Original Article

arm
Annals of Rehabilitation Medicine

Reliability and Applicability of the Bayley Scale of Infant Development-II for Children With Cerebral Palsy

Ji Hyun Lee, MD, Hye Kyung Lim, EunYoung Park, Junyoung Song, MD,
Hee Song Lee, MD, Jooyeon Ko, PhD, MinYoung Kim, MD

Department of Rehabilitation Medicine, CHA Bundang Medical Center, CHA University School of Medicine, Seongnam, Korea

Objective To obtain reliability and applicability of the Korean version Bayley Scale of Infant Development-II (BSID-II) in evaluating the developmental status of children with cerebral palsy (CP).

Methods The inter-rater reliability of BSID-II scores from 68 children with CP (46 boys and 22 girls; mean age, 32.54±16.76 months; age range, 4 to 78 months) was evaluated by 10 pediatric occupational therapists. Patients were classified in several ways according to age group, typology, and the severity of motor impairment by the level of the Gross Motor Function Classification System (GMFCS). The measures were performed by video analysis, and the results of intraclass correlation (ICC) were obtained for each of the above classifications. To evaluate the clinical applicability of BSID-II for CP, its correlation with the Gross Motor Function Measure (GMFM), which has been known as the standard motor assessment for CP, was investigated.

Results ICC was 0.99 for the Mental scale and 0.98 for the Motor scale in all subjects. The values of ICC ranged from 0.92 to 0.99 for each age group, 0.93 to 0.99 for each typology, and 0.99 to 1.00 for each GMFCS level. A strong positive correlation was found between the BSID-II Motor raw score and the GMFM total score (r=0.84, p<0.001), and a moderate correlation was observed between the BSID-II Mental raw score and the GMFM total score (r=0.65, p<0.001).

Conclusion The Korean version of BSID-II is a reliable tool to measure the functional status of children with CP. The raw scores of BSID-II showed a great correlation with GMFM, indicating validity of this measure for children with CP on clinical basis.

Keywords Cerebral palsy, Child development, Reliability, Gross motor function

INTRODUCTION

Cerebral palsy (CP) is a group of disorders due to impaired development of movement and posture, which are attributed to various non-progressive damages in the developing fetal or infant brain. Motor dysfunctions in CP are often accompanied by various problems involving sensation, cognition, communication, strabismus, perception, and behavior [1]. Recently, due to the con-

spicuous development of pediatric medicine, the survival rate of children with high risks, such as premature-born children or those with traumatic brain injury has increased [2]. When treating children with neurological or developmental disorders, the exact assessment of their current developmental status is essential for planning the strategy of therapy and for determining therapeutic efficacy. However, the functional implications of CP involves various developmental domains as listed above, that is, the gross motor domain, cognition, communication, perception, etc. It should be also noted that the severity of involvement varies in each domain. Moreover, the functional development process of children with CP does not always follow the routine developmental stages in normal children. Thus, it is difficult to assess variable status of CP exactly, and the issues involving the assessment of children with CP using a tool are reliability and clinical applicability.

Reliability indicates reproducibility of the same value through repetitive assessment. When reproducibility is measured by different raters, it is called "inter-rater reliability". It reflects the degree of standardization of the test and the ability of the raters to perform the evaluation correctly [3]. Therefore, confirming the reliability of a tool is mandatory before its use in practice. There are some reports about reliable measures of motor function and performances in children with CP [3,4]. However, these reports have focused on gross motor function only, which is related to ambulation ability. Few studies have reported the reliability of measurement tools in regard to other important aspects of CP including cognition, communication, perception, and fine motor function [5].

The Bayley Scales of Infant Development (BSID) is a world-wide tool used to evaluate the development of variable cognitive function as well as motor performances of infants and children [6]. The BSID was published on the basis of the California Mental scale and Motor scale that have been used since the 1930s. Through revision, the second edition was released in 1993, and is currently in use around the world [6]. This version was adopted and standardized as the Korean version of Bayley Scale of Infant Development-II (BSID-II) in Korea and has been used since 2004 [7].

