⁶⁻¹⁸ as well. The benchmark criterion for gross motor impairment in this study was defined as a GMFCS level 2 per NRN protocol³ and other studies.¹⁹

Fine Motor

While the GMFCS has been shown to reliably determine gross motor impairment,^{3,12} no equivalent test has been published to assist in defining a benchmark for fine motor impairment in children <2 years of age. In order to define this benchmark for this study, expert consensus was utilized. Four sets of Bayley-III fine motor test items corresponding to an FMss at the lowest end of the score-able range (FMss of 1 to 4) were identified and compiled. The skill sets were chosen from the 19.5 to 20.5 month age bracket, since that was the middle of the NRN testing window of 18-22 months adjusted age. All four skill set descriptions (Table 3) were submitted to 17 NRN-certified neurological examiners and site-based NRN-site based occupational and physical therapists, who independently, and masked to FMss scores, rated the impairment level of each item set, using a rating scale similar to the one used to define GMFCS levels, namely: level 0- none or minimal, level 1- mild, level 2- moderate, and 3- severe. In order to determine whether the use of expert opinion was sound, we performed a limited validation of the method, utilizing the percent decrement method defined by the WHO International Classification of Functioning, Disability and Health (ICF) and the International Classification of Functioning, Disability and Health: Children & Youth Version (ICF-CY).¹⁹ This method uses percent decrements from normal functioning to define degrees of impairment, in lieu of calculating scores like the Bayley does for specific functional sections; and the tool has been used extensively in children worldwide and found to be valid.¹⁹ According to the ICF-CY, a 5-24% decrement signifies mild impairment, 25-49% moderate impairment, and 50-95% severe. This was applied to the raw scores corresponding to each of the four Bayley fine motor skill set levels noted above, by calculating the percent decrement in raw score from the mean raw score for each of the FMss.

Statistical Analyses

To compare the estimation of impairment obtained using the proposed Motor Composite cut-point of 85 with alternative cut-points, we conducted a series of analyses to (1) identify the FMss cut-point which differentiated mild from moderate to severe fine motor impairment; (2) determine the percentage of variance in the Motor Composite score accounted for by the Bayley-III Gross and Fine Motor Scaled Scores (GMss and FMss); (3) examine the relationship between Motor Composite scores and measures of gross and/or fine motor impairment; and (4) assess how various cut-points on the Motor Composite score would affect NDI rates.

To determine the FMss cut-point for moderate to severe impairment, percent distributions of expert ratings for each FMss were calculated. Percent decrement in Bayley-III fine motor raw scores were calculated for each FMss as follows: (raw score for mean- raw score for given FMss) ÷ (raw score for mean). Linear regression models were used to identify the unique contribution of Bayley-III GMss and FMss to Motor Composite scores (i.e. percentage of variance accounted for), controlling for covariates previously shown to adversely affect neurodevelopmental outcomes.⁴

Receiver operating characteristic (ROC) curve analyses and chi-square tests were conducted to determine the relationship between Bayley-III motor composite scores and indicators of

moderate/severe impairment in (a) gross motor skills (defined as GMFCS level ≥ 2), (b) fine motor skills (defined above), or (c) either gross and/or fine motor skills. We then estimated the sensitivity, specificity, and positive and negative predictive values of various Motor Composite score cut-points as indicators of moderate/severe gross and/or fine motor impairment (defined as above), and repeated the analyses for subsets with diagnoses of normal, CP and suspect/abnormal non-CP. Finally, we computed the rates of NDI in our sample for alternative Motor Composite score cut-points, using two different Cognitive Composite score cut-points (<70 vs. <85), while keeping the other criteria for the NRN definition consistent (i.e., bilateral deafness, bilateral blindness; moderate-severe CP). All analyses were conducted using SAS version 9.3.

Results

Of 2303 children eligible for follow-up, the final sample of 1183 excluded 919 deaths, 99 lost, and 102 with missing data (Figure 1). Demographic characteristics of subjects are described in Table 1.

Analysis of 17 expert ratings of the four fine motor skill sets corresponding to FMss of 1-4 indicated that 100% rated FMss of 3-4 as normal/minimal or mild (thus none as moderate-severe), whereas 47% rated FMss=2 as moderate to severe, and all rated FMss=1 moderate-severe. Thus, an FMss cut-point of <3 conservatively differentiated between mild and moderate-severe impairment ($\chi^2(1)=39.54$, $p < 0.001$). Percent decrement in Bayley-III raw fine motor scores for the range of FMss of interest was as follows: FMss=1: 31% (moderate), FMss=2: 26% (moderate), FMss=3: 23% (mild), and FMss=4: 20% (mild), which supports a mild vs. moderate cut-point of <3 using ICF-CY criteria.

