Original Article

arm
Annals of Rehabilitation Medicine

Assessment of Preterm Infants Using the Bayley-III Scales in Korea

Sung Ho Ahn, MD, Soo A Kim, MD

Department of Physical Medicine and Rehabilitation, Soonchunhyang University Cheonan Hospital,
Soonchunhyang University College of Medicine, Cheonan, Korea

Objective To assess the well-being of preterm newborns using the Bayley-III scales in a Korean-based population, and to evaluate the perinatal risk factors influencing developmental outcome.

Methods Using the Bayley-III scales, we assessed 120 preterm infants who were referred for evaluation of neurodevelopmental performance. We subdivided them into an extremely preterm group (n=18) and a very/moderate to late preterm group (n=102). Bayley-III mean scores and the rate of infants showing a delay were compared for both groups. The relationship between perinatal risk factors and Bayley-III scores was analyzed. The risk factors were considered as very low birth weight, history of neonatal medical problems, and abnormal radiologic findings in brain magnetic resonance images (MRIs).

Results Although no significant differences in mean scores were observed between the extremely preterm group and the very/moderate to late preterm group, the rate of babies showing developmental delay in motor composite scores was significantly higher in the extremely preterm group. The proportions of preterm infants with cognitive, language, and motor delays were 38.3%, 26.7%, and 35.0%, respectively. Very low birth weight was a significant risk factor for low cognitive, language, and motor composite scores. Also, abnormal radiologic findings on brain MRI were significant indicators of lower motor composite scores.

Conclusion Cognitive development was the most frequently delayed domain in preterm infants and motor development was more frequently delayed in the extremely preterm group. The very low birth weight and abnormal radiologic findings in brain MRI were predictive factors for neurodevelopmental outcome.

Keywords Bayley scale, Preterm, Perinatal risk factor, Neurodevelopmental outcome

INTRODUCTION

Preterm infants (born before 37 weeks gestational age) are known to be at high risk for neurodevelopmental delay [1]. These infants have significantly less developed executive functioning, smaller receptive and expressive

language lexicons, and weaker motor skills as compared with full-term infants [2]. The importance of early neurodevelopmental assessment for preterm infants has been highlighted because delayed intervention for neurodevelopmental delay may result in high rates of cognitive impairment, motor impairment, learning disabilities, and the need for special education placement [3].

The most widely used developmental assessment scheme is the Bayley Scales of Infant Development (BSID). The first two editions of the scales produce 2 scores: the Mental Developmental Index (MDI) was used to assess cognitive and language development, and the Psychomotor Developmental Index (PDI) to assess motor skills. The revised Bayley Scales of Infant and Toddler Development, third edition (Bayley-III, 2006) separated the MDI into distinct cognitive, receptive language, and expressive language scales, and the PDI into fine motor and gross motor scales [4].

Although the new edition provides indices of performance over separable domains, concerns have arisen regarding the interpretation of scores. Recent studies in the United States, United Kingdom, and Australia have reported that BSID-III assessments significantly underestimate the rate of developmental delay in preterm and full-term infants [5-7].

This study aims to investigate neurodevelopmental assessment of Korean preterm infants using Bayley-III scales. We also explore within-group patterns of performance and correlation between the perinatal risk factors and the Bayley-III scores.

MATERIALS AND METHODS

Study design

The study is a retrospective chart review of inpatient and outpatient medical records.

Study population

We identified all preterm infants who visited the Department of Physical Medicine and Rehabilitation and the Department of Pediatrics, Soonchunhyang University Hospital for assessment of developmental delay from July 1, 2014 to June 30, 2016. A total of 120 preterm infants screened for developmental status using Bayley-III scales. The study population consisted of 18 extremely preterm (gestational age [GA] <28 weeks), 29 very preterm (28 weeks ≤ GA < 32 weeks), and 73 moderate to late preterm (32 weeks ≤ GA < 37 weeks) infants. We divided preterm infants into extremely preterm group (n=18) and very/moderate to late preterm group (n=102). Although there was no exclusion based on perinatal characteristics, infants lacking medical records were excluded. This study was approved by the Soonchunhyang University Cheonan Hospital Institutional Review Board (No. 2016-08-009).