The BSID-II itself was not purposed to diagnose impairments, but rather focuses on providing normative data to assess the current status by comparing with the norms [8].

For its applicability, only those with Down syndrome [9], prematurity [10], and prenatal drug exposure [11] were validated as the subjects in the original English version. The Korean interpretation version provided validity from just 14 disabled children without a specific diagnosis [7]. Although some studies used BSID-II for the assessment of the functional status of CP, no study revealed the validity of BSID-II application for CP subjects before enrollment [5,12-18]. Since the BSID-II was not developed to quantify the ability of children with CP with a wide range of motor and cognitive dysfunction [8], the scale needs to be validated for CP subjects prior to interpretation of the results. To validate a function measuring instrument, determination of its reliability and ability in assessing the purposed functional status is essential.

The goal of this study was to evaluate the reliability and ability of the function, i.e., clinical applicability, of the Korean version of BSID-II for children with CP. To confirm the reliability for a whole range of conditions, each stratified group was examined according to age, typology, and severity. To evaluate clinical applicability, the correlation between the raw scores of BSID-II and Gross Motor Function Measurement (GMFM), which is known as the touchstone of the functional evaluation of CP, was investigated. The correlation between the Mental and Motor scales was also assessed.

MATERIALS AND METHODS

This study was commenced after approval of the Institutional Review Board of CHA Medical Center, Republic of Korea. The parents of the participants provided written informed consent for this study before enrollment. Performance during each assessment of BSID-II was video-recorded, and each video record was used for BSID-II scoring by the other enrolled evaluators.

Subjects

The participants were children with CP who were receiving rehabilitation treatment from December 2010 to January 2011. The inclusion criteria were: a diagnosis of CP, who showed abnormalities in movement and posture, muscle tonus, and detectable brain lesion by imaging studies which correlated with physical impairment. The exclusion criteria were the presence of congenital anomalies or highly possible genetic syndrome and any

medico-surgical condition affecting the analysis. The defining diagnosis of CP was made by a pediatric rehabilitation medicine doctor. By referring to Shoukri's study [19], the sample size was elicited. Ten occupational therapists conducted BSID-II assessment, and 68 children with CP (46 males and 22 females), whose function was scored under 42 months of age using both the Motor and Mental scales, participated in the present study. The subjects were classified by three kinds of criteria for further analyses. Firstly, the classification was made according to their chronological age as five subgroups: 0–12, 13–24, 25–36, 37–42, and more than 42 months. They were also classified into five subgroups by typology: spastic bilateral, spastic unilateral, dystonia, chorea athetosis, and ataxia [1]. Classification in terms of the severity of functional motor abilities was made according to the Gross Motor Function Classification System (GMFCS) from the least impaired as level I to the most severely impaired as level V [20].

Evaluation tools

All subjects were evaluated with BSID-II, GMFM-88, and GMFCS.

Korean version of the BSID-II

The Korean BSID-II is an evaluation tool for the assessment of the developmental status of individual children, which was standardized by Park and Cho [7]. The Korean BSID-II is used for children in the age range of 1 to 42 months, however, it is also applicable to children over 42 months of age with developmental delay aged if their function is below than their normal counterparts [7,8]. BSID-II consists of Mental, Motor, and behavior rating scales. The Mental scale provides a raw score, a developmental age of mental status, and a Mental Developmental Index (MDI); the Motor scale provides a raw score, a developmental age of motor related function, and a Psychomotor Developmental Index (PDI) [6,21]. The behavior scale is about the quality of patient behavior during the test [7]. The Mental scale was designed to assess mainly cognition through evaluation of sensory/perception, knowledge, memory, problem solving skills, and language. The Motor scale tests evaluate the ability to control gross muscle groups responsible for movements associated with crawling, sitting, walking, and jumping and tests fine motor manipulations involved in prehen-

sion, adaptative use of writing implements and imitation of hand movements [22].