The GMss accounted for 48% of variance in the Motor Composite and FMss for 50% after adjusting for control variables. Motor Composite scores were strong indicators of gross motor impairment (GMFCS level ≥ 2) with an area under the ROC curve (AUC) of 0.98, fine motor impairment (Bayley-III FMss <3) (AUC=0.99), and either type of impairment (AUC=0.98). Over 90% of children with a GMFCS ≥ 2 , FMss <3 , or composite of the two had a Motor Composite <73 ; and motor scores were significantly associated with each of these outcomes ($p < 0.001$ for each comparison; see Supplemental Figure 2).

Table 2 shows the sensitivity, specificity, and positive and negative predictive values of Motor Composite scores ranging from 70-85 (in Bayley scaled increments of 3) for identifying a GMFCS ≥ 2 and/or FMss <3 . A cut-point of 73 maximized the sum of sensitivity and specificity for all children; this finding remained when children with CP were removed. This cut-point also provided high sensitivity and specificity for identifying subjects with normal neurological findings (sensitivity=100%, specificity=98%) and those with abnormal or suspect function not due to cerebral palsy (sensitivity and specificity=88%). For those with CP, sensitivity was high (96%), though specificity was lower (65%), but only a cut-point of 70 had a higher specificity (Supplementary Tables 4-6). Though a cut-point of 70 had a slightly higher sum of sensitivity and specificity for identifying fine motor impairment alone, an optimal cut-point of 73 for gross motor impairment more than offset this, accounting for the optimal cut-point of 73 for the composite of gross/fine motor impairment

(Table 2). This table also shows that a cut-point of 73 has much better positive predictive value than a cut-point of 85.

The percentage of children categorized as having NDI, using a Motor Composite cutpoint of 73 determined by this study, along with a cognitive cut-point of 85 proposed by other studies (leaving the other standard criteria the same-- moderate to severe CP, blindness, and/or deafness), is 33%.

Discussion

Testing with the Bayley-III yields higher cognitive and motor scores than those obtained with the Bayley-II, despite use in similar populations. This finding has resulted in dilemmas regarding what score should be considered an appropriate threshold for determining impairment. The higher observed cognitive scores have been the focus of several recently published studies,⁴⁻¹⁰ many of which have attempted to define alternate Bayley-III cognitive score cut-points that would produce similar rates of NDI to those observed based on Bayley-II.^{4,7, 9,10} However, there has been limited focus on the motor discrepancies, and limited proposals and rationales for alternate motor score thresholds for impairment.

A study of extremely preterm children in Australia used a contemporary normal term control group to determine an alternate 2 SD cutoff of 85 based on the control Motor Composite mean of 118.4 (SD 16.7).⁵ Another study of extremely preterm infants found that a Bayley-III Motor Composite score of 94 at 2 years of age was the best predictor of impairment on the Movement ABC-2 test of motor function at 4 years of age.²⁰

The current study has developed a novel method for determining a meaningful cut-point for the Bayley-III Motor Composite score using specific reference points for gross and fine motor components. In contrast to other studies, we have shown that a Motor Composite cut-point of 73 had 95% sensitivity and 94% specificity for identifying those with motor impairment based on a GMFCS ≥ 2 and/or FMss < 3 . Sensitivity is improved over that found with a Motor Composite cut-point of 70 and specificity over that found with cut-point of 85. Moreover, this motor cutoff yields an NDI rate of 33%; though this is not as high as the 42% obtained by raising both motor and cognitive cut-points to 85, it is less likely to overstate motor impairment.

Despite efforts to determine the ideal cut-point for BSID-III scores, there remains ambiguity regarding which version of the Bayley is actually more accurate, as Aylward has recently pointed out.²¹ This is related to the unclear impact of a spectrum of changes in test format, content and scoring, and of the expansion of the normative sample to include 9.8% children at risk for developmental delay. Though the format of the Motor Composite preserved the inclusion of both gross and fine motor items, many changes were made, as detailed in the Bayley-III Technical Manual,^{2,22} including separate rather than integrated gross and fine motor subtests, deletion of 22 of 111 Bayley-II motor items, modification of 65 of 89 remaining items (53 gross motor and 13 fine motor), transfer of 25 items from the Bayley-II MDI to the Bayley-III fine motor subtest (including 14 with modifications) and addition of 18 new fine motor and 4 new gross motor items. The exact impact of these changes cannot

be determined, but they likely contribute to observed shift in overall score, as do concomitant changes in the order of administration and in rules for determining start points, floors, ceilings and credit for individual items.