Perinatal and demographic data collection

Inpatient and outpatient charts were reviewed to determine demographic and perinatal characteristics, such as GA, corrected age at test, birthweight, sex, neonatal medical problems, radiologic findings, and prenatal maternal problem. Infants with very low birth weight less than 1,500 g were separated from the others. Neonatal medical problems were classified as pulmonary disease (respiratory distress syndrome, hyaline membrane disease, meconium aspiration, bronchopulmonary dysplasia, pneumonia, etc), cardiac disease (tetralogy of Fallot, atrial septal defect, ventricular septal defect, patent foramen ovale, etc.), seizure, and others. Prenatal maternal problems were classified as intrauterine problem (premature rupture of membranes, placenta previa, incompetent internal os of cervix, etc), maternal medical disease (hypertension, diabetes mellitus, hypothyroidism, etc.) and others. Radiologic abnormalities in brain magnetic resonance imaging (MRI) such as periventricular leukomalacia, intraventricular hemorrhage, or cerebral infarction were considered as perinatal risk factors. We then analyzed the relationship between perinatal risk factors and Bayley-III scores.

Outcome measures

Neurodevelopmental outcome was assessed at 1–48 months of corrected GA using the Bayley-III scales administered by a licensed occupational therapist. The examiner was rigorously trained prior to conducting the study as recommended in the guidelines of Bayley-III. The assessment consists of five scales: cognition, receptive language, expressive language, gross motor, and fine motor. The language composite scores consist of the combination of expressive and receptive language subscales; motor composite scores consist of gross and fine motor subscales [4]. The five subscales were used

independently as they identify differences in the language and motor scales which are not evident with the composite scores. The rates of developmental delay were determined using the Bayley-III test norms, indicating mild, moderate and severe delay, respectively [4]. We compared the Bayley-III mean scores and rate of infants showing delay for extremely preterm and very/moderate to late preterm groups.

Statistical analysis

All statistical analyses were performed using SPSS ver. 14.0 (SPSS Inc., Chicago, IL, USA) and R ver. 3.1.3 (The R Foundation for Statistical Computing, Vienna, Austria). Demographic characteristics were described by mean±standard deviation for continuous variables using an independent t-test and frequency (percentage, %) for categorical variables using a chi-square test or Fisher exact test. A Mann-Whitney U-test was used to compare the Bayley-III scores for the extremely preterm and very/moderate to late preterm groups. Fisher exact test was used to compare the rates of mild, moderate, and severe delay for both groups. The relationship between perinatal risk factors and Bayley-III scores was calculated. We used an independent t-test for comparing birthweight groups and a Kruskal-Wallis test for comparing radiologic findings and medical problems. A p-value of less than 0.05 was considered statistically significant.

RESULTS

Demographic characteristics

The perinatal and demographic characteristics of the two groups are shown in Table 1. As expected, there were significant group differences regarding mean GA and