GMFM-88 and GMFCS

The GMFM-88 is a criterion-referenced observational measure for the assessment of children with CP [23]; it consists of 88 items grouped into 5 dimensions: lying and rolling; sitting; crawling and kneeling; standing; and walking, running, and jumping. The scale was proposed to quantitatively evaluate gross motor function. Score for each dimension is expressed as a percentage of the maximum score for that dimension. The total score is calculated by averaging the percentage scores across the 5 dimensions, range from 0 to 100.

During the GMFM evaluation, GMFCS is also assessed. GMFCS was developed by Palisano et al. [24] to classify the degree of gross motor impairment of children with CP into five levels. The distinction between each level is based on the ability to move and the need of supporting devices.

Methods to evaluate inter-rater reliability of BSID-II

Ten pediatric occupational therapists who were well-educated in conducting BSID-II served as raters in this study. Three of them had experience exclusively in the pediatric setting for more than 8 years, 2 had 7 years, and 5 had 2 years of experience. Before the study began, they went through a training session for about 1 month, 3 days a week, for more than 4 hours a day by watching and scoring video recordings of actual tests by each therapist.

The testing time for one child was approximately 40 minutes to 1 hour. While each therapist carried out the BSID-II, the whole process was video-recorded by an assistant therapist who sorted the process into mental and motor parts afterwards. The other 9 therapists then assessed the same patient by watching the video recordings. The pediatric physical therapist conducted GMFM and GMFCS for the same patients within 7 days after the BSID-II exam.

In addition, the raters searched for the limitations of the Korean BSID-II in children with CP as complementary to the original version of BSID-II.

Statistical analysis

Intraclass correlations (ICCs) for inter-rater reliability were analyzed according to age groups, the typology of

CP, GMFCS levels, and career of the raters. Correlation analyses were performed to evaluate the relationship

Table 1. General characteristics of subjects (n=68)

Characteristic	No. (%)
Sex	
Male	48 (70.6)
Female	20 (29.4)
Chronological age (mo), mean±SD	32.54±16.76
0–12	6 (8.8)
13–24	21 (30.9)
25–36	17 (25.0)
37–42	11 (16.2)
>42	13 (19.1)
Birth weight (kg), mean±SD	2.43±0.94
<1.0	6
1.0–1.5	8
1.5–2.5	20
>2.5	34
Gestational age (wk)	
<32	22
33–37	12
37–42	34
Typology	
Spastic bilateral (SB)	50 (73.5)
GMFCS level I/II/III/IV/V	10/3/11/15/11
Spastic unilateral (SU)	9 (13.2)
GMFCS level I/II	7/2
Dystonia (D)	6 (8.8)
GMFCS level II/IV/V	1/1/4
Chorea athetosis (CA)	0 (0)
Ataxia (A)	3 (4.4)
GMFCS level II/IV	1/2
GMFCS level	
I	17 (25)
SB/SU	10/7
II	7 (10.3)
SB/SU/D/A	3/2/1/1
III	11 (16.2)
SB	11
IV	18 (26.5)
SB/D/A	15/1/2
V	15 (22.1)
SB/D	11/4

GMFCS, Gross Motor Function Classification System; SD, standard deviation.

between the GMFM total score and the raw scores of the Motor and Mental scales in BSID-II with the Pearson correlation coefficient or the Spearman rank coefficient according to the number of samples. The correlation between MDI and PDI was also obtained with the Pearson correlation coefficient or the Spearman rank coefficient. The correlation between GMFCS levels and the raw scores of the Motor and Mental scales in BSID-II were analyzed with the Spearman rank coefficient.

For the analyses, SPSS ver. 19.0 (IBM, Armonk, NY, USA) program in CHA Medical Center was used. For this study, ICCs below 0.75 were considered 'poor to fair', those above 0.75 were considered 'good', and above 0.90 'excellent' [25]. In terms of correlation, the coefficient $r \geq 0.8$ indicated 'high' correlation, 0.6–0.8 'good', 0.4–0.6 'moderate', and ≤ 0.4 'poor' [26].

RESULTS

Characteristics of population

Sixty-eight children with CP, 48 boys (70.6%) and 20 girls (29.4%) participated in the present study. Their mean age was 32.54±16.76 months (range, 8 to 78 months). Demographic data are presented in Table 1.