Authors of studies referenced above⁴⁻¹⁰ have also speculated about the impact of these and other factors, as have various commentators.²³⁻²⁷ However, it remains unclear how these various factors affect scores, how accurately the Bayley-III measures performance or estimates impairment and whether or how much thresholds should be adjusted.

While our study sheds further light on this matter, it has several limitations. First, the study is retrospective. Secondly, the lack of a concurrent term control group limits the ability to determine whether the Motor Composite over or underestimates motor impairment compared to a normal population.

Thirdly, the use of expert opinion to define fine motor impairment does not employ an independent, validated tool. However, this method was utilized in an attempt to overcome the lack of any published data establishing a fine motor impairment benchmark for comparison and was deemed appropriate, as the FMss is a component of the Motor Composite score and so had to be taken into consideration in the analyses. Though several fine motor tools have been published for 4-18 year-old children, they have not been standardized for younger children.²⁸⁻²⁹ Plasschaert et al³⁰ studied the inter-observer reliability of the Manual Ability Classification System (MACS) in children between 1 and 5 years of age and found a kappa of 0.55 for children under two, which they attributed in part to difficulty applying a classification system developed for an older age group to children under two; accordingly, they recommended caution in using the tool for younger children. In the case of our study, the cut-point determined by expert consensus was further supported by application of the ICF-CY rating system to FMss. Validity of the expert opinion method utilized could ostensibly be assessed in future studies by determining whether the results obtained at this early age predict/correlate with fine motor function later in childhood. Finally, use of the GMFCS at the young age of 18-22 months may or may not be indicative of later function and thus may be a limitation. However, this is the standard for long-term neurodevelopmental assessment of preterm children at this time.

Strengths of the study are the large sample size, the inter-rater reliability for neurodevelopmental examiners, and the in-depth comprehensive analyses performed employing available reference tools.

In conclusion, a Bayley-III Motor Composite cut-point of 73 has improved sensitivity and specificity over other cut-points, and is less likely to overstate impairment. We would thus suggest that a Motor Composite cut-point of 73 would be a more appropriate motor criteria for determining NDI than a score of 85, at least for extremely preterm infants. Longitudinal studies and studies including a comparative normal term-born sample are now needed to validate these results.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The National Institutes of Health, the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD), the National Center for Research Resources, and the National Center for Advancing Translational Sciences provided grant support for the Neonatal Research Network's Generic Database and Follow-up Studies. The content of the publication is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