Table 1. Demographic characteristics of preterm infants

Variable	Total (n=120)	GA <28 weeks (n=18)	GA ≥28 weeks (n=102)	p-value
GA (wk)	32.2±3.3	26.7±0.9	33.2±2.5	<0.001
Age at test (mo)	11.7±8.2	9.6±5.6	12.1±8.5	0.046
Birthweight (g)	1,814.0±695.5	911.4±196.7	1,973.2±626.5	<0.001
Sex				0.318
Male	63 (52.5)	7 (38.9)	56 (54.9)	
Female	57 (47.5)	11 (61.1)	46 (45.1)	
Medical problem				0.177
No	22 (18.3)	0 (0)	22 (21.6)	
Pulmonary disease	19 (15.8)	2 (11.1)	17 (16.7)	
Cardiac disease	3 (2.5)	1 (5.6)	2 (2.0)	
Seizure	5 (4.2)	1 (5.6)	4 (3.9)	
Others	71 (59.2)	14 (77.8)	57 (55.9)	
Radiologic finding				0.054
No	97 (80.8)	13 (72.2)	84 (82.4)	
PVL	11 (9.2)	1 (5.6)	10 (9.8)	
IVH	5 (4.2)	3 (16.7)	2 (2.0)	
Cerebral infarction	3 (2.5)	0 (0)	3 (2.9)	
Others	4 (3.3)	1 (5.6)	3 (2.9)	
Maternal problem				0.303
No	41 (34.2)	6 (33.3)	35 (34.3)	
Intrauterine problem	54 (45.0)	11 (61.1)	43 (42.2)	
Medical problem	22 (18.3)	1 (5.6)	21 (20.6)	
Others	3 (2.5)	0 (0)	3 (2.9)	

Values are presented as mean±standard deviation or number (%).
GA, gestational age; PVL, periventricular white matter leukomalacia; IVH, intraventricular hemorrhage.

birthweight (p<0.001). Also, the extremely preterm infants were assessed at a significantly earlier age (p<0.05).

Assessment of Bayley-III performance scores and developmental delay patterns

Table 2 presents Bayley-III composite and subscale scores per group along with group differences between the extremely preterm (GA <28 weeks) group and the very/moderate to late preterm (28 weeks ≤ GA < 37 weeks) group. In both groups, mean cognitive, language, and motor composite scores were 88.8±18.6, 92.0±17.5,

and 91.8±23.0, respectively. No significant differences in scores were observed between the extremely preterm group and the very/moderate to late preterm group.

Using normative criteria, the proportions of preterm infants with cognitive, language, and motor delay were 38.3%, 26.7%, and 35.0%, respectively. The rates for the extremely preterm group were 50.0%, 22.2%, and 55.6%, and those for the very/moderate to late preterm group were 36.3%, 27.5%, and 31.4% respectively. The rate of infants showing delays in motor composite scores was significantly higher in the extremely preterm group (Table 3).

Table 2. Bayley-III performance scores in preterm infants

Variable	Total (n=120)	GA <28 weeks (n=18)	GA ≥28 weeks (n=102)	p-value
CC (cognitive-composite)	88.8±18.6	85.1±17.8	89.5±18.8	0.278
CS (cognitive-subscale)	7.9±3.4	7.0±3.6	8.1±3.4	0.23
LC (language-composite)	92.0±17.5	94.3±11.3	91.5±18.4	0.807
RL (receptive language-subscale)	8.6±3.2	8.9±2.6	8.6±3.4	0.59
EL (expressive language-subscale)	9.0±3.6	9.2±2.1	9.0±3.8	0.64
MC (motor-composite)	91.8±23.0	87.3±20.3	92.6±23.5	0.11
FM (fine motor-subscale)	8.5±3.6	7.7±3.2	8.6±3.6	0.132
GM (gross motor-subscale)	9.1±4.2	8.1±4.0	9.2±4.2	0.233

Values are presented as mean±standard deviation.
GA, gestational age.

Table 3. Rates of developmental delay according to BSID-III composite scores

Variable	Total (n=120)	GA <28 weeks (n=18)	GA ≥28 weeks (n=102)	p-value
Cognitive scores				0.377
No delay	74 (61.7)	9 (50.0)	65 (63.7)	
Mild	22 (18.3)	5 (27.8)	17 (16.7)	
Moderate	6 (5.0)	0 (0)	6 (5.9)	
Severe	18 (15.0)	4 (22.2)	14 (13.7)	
Language scores				0.482
No delay	88 (73.3)	14 (77.8)	74 (72.5)	
Mild	7 (5.8)	1 (5.6)	6 (5.9)	
Moderate	14 (11.7)	3 (16.7)	11 (10.8)	
Severe	11 (9.2)	0 (0)	11 (10.8)	
Motor scores				0.025
No delay	78 (65.0)	8 (44.4)	70 (68.6)	
Mild	11 (9.2)	5 (27.8)	6 (5.9)	
Moderate	11 (9.2)	2 (11.1)	9 (8.8)	
Severe	20 (16.7)	3 (16.7)	17 (16.7)	

Values are presented as frequency (%).
GA, gestational age; BSID, Bayley Scales of Infant Development.