The mean raw scores of the BSID-II Mental and Motor scales measured by the raters are presented in Table 2.

Inter-rater reliability of BSID-II

The ICC values for inter-rater reliability of BSID-II

Table 2. Measured mean raw scores of the BSID-II Mental and Motor Scales (n=68)

Examiner	Motor scale	Mental scale
A	50.90±24.56	95.24±42.86
B	52.44±24.84	96.32±42.06
C	52.09±24.58	95.62±42.41
D	51.74±24.59	95.75±42.33
E	51.90±24.85	94.32±42.34
F	51.69±24.12	95.96±42.28
G	51.93±24.77	95.68±41.99
H	52.25±25.23	95.74±41.68
I	51.94±25.40	95.09±41.76
J	51.91±24.35	95.93±42.13

BSID-II, Bayley Scale of Infant Development-II of Korean version (full score of Motor scale is 112 and full score of Mental scale is 178).

scores assessed by 10 raters were 0.99 for the Mental scale and 0.98 for the Motor scale in all subjects, which was interpreted as having 'excellent' reliability (Table 3). The ICC values of BSID-II scores for each group divided by age, typology, GMFCS levels, and career of the raters also demonstrated 'excellent' reliability (Table 3).

Correlation between GMFCS and raw scores of BSID-II

The GMFCS levels had a significant negative correlation with both Motor scale (r=-0.86, p<0.001) and Mental scale (r=-0.60, p<0.001) raw scores of BSID-II.

Correlation between GMFM and raw scores of BSID-II

The total score of GMFM showed a positive correlation with both Motor scale (r=0.84, p<0.001) and Mental scale (r=0.65, p<0.001) scores of BSID-II in all subjects. Analysis according to typology revealed a high or good correlation between the GMFM total score and the Motor scale scores in the spastic bilateral and spastic unilateral subgroups (rs>0.77, p<0.001), while dystonia and ataxia did not (Table 4). A moderate degree and high degree correlation between GMFM and Mental scale scores was observed in the spastic bilateral and spastic unilateral subgroups, respectively. When analysis was conducted according to GMFCS level, GMFCS levels I, IV, and V showed a moderate to high degree correlation between the GMFM and both BSID-II Motor and Mental scale

Table 3. Intraclass coefficients of BSID-II

	No.	Motor scale	Mental scale
Total subjects	68	0.98	0.99
Age (mo)			
0–12	6	0.92	0.93
13–24	21	0.99	0.98
25–36	17	0.99	0.99
37–42	11	0.97	0.99
>42	13	0.99	0.99
Typology			
Spastic bilateral	50	0.99	0.99
Spastic unilateral	9	0.97	0.99
Dystonia	6	0.97	0.96
Ataxia	3	0.93	0.93
GMFCS level			
I	17	0.99	0.99
II	7	0.99	1.00
III	11	1.00	1.00
IV	18	0.99	0.99
V	15	0.99	1.00
Career duration of therapist (yr)			
≥8	3	0.99	0.99
1–2	5	0.99	0.99

BSID-II, Bayley Scale of Infant Development-II of Korean version; GMFCS, Gross Motor Function Classification System.

Table 4. Correlation between BSID-II and GMFM total scores

	No.	GMFM (total score) Mean±SD	Motor (raw score) Mean±SD	r	p-value	Mental (raw score) Mean±SD	r	p-value
Total subjects[a]	68	45.6±28.24	51.88±24.61	0.84	<0.001*	95.94±42.06	0.65	<0.001*
Typology								
Spastic bilateral[a]	50	43.62±27.58	50.25±23.82	0.78	<0.001*	93.16±43.27	0.59	<0.001*
Spastic unilateral[b]	9	75.89±16.05	77.62±18.60	0.98	<0.001*	128.11±31.21	0.98	<0.001*
Dystonia[b]	6	20.83±16.90	31.32±14.15	0.64	0.170	73.87±33.32	0.64	0.170
Ataxia[b]	3	37.33±12.74	43.00±8.73	0.50	0.670	81.07±9.40	-1.00	<0.001*
GMFCS level[b]								
I	17	81.61±11.13	78.68±16.42	0.73	0.001*	125.86±29.17	0.63	0.007*
II	7	57.57±9.98	60.79±16.71	-0.12	0.800	93.94±30.77	-0.19	0.680
III	11	51.91±17.81	60.99±14.22	-0.52	0.100	107.08±37.20	-0.08	0.820
IV	18	32.00±12.07	42.16±13.64	0.91	<0.001*	98.17±40.66	0.77	<0.001*
V	15	10.89±4.77	22.35±8.29	0.61	0.020*	50.33±25.92	0.59	0.020*