REFERENCES

1. Bayley, N. Bayley Scales of Infant Development. 2nd Edition.. The Psychological Corporation; San Antonio: 1993.
2. Bayley, N. (BSIDIII): Technical Manual. 3rd Edition. Harcourt Assessment/Psychological Corporation; San Antonio: 2006. Bayley Scales of Infant and Toddler Development.
3. Palisano R, Rosenbaum P, Walter S, Russell D, Wood E, Galuppi B. Development and reliability of a system to classify gross motor function in children with cerebral palsy. *Dev Med Child Neurol*. 1997; 39:214–23. [PubMed: 9183258]
4. Vohr BR, Stephens BE, Higgins RD, et al. Eunice Kennedy Shriver National Institute of Child Health and Human Development Neonatal Research Network. Are outcomes of extremely preterm infants improving? Impact of Bayley assessment on outcomes. *J Pediatr*. 2012; 161:222–8. [PubMed: 22421261]
5. Anderson PJ, De Luca CR, Hutchinson E, Roberts G, Doyle LW. Victorian Infant Collaborative Group. Underestimation of developmental delay by the new Bayley-III Scale. *Arch Pediatr Adolesc Med*. 2010; 164:352–6. [PubMed: 20368488]
6. Acton BV, Biggs WS, Creighton DE, et al. Overestimating neurodevelopment using the Bayley-III after early complex cardiac surgery. *Pediatrics*. 2011; 128:e794–800. [PubMed: 21949148]
7. Long SH, Galea MP, Eldridge BJ, Harris SR. Performance of 2-year-old children after early surgery for congenital heart disease on the Bayley Scales of Infant and Toddler Development. *Early Hum Dev (Third Edition)*. 2012; 88:603–7. [PubMed: 22336496]
8. Skjöld B, Vollmer B, Böhm B, Halberg B, et al. Neonatal magnetic resonance imaging and outcome at age 30 months in extremely preterm infants. *J Pediatr*. 2012; 160:559–566. [PubMed: 22056283]
9. 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 preterm children. *J Pediatr*. 2012; 160:553–8. [PubMed: 22048046]
10. Lowe JR, Erickson SJ, Schrader R, Duncan AF. Comparison of the Bayley II Mental Developmental Index and the Bayley III Cognitive Scale: are we measuring the same thing? *Acta Paediatr*. 2012; 101:e55–8. [PubMed: 22054168]
11. Newman JE, Bann CM, Vohr BR, Dusick AM, Higgins RD. Follow-Up Study Group of Eunice Kennedy Shriver National Institute of Child Health and Human Development Neonatal Research Network. Improving the Neonatal Research Network annual certification for neurologic examination of the 18-22 month child. *J Pediatr*. 2012; 161:1041–6. [PubMed: 22748517]
12. Bodkin A, Robinson C, Perales F. Reliability and Validity of the Gross Motor Function Classification System for Cerebral Palsy. *Pediatr Phys Ther*. 2003; 15:247–252. [PubMed: 17057460]
13. Peralta-Carcelen M, Moses M, Adams-Chapman I, Gantz M, Vohr BR. Stability of neuromotor outcomes at 18 and 30 months of age after extremely low birth weight status. *Pediatrics*. 2009; 123:e887–e895. [PubMed: 19403482]
14. Pappas A, Kendrick DE, Shankaran S, Stoll BJ, Bell EF, Laptook AR, et al. Chorioamnionitis and early childhood outcomes among extremely low gestational age neonates. *JAMA Pediatr*. 2014; 168:137–147. [PubMed: 24378638]
15. Chapman-Adams I, Bann CM, Vaucher YE, Stoll BJ. Association between feeding difficulties and language delay in preterm infants using Bayley Scales of Infant Development-Third Edition. *J Pediatr*. 2013; 163:680–5. [PubMed: 23582139]

16. Schmidt B, Anderson PJ, Doyle LW, Dewey D, Grunau RE, Asztalos EV, et al. Survival without disability to age 5 years after neonatal caffeine therapy for apnea of prematurity. *JAMA*. 2012; 307:275–282. [PubMed: 22253394]
17. Azzopardi DV, Strohm B, Edwards AD, Dyet L, Halliday HH, Juszczak E, et al. Moderate hypothermia to treat perianal asphyxia encephalopathy. *NEJM*. 2009; 361:1349–1358. [PubMed: 19797281]
18. Kim CT, Han J, Heakyung K. Pediatric stroke recovery: a descriptive analysis. *Arch Phys Med Rehabil*. 2009; 90:657–662. [PubMed: 19345783]
19. Boyd LAC, Msall ME, O'Shea TM, Allred EN, Hounshell G, Leviton A. Social-emotional delays at 2 years in extremely low gestational age survivors: Correlates of impaired orientation/engagement and emotional regulation. *Early Hum Dev*. 2013; 89:925–930. [PubMed: 24144915]
20. WHO International Classification of Functioning. Disability and Health: Children & Youth Version (ICF-CY). WHO; Geneva: 2001.
21. Spittle AJ, Spencer-Smith MM, Eeles AL, Lee KJ, Lorefice LE, Anderson PJ, et al. Does the Bayley-III Motor Scale at 2 years predict motor outcome at 4 years in very preterm children? *Dev Med Child Neurol*. 2013; 55:448–52. [PubMed: 23216518]
22. Aylward GP. Continuing issues with the Bayley-III: Where to go from here. *J Dev Behav Pediatr*. 2013; 34:697–701. [PubMed: 24217027]
23. Bayley III Technical Report 2: Factors contributing to differences between Bayley-III and BSID-II scores. Harcourt Assessment; San Antonio: 2007.
24. Hack M. Dilemmas in the measurement of developmental outcomes of preterm children. *J Pediatr*. 2012; 160:537–38. [PubMed: 22188782]
25. Msall ME. Measuring outcomes after extreme prematurity with the Bayley-III Scales of infant and toddler development: a cautionary tale from Australia. *Arch Pediatr Adolesc Med*. 2010; 164(4): 391–3. [PubMed: 20368495]
26. Msall ME. The Bayley-III scale underestimates developmental delay in extremely premature and extremely low birth weight infants. *J Pediatr*. 2010; 157(5):863–4. [PubMed: 20955860]
27. Msall ME. Overestimating neuroprotection in congenital heart disease: problems with Bayley III Outcomes. *Pediatrics*. 2011; 128:e993–e994. [PubMed: 21949147]
28. Aylward GP, Sylward BS. The changing yardstick in measurement of cognitive abilities in infancy. *J Dev Behav Pediatr*. 2011; 32:465–468. [PubMed: 21555956]
29. Morris C, Kurinczuk JJ, Fitzpatrick R, Rosenbaum P. Reliability of the manual ability classification system for children with cerebral palsy. *Dev Med Child Neurol*. 2006; 48:950–953. [PubMed: 17109781]
30. Beckung E, Hagberg G. Neuroimpairments, activity limitations, and participation restrictions in children with cerebral palsy. *Dev Med Child Neurol*. 2002; 44:309–316. [PubMed: 12033716]
31. Plasschaert VFP, Ketelar M, Nijhuis MG, Enkelaar L, Gorter JW. Classification of manual abilities in children with cerebral palsy under 5 years of age: how reliable is the Manual Ability Classification System? *Clinical Rehabilitation*. 2009; 23:164–170. [PubMed: 19129264]