Perinatal and demographic correlation with Bayley-III scores

Very low birth weight (<1,500 g) was significantly related to lower cognitive, language, and motor composite scores on the Bayley-III scales. Abnormal radiologic findings in brain MRI were also significant risk factors for lower motor composite scores (Table 4). Otherwise, the history of medical problems was not significantly related to Bayley-III scores.

DISCUSSION

Understanding neurodevelopmental patterns is necessary since sensitive identification of cognitive, language, and motor delay is required for enrollment in early. Previous research has revealed that preterm infants demonstrate cognitive, language, and motor delays during the first postnatal year and into adolescence [2,8-10]. Our study examined neurodevelopmental patterns of preterm infants using the Bayley-III scales to quantify differences between extremely preterm and very/moderate to late preterm infants in a Korean-based population.

Previous studies show that extremely preterm infants have significantly lower function levels in cognitive, language, and motor areas compared with full-term infants [11,12]. We subdivided the preterm group into an extremely preterm group and a very/moderate to late preterm group. We investigated differences between the groups, unlike previous studies which analyzed extremely preterm or preterm infants separately while comparing with a full-term or control group. In the present study there is no significant difference of mean scores for all domains between the extremely preterm group and the very/moderate to late preterm group, although the frequency of infants showing developmental delay in motor composite scores was significantly higher in the extremely preterm group. However, this is less meaningful due to the small sample size of the extremely preterm group (n=18).

The Bayley-III scales were developed and standardized in 2006 using a random sample of 1,700 American children. The sample was reported to be representative of the 2000 US Bureau of the Census population survey data in terms of parent education, ethnicity, and geographic region [4]. It seems doubtful that US norms would be adequate in populations with different characteristics and

Table 4. Correlation between perinatal risk factors and BSID-III scores

Variable	No. of patients	Cognitive	p-value	Language	p-value	Motor	p-value
Birthweight (g)							
<1,500	41	84.8±20.4	<0.001	85.9±17.6	<0.001	87.3±24.4	<0.001[a]*
≥1,500	79	90.9±17.4		92.2±17.3		94.2±22.0	
Radiologic finding							
No	97	91.2±17.5	0.095	93.3±15.8	0.455	95.0±22.1	0.011[b]*
PVL	11	75.1±18.1		84.8±19.7		74.1±21.8	
IVH	5	86.0±18.2		98.2±5.0		92.2±8.1	
ICH, infarction	3	80.0±27.8		76.7±27.1		79.3±30.9	
Other radiologic finding	4	80.2±29.7		82.5±41.0		73.8±32.1	
Medical problem							
No	22	85.5±17.3	0.373	89.2±20.6	0.345	84.78± 21.3	0.135[b]
Pulmonary	19	87.1±17.1		92.8±13.7		91.5±22.4	
Seizure	3	76.3±58.5		107.0±9.6		103.0±12.0	
Cardiac	5	97.0±9.1		100.0±9.5		104.0±25.2	
Other medical problem	71	90.3±17.4		91.4±17.9		92.8±23.7	

Values are presented as mean±standard deviation.
BSID, Bayley Scales of Infant Development; PVL, periventricular white matter leukomalacia; IVH, intraventricular hemorrhage; ICH, intracerebral hemorrhage.
[a]t-test for between-group comparison, [b]Kruskal-Wallis test for between-group comparison.
*p<0.05.