BSID-II, Bayley Scale of Infant Development-II of Korean version; GMFM, Gross Motor Function Measure; SD, standard deviation.
*p<0.05. [a]Analyzed by Pearson correlation coefficient. [b]Analyzed by Spearman rank coefficient.

Table 5. Correlation between MDI and PDI of BSID-II

	No.	MDI (mean±SD)	PDI (mean±SD)	r	p-value
Total subjects[a]	55	59.36±17.25	53.38±9.70	0.59	<0.001*
Typology					
Spastic bilateral[a]	40	46.06±26.92	52.83±9.92	0.61	<0.001*
Spastic unilateral[b]	8	67.89±33.95	59.13±11.27	0.91	0.002*
Dystonia[b]	4	33.50±25.95	50.00±0.00	1.00	<0.001*
Ataxia[b]	3	50.00±0.00	50.00±0.00	NA	NA

MDI, mental developmental index; PDI, psychomotor developmental index; BSID-II, Bayley Scale of Infant Development-II of Korean version; SD, standard deviation; NA, not applicable.
*p<0.05. [a]Analyzed by Pearson correlation coefficient. [b]Analyzed by Spearman rank coefficient.

scores (rs>0.58, p<0.030) (Table 4).

Correlation between PDI and MDI of BSID-II

Of a total of 68 participants, 55 were younger than 42 months of age and were enrolled for PDI and MDI correlation analysis. A moderate degree of positive correlation between the PDI and MDI was shown in all subjects. The correlations were good to high in all the subgroups of typology, except for ataxia (rs>0.60, p<0.030) (Table 5).

Comments about considerable points of Korean BSID-II in evaluating children with CP with reference to the original version

In the research process, we discovered several possible inaccuracies or ambiguous terminologies in the Motor scale of the Korean BSID-II with reference to the original English version of BSID-II. In the 23rd, 29th, and 32nd items in the Motor scale, the given condition for scoring could be misunderstood as 'bench sitting' by describing it as 'sitting on a table' in the Korean version. While the original version said 'sitting on the bench surface', Korean version could be mean 'sitting mat surface'. In the 27th item, the description 'moves wrist' in the Korean version could include wrist flexion and extension, but in the original version, it is expressed as 'rotates wrist'. In the 67th item, 'If the child stands up after lying on the belly' is not a clear expression, whereas it was described as 'If the child rolls into the prone position before standing up' in the original version. In the 78th item, the height of rope for jumping in the Korean version is 5 cm lower than the original version's 8 inch, which might be a misprint of 15 cm. In the 91st item, 'manipulates pencil in hand' is represented as 'involvement of any fingers fulfills the score', while the original version accepts 'using only the

fingers to position the pencil'. In the present study, we chose to follow the policy of the original version of BSID-II for such ambiguous items.

Four pivotal typing errors were found in the scoring system and have been reported to the press company.

DISCUSSION

In this study, inter-rater reliability of the Korean version of BSID-II was examined by ten raters, based on the scores from the video recordings of children with CP. The inter-rater reliability was found to be very excellent for Korean BSID-II in both the Motor and Mental scales when applied to children with CP. High values of ICC results were not affected by age, typology or severity. This shows significance because BSID-II enables quantitative observation of various abnormal functions of children with CP during their development in the absence of available tools, especially for cognitive function [27]. To date, only the GMFM has been validated specifically for children with CP and has been used as a standardized measurement for observing gross motor function [23,28]. In the present study, we examined the clinical availability of the Motor and Mental scales of BSID-II by analyzing the correlation with GMFM score. The results showed a good correlation between the Motor scale of BSID-II and GMFM in all subjects. The results induce similarity in the observational power of the BSID-II Motor scale with that of GMFM. The Mental scale also showed a correlation with the GMFM to a moderate degree.

Further subgroup analysis according to typology and severity revealed a strong correlation of the BSID-II Motor scale with GMFM in the spastic bilateral and spastic unilateral subgroups, and the GMFCS level I, IV, and V,

which also showed positive correlations even with Mental scales. However, the GMFCS level II and III groups were different by showing no correlation. This can be interpreted by the idea that there are obviously different observational points which were brought by differences in the scoring point of the two independent systems which were accentuated in those groups. Although the GMFCS level II and II groups involved other typologies, the majority of them were spastic bilateral (Table 1). Considering their function level, they mostly corresponded to 'spastic diplegia' by the old CP classification that has better motor function in upper limbs and poorer motor function in lower limbs. Considering assessment points of each measure, it is understandable that the GMFCS level II and III groups showed no correlation of the BSID-II Motor scale with GMFM. The GMFM mainly assesses the function of the lower extremities whereas the BSID-II Motor scales consider the function of the upper extremities as one of the important determinant.

The Mental scale of BSID-II or MDI measures a broad range of cognition, including sensory/perception, language, and social skills [8]. We did not assess cognitive function with other instruments besides the BSID-II Mental scale. Thus, validation for cognition could not be fully provided in this study. However, considering the high reliability, its significant correlation with motor variables and the validity of BSID-II Motor scale which has the same scoring system, the BSID-II Mental scale can be suggested as a cognition evaluating tool for CP subjects.

All subject groups showed a moderate degree of correlation between MDI and PDI [26]. According to the subgroup analysis, the spastic unilateral subgroup showed a strong relationship between MDI and PDI. However, the number of subjects was too small to draw any conclusions. Good correlation in the spastic bilateral subgroup seems rational, because the severity of brain damage determines both motor and cognitive sequalae.

Taken together, BSID-II is a useful tool for the evaluation of children with CP who are under the developmental status for motor and other functions including cognition. Moreover, the scale is useful for the evaluation of children aged more than 42 months whose functions are lower than their normal counterparts. The overall correlation between BSID-II scores and GMFM was high (Table 4). Thus, this confirms the validity of the score system for children with CP.

In conclsion, inter-rater reliability of the BSID-II Motor and Mental scales by ten raters was shown to be very high. The BSID-II Motor scales showed a high correlation with the GMFM in all subjects. MDI and PDI were also well correlated. The results of this study show that BSID-II is a valid tool for the evaluation of various functions of CP.

CONFLICT OF INTEREST

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

REFERENCES

1. Rosenbaum P, Paneth N, Leviton A, Goldstein M, Bax M, Damiano D, et al. A report: the definition and classification of cerebral palsy April 2006. Dev Med Child Neurol Suppl 2007;109:8-14.

2. Vohr BR, Wright LL, Dusick AM, Mele L, Verter J, Steichen JJ, et al. Neurodevelopmental and functional outcomes of extremely low birth weight infants in the National Institute of Child Health and Human Development Neonatal Research Network, 1993-1994. Pediatrics 2000;105:1216-26.

3. Harris SR, Haley SM, Tada WL, Swanson MW. Reliability of observational measures of the Movement Assessment of Infants. Phys Ther 1984;64:471-7.

4. Thomas SS, Buckon CE, Phillips DS, Aiona MD, Sussman MD. Interobserver reliability of the gross motor performance measure: preliminary results. Dev Med Child Neurol 2001;43:97-102.

5. Enkelaar L, Ketelaar M, Gorter JW. Association between motor and mental functioning in toddlers with cerebral palsy. Dev Neurorehabil 2008;11:276-82.

6. 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.

7. Park HW, Cho BH. Korean Bayley Scales of Infant Development: interpretation manual. 2nd ed. Seoul: Kidspop Publishing Co.; 2006.

8. Bayley N. Bayley Scales of Infant Development: manual. 2nd ed. San Antonio, TX: Psychological Corp.; 1993.

9. Moore DG, Goodwin JE, Oates JM. A modified ver-

sion of the Bayley Scales of Infant Development-II for cognitive matching of infants with and without Down syndrome. J Intellect Disabil Res 2008;52 Pt 6:554-61.

10. Feldman R, Eidelman AI. Direct and indirect effects of breast milk on the neurobehavioral and cognitive development of premature infants. Dev Psychobiol 2003;43:109-19.

11. Schuler ME, Nair P, Harrington D. Developmental outcome of drug-exposed children through 30 months: a comparison of Bayley and Bayley-II. Psychol Assess 2003;15:435-8.

12. Guillen U, DeMauro S, Ma L, Zupancic J, Roberts R, Schmidt B, et al. Relationship between attrition and neurodevelopmental impairment rates in extremely preterm infants at 18 to 24 months: a systematic review. Arch Pediatr Adolesc Med 2012;166:178-84.

13. Ballot DE, Potterton J, Chirwa T, Hilburn N, Cooper PA. Developmental outcome of very low birth weight infants in a developing country. BMC Pediatr 2012;12:11.

14. O'Shea TM, Allred EN, Kuban KC, Hirtz D, Specter B, Durfee S, et al. Intraventricular hemorrhage and developmental outcomes at 24 months of age in extremely preterm infants. J Child Neurol 2012;27:22-9.

15. Skiold B, Vollmer B, Bohm B, Hallberg B, Horsch S, Mosskin M, et al. Neonatal magnetic resonance imaging and outcome at age 30 months in extremely preterm infants. J Pediatr 2012;160:559-66.

16. Hamer EG, Bos AF, Hadders-Algra M. Assessment of specific characteristics of abnormal general movements: does it enhance the prediction of cerebral palsy? Dev Med Child Neurol 2011;53:751-6.

17. Romeo DM, Cioni M, Battaglia LR, Palermo F, Mazzone D. Spectrum of gross motor and cognitive functions in children with cerebral palsy: gender differences. Eur J Paediatr Neurol 2011;15:53-8.

18. Constantinou JC, Adamson-Macedo EN, Mirmiran M, Fleisher BE. Movement, imaging, and neurobehavioral assessment as predictors of cerebral palsy in preterm infants. J Perinatol 2007;27:225-9.

19. Shoukri MM, Asyali MH, Donner A. Sample size requirements for the design of reliability study: review and new results. Stat Methods Med Res 2004;13:251-71.

20. Gunel MK, Mutlu A, Tarsuslu T, Livanelioglu A. Relationship among the Manual Ability Classification System (MACS), the Gross Motor Function Classification System (GMFCS), and the functional status (WeeFIM) in children with spastic cerebral palsy. Eur J Pediatr 2009;168:477-85.

21. 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.

22. Morley R, Fewtrell MS, Abbott RA, Stephenson T, MacFadyen U, Lucas A. Neurodevelopment in children born small for gestational age: a randomized trial of nutrient-enriched versus standard formula and comparison with a reference breastfed group. Pediatrics 2004;113(3 Pt 1):515-21.

23. Russell DJ, Avery LM, Rosenbaum PL, Raina PS, Walter SD, Palisano RJ. Improved scaling of the gross motor function measure for children with cerebral palsy: evidence of reliability and validity. Phys Ther 2000;80:873-85.

24. 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.

25. Portney LG, Watkins MP. Foundations of clinical research: applications to practice. 2nd ed. Upper Saddle River: Prentice-Hall; 2000.

26. Meyers CR. Measurement in physical education. 1st ed. New York: Ronald Press Co.; 1974.

27. Munck P, Niemi P, Lapinleimu H, Lehtonen L, Haataja L; PIPARI Study Group. Stability of cognitive outcome from 2 to 5 years of age in very low birth weight children. Pediatrics 2012;129:503-8.

28. Wong EC, Man DW. Gross motor function measure for children with cerebral palsy. Int J Rehabil Res 2005;28:355-9.