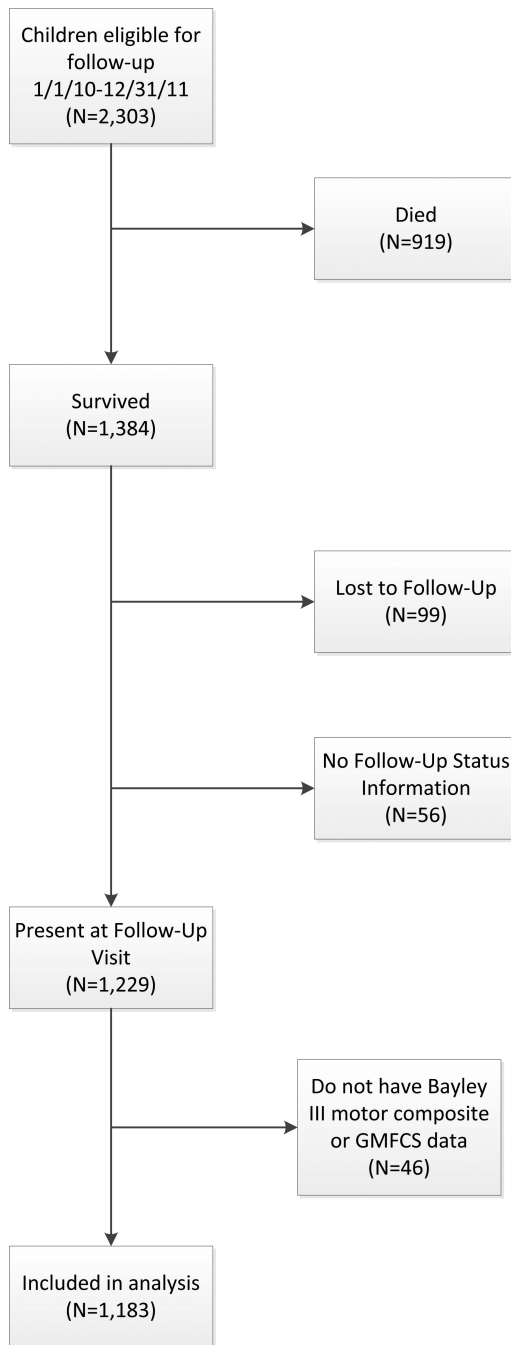


Figure 1. Sample Selection Flowchart

Eighty-five percent of survivors were included in follow-up analysis.