languages, leading to potential inaccuracies in identification of referral thresholds. Therefore, we compared the Bayley-III performance of Korean preterm infants with that of US preterm infants. In a retrospective cohort of 85 preterm infants in United States, mean Bayley-III cognitive, language, and motor composite scores were 102±15, 91±15 and 94±17. In patterns of Bayley-III performance, mean language index and motor index were significantly lower than the cognitive index. For the majority (53%) of preterm infants, language development was the weakest domain; for another 39%, motor skill was the weakest area of development [13]. In our study, mean cognitive, language, and motor composite scores were 88.8±18.6, 92.0±17.5, and 91.8±23.0. The proportions of cognitive, language, and motor delay in the preterm group were 38.3%, 26.7%, and 35.0%, respectively. Cognitive development was the most frequently delayed domain. Although we could not statistically analyze the Bayley-III scales difference among US preterm infants and the preterm infants enrolled in our study, cognitive composite scores seemed to be relatively lower in our study, compared to the US study. Previous studies in other countries showed some differences between children of their country and US children in Bayley-III performance. For example, Taiwanese children had lower scores on the Bayley cognitive and motor scales when compared to the US norms at 6 and 24 months [14]. A Dutch cohort study revealed that the Dutch norms differ from the US norms for all subtests, therefore population specific norms appear necessary [15]. Hence it is important to establish normative data for Korean children and further large cohort study is needed to compare the Bayley-III scales of Korean preterm or full-term infants with those of US infants in order to establish appropriate norms of the Bayley-III scales in Korea.

The Bayley-III scales assessment has been reported to underestimate rates of delay [6]. Recently published cohort studies reporting preterm outcomes confirmed that if the performance of the preterm infants had been compared with the normative means, there would have been misleading frequencies of delay [5,7]. Although we could not enroll control group participants due to the characteristics of a retrospective study, doing so might have contributed no significant effect on the interpretation of our results. Most previous studies in Korea examining developmental outcomes have used the Bayley-II

scales until Kim and Shin [16] recently investigated developmental assessment using the Bayley-III scales. The Bayley-III scales demonstrated strong correlations with the Korean Infant and Child Development Test (KICDT), except in the social-emotional category. Further study is needed to assess the validity of the Bayley-III scales for Korean infants.

The present study also examined correlation between the perinatal risk factors and the Bayley-III subtest scores. Previous research has revealed that very low birth weight [5], low Apgar score [17], abnormal spike waves in electroencephalography [18], and abnormal findings in brain MRI [19-21] were related to poor outcomes in neurodevelopment. In addition to the severity of periventricular leukomalacia in brain MRI [20], the degree of intraventricular hemorrhage [21] was consistently related to neurodevelopmental outcome. Our results showed that very low birthweight (<1,500 g) was significantly related to lower cognitive, language, and motor composite scores, and abnormal findings in brain MRI significantly predicted lower motor composite scores. Medical problems such as pulmonary disease, seizure, cardiac disease, and others were not significantly related to Bayley-III scales (Table 4). Future investigations should include detailed socio-demographic variables including parental education and economic status.

This study has several limitations. First, the subjects were recruited from a tertiary care hospital in a neonatal clinical setting, meaning our results may not be appropriate to the general population. Second, instead of control group performance values we used the test norms as reference values since this retrospective study had no inclusion of a contemporary control group. Third, the sample of the extremely preterm group was too small to analyze group. However, there are some strengths of this study. Our study had a relatively large sample size in comparison to a previous study on Korean infants [16]. Also, this study targeted preterm infants to evaluate using the Bayley-III scales in Korea.

In conclusion, we assessed neurodevelopmental patterns of both extremely preterm and very/moderate to late preterm infants using the Bayley-III scales. Contrary to US studies, cognitive development was the most frequently delayed domain in our study. No significant difference in mean scores was observed between the extremely preterm group and the very/moderate to late

preterm group. However, the frequency of infants showing developmental delay in motor composite scores was significantly higher in the extremely preterm group, but this is less meaningful due to the small sample size. Very low birthweight (<1,500 g) was significantly related to lower cognitive, language, and motor composite scores, and abnormal findings in brain MRI significantly predicted lower motor composite scores. Further prospective studies including large enough sample size and a control group are needed.

CONFLICT OF INTEREST

No potential conflict of interest relevant to this article was reported.

REFERENCES

1. Stephens BE, Vohr BR. Neurodevelopmental outcome of the premature infant. Pediatr Clin North Am 2009; 56:631-46.

2. Evensen KA, Skranes J, Brubakk AM, Vik T. Predictive value of early motor evaluation in preterm very low birth weight and term small for gestational age children. Early Hum Dev 2009;85:511-8.

3. Aylward GP. Cognitive and neuropsychological outcomes: more than IQ scores. Ment Retard Dev Disabil Res Rev 2002;8:234-40.

4. Bayley N. Bayley scales of infant and toddler development. 3rd ed. San Antonio, TX: Harcourt Assessment; 2006.

5. Vohr BR, Stephens BE, Higgins RD, Bann CM, Hintz SR, Das A, et al. Are outcomes of extremely preterm infants improving? Impact of Bayley assessment on outcomes. J Pediatr 2012;161:222-8.

6. 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.

7. 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.

8. Bhutta AT, Cleves MA, Casey PH, Cradock MM, Anand KJ. Cognitive and behavioral outcomes of school-aged children who were born preterm: a meta-analysis. JAMA 2002;288:728-37.

9. Stolt S, Haataja L, Lapinleimu H, Lehtonen L. The early lexical development and its predictive value to language skills at 2 years in very-low-birth-weight children. J Commun Disord 2009;42:107-23.

10. Sun J, Mohay H, O'Callaghan M. A comparison of executive function in very preterm and term infants at 8 months corrected age. Early Hum Dev 2009;85:225-30.

11. Mansson J, Stjernqvist K. Children born extremely preterm show significant lower cognitive, language and motor function levels compared with children born at term, as measured by the Bayley-III at 2.5 years. Acta Paediatr 2014;103:504-11.

12. L Orton J, McGinley JL, Fox LM, Spittle AJ. Challenges of neurodevelopmental follow-up for extremely preterm infants at two years. Early Hum Dev 2015;91:689-94.

13. Greene MM, Patra K, Nelson MN, Silvestri JM. Evaluating preterm infants with the Bayley-III: patterns and correlates of development. Res Dev Disabil 2012;33: 1948-56.

14. Wu YT, Tsou KI, Hsu CH, Fang LJ, Yao G, Jeng SF et al. Brief report: Taiwanese infants' mental and motor development: 6-24 months. J Pediatr Psychol 2008;33: 102-8.

15. Steenis LJ, Verhoeven M, Hessen DJ, van Baar AL. Performance of Dutch children on the Bayley III: a comparison study of US and Dutch norms. PLoS One 2015;10:e0132871.

16. Kim HJ, Shin JI. A study of the correlation between BSID-III and KICDT for children with developmental delay. J Phys Ther Sci 2015;27:269-71.

17. Woo MK, Kim DW, Huh K, Shim GH, Chey MJ. Study on the neurodevelopmental predictors for the results of the Bayley Scales of Infant Development II in high-risk neonates. Korean J Pediatr 2009;52:1221-7.

18. Iype M, Prasad M, Nair PM, Geetha S, Kailas L. The newborn with seizures: a follow-up study. Indian Pediatr 2008;45:749-52.

19. van de Bor M, den Ouden L, Guit GL. Value of cranial ultrasound and magnetic resonance imaging in predicting neurodevelopmental outcome in preterm infants. Pediatrics 1992;90(2 Pt 1):196-9.

20. Lee HJ, Park SH, Na KH, Park SY, Kim EY, Kim KS, et al. Developmental assessment of preterm infants at

two years of age with periventricular leukomalacia. J Korean Soc Neonatol 2002;9:167-75.

21. Patra K, Wilson-Costello D, Taylor HG, Mercuri-Minich N, Hack M. Grades I-II intraventricular hemorrhage in extremely low birth weight infants: effects on neurodevelopment. J Pediatr 2006;149:169-73.