Table 1

Demographic and Medical Characteristics of Study Sample

| Characteristic | All (N=1,183) | Neurological Classification | | |
|------------------------------|---------------|-----------------------------|--------------------------|----------------------|
| | | Normal (N=780) | Abnormal: Non-CP (N=259) | Abnormal: CP (N=143) |
| | n(%)* | n(%)* | n(%)* | n(%)* |
| Male | 583 (49) | 374 (48) | 126 (49) | 82 (57) |
| Race/Ethnicity | | | | |
| Black | 470 (40) | 317 (41) | 95 (37) | 58 (41) |
| White | 626 (54) | 412 (53) | 141 (55) | 72 (51) |
| Other | 74 (6) | 45 (6) | 19 (7) | 10 (7) |
| Gestational age...mean (SD) | 25 (1) | 25 (1) | 25 (1) | 25 (1) |
| Birth weight...mean (SD) | 775 (169) | 794 (164) | 734 (176) | 747 (170) |
| SGA | 74 (6) | 35 (4) | 30 (12) | 9 (6) |
| Congenital anomaly | 25 (2) | 15 (2) | 6 (2) | 4 (3) |
| Postnatal steroids | 191 (16) | 106 (14) | 47 (18) | 37 (26) |
| Surfactant | 1053 (89) | 684 (88) | 235 (91) | 133 (93) |
| BPD (traditional definition) | 633 (54) | 385 (50) | 159 (62) | 88 (62) |
| Sepsis | 402 (34) | 234 (30) | 104 (40) | 64 (45) |
| IVH 3-4/PVL | 207 (18) | 84 (11) | 48 (19) | 75 (52) |
| Multiple birth | 263 (22) | 181 (23) | 49 (19) | 33 (23) |
| Proven NEC | 125 (11) | 76 (10) | 29 (11) | 20 (14) |
| ROP | 850 (72) | 524 (68) | 208 (81) | 117 (82) |
| Apgar...mean (SD) | | | | |
| 1-minute | 4 (2) | 4 (2) | 4 (2) | 3 (2) |
| 5-minute | 6 (2) | 6 (2) | 6 (2) | 6 (2) |
| Hospital stay...mean (SD) | 116 (47) | 105 (35) | 130 (57) | 145 (59) |

Acronyms: SGA= Small for Gestational Age, BPD= Bronchopulmonary Dysplasia, IVH= Intraventricular Hemorrhage; PVL= Periventricular Leukomalacia, NEC= Necrotizing Enterocolitis, ROP= Retinopathy of Prematurity

* Results presented as n (% of column N) unless otherwise specified as mean (SD).

Table 2

Sensitivity, Specificity, Positive and Negative Predictive Values of Bayley III Motor Composite Scores for Indicating Gross and/or Fine Motor Impairment

| Outcome | Motor Composite Cut Point | Sensitivity | Specificity | Positive Predictive Value | Negative Predictive Value |
|--|---------------------------|-------------|-------------|---------------------------|---------------------------|
| Gross Motor Impairment | | | | | |
| GMFCS Level 2 | 70 | 88 | 94 | 58 | 99 |
| | 73 | 94 | 93 | 54 | 99 |
| | 76 | 95 | 90 | 46 | 99 |
| | 79 | 97 | 86 | 38 | 100 |
| | 82 | 98 | 80 | 31 | 100 |
| | 85 | 99 | 75 | 26 | 100 |
| Fine Motor Impairment | | | | | |
| FMss < 3 | 70 | 100 | 93 | 38 | 100 |
| | 73 | 100 | 92 | 32 | 100 |
| | 76 | 100 | 89 | 26 | 100 |
| | 79 | 100 | 84 | 21 | 100 |
| | 82 | 100 | 79 | 16 | 100 |
| | 85 | 100 | 73 | 13 | 100 |
| Gross and/or Fine Motor Impairment* | | | | | |
| GMFCS Level 2 and/or FMss < 3 | 70 | 89 | 95 | 67 | 99 |
| | 73 | 95 | 94 | 62 | 99 |
| | 76 | 96 | 91 | 53 | 99 |
| | 79 | 97 | 87 | 44 | 100 |
| | 82 | 98 | 81 | 36 | 100 |
| | 85 | 99 | 76 | 30 | 100 |

* Of the 98 children with GMFCS 2, 71 had data for the fine motor composite scores. Of those 71, 31 (44%) had a fine motor score < 3. Of the 45 children with fine motor impairment, 31 (69%) had GMFCS 2.

Table 3

Fine Motor Skills Sets from Bayley-III for 19.5 month to 20.5 month Age Bracket

| Fine Motor Scaled Score | Raw Score | Skill Set: Tasks child likely to be able to do to get given score |
|--------------------------------|------------------|---|
| 1 | 25 | Bring blocks to midline with both hands simultaneously and Use pads of thumb and any finger-tip to grasp a pellet |
| 2 | 26 | Above plus: Isolate extended index finger or scribble spontaneously |
| 3 | 27 | Above plus: Stack two blocks or imitate a random stroke |
| 4 | 28 | Above plus: Place ten pellets in a bottle within 60 seconds |

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript