CONCURRENT VALIDITY OF THE COMPREHENSIVE DEVELOPMENTAL INVENTORY FOR INFANTS AND TODDLERS WITH THE BAYLEY SCALES OF INFANT DEVELOPMENT-II IN PRETERM INFANTS

Hua-Fang Liao,[1] Tien-Miau Wang,[2] Grace Yao,[3] and Wang-Tso Lee[4]

Background and Purpose: The Comprehensive Developmental Inventory for Infants and Toddlers (CDIIT) is a new developmental test designed in Taiwan and lacks concurrent validity information. This study investigated the concurrent validity of the CDIIT with the Bayley Scales of Infant Development-II (BSID-II) in preterm infants aged 6-18 and 21-40 months, respectively.

Methods: We recruited 160 preterm infants (84 boys, 76 girls) with a corrected age of 6-18 months and followed them up until 21-40 months of age. One tester administered the CDIIT and BSID-II to all infants. Developmental ages (DAs) and developmental quotients (DQs) from both tests were analyzed with Pearson correlation and κ tests.

Results: Correlation coefficients for DAs and DQs between the 2 cognitive and motor subtests and classification agreements were high at 6-18 months (r = 0.80-0.97; κ = 0.80, 0.85) and moderate to high at 21-40 months (r = 0.60-0.77; κ = 0.44, 0.57). DQ classification for the CDIIT motor subtests tended to be higher than for the BSID-II motor scales at 21-40 months.

Conclusions: In preterm infants, concurrent validity between the motor and cognitive subtests of the CDIIT and the BSID-II was acceptable. The CDIIT can be thus used in clinics for the early identification of developmental delay in infants and toddlers.

Key words: Child development; Diagnostic tests; Longitudinal studies; Mass screening; Premature birth

J Formos Med Assoc 2005;**104**:731-7

About 6-9% of children aged 0-6 years may have developmental delays requiring early intervention.[1] A valid developmental diagnostic test can help clinicians to identify psychomotor problems early[2] and to predict later school performance.[3] Approximately 5-10% of all births are preterm,[4] and these infants usually have an increased risk of poor motor, mental, and behavior-related developmental outcomes.[5,6] Lower mental or motor developmental quotients (DQs) than those of full-term infants since 24 weeks of age are found, even when corrected ages are used.[7] Therefore, preterm infants usually need an early developmental assessment.[7]

In the western literature, the Bayley Scales of Infant Development-II (BSID-II) is regarded as the best measure of developmental status of infants.[8] For specialists in Taiwan, however, the BSID-II is considered lengthy and expensive. More important, it does not have Taiwanese norms, and it does not address developmental skills related to traditional Chinese culture, such as the use of chopsticks for feeding and Mandarin language development. The Comprehensive Developmental Inventory for Infants and Toddlers (CDIIT) is a new developmental test in Taiwan. Standardization of the CDIIT was completed in 1996 in compliance with the Child Welfare Act of 1993 and the Special Education Act of 1984, which call for the early identification of Taiwanese children with developmental delays.[9] The reliability and validity of any new diagnostic test should be thoroughly evaluated before its introduction into daily practice.[8] Validity includes content, concurrent, predictive, and construct validity, as the American Psychological Association recommends.[10] A new test must have evidence of concurrent validity, which refers to the

[1]School and Graduate Institute of Physical Therapy, College of Medicine, Taipei; [2]Department of Special Education, National Taiwan Normal University, Taipei, [3]Department of Psychology, National Taiwan University, Taipei; [4]Department of Pediatrics, National Taiwan University Hospital, Taipei, Taiwan.
Received: 30 October 2004 Revised: 3 January 2005 Accepted: 12 April 2005
Reprint requests and correspondence to: Associate Professor Hua-Fang Liao, School and Graduate Institute of Physical Therapy, College of Medicine, National Taiwan University, 7 Chung-Shan S. Road, Taipei, Taiwan.

relationship between its scores and criterion measures of similar constructs at the same time.[8] The CDIIT has acceptable test-retest reliability, content validity, and construct validity.[9] However, its concurrent validity has not been reported.

To provide evidence of the concurrent validation of a test, Pearson product-moment correlation analyses between the test and a criterion test are frequently applied.[8] In some studies, correlations between tests are analyzed by using the developmental age (DA), which is the DA-equivalent score calculated from a norm, or the DQ, which is the derived score from the mean and standard deviation (SD) of the norm sample.[11,12] The correlation coefficient is only one of many factors that determine the degree to which a test may improve or detract from the quality of decisions. To fully evaluate the effect of a test on decisions, one must also understand the possible outcomes of decisions.[13] Classification agreement can be expressed as agreement percentages (true-positive and true-negative percentages) and κ coefficients. The purpose of this study was to investigate the relationship between DQs and DAs of the CDIIT and the BSID-II and to analyze the classification agreement between these scales in preterm infants.

Methods

Reliability pilot study

Because the reliability of the CDIIT has not been established for preterm infants, we conducted a pilot reliability study in 16 preterm infants (aged 6-18 months) before the concurrent validity study. To determine inter-rater reliability, 2 examiners simultaneously tested and recorded scores for each child. In the investigation of test-retest reliability, each child received the test twice from the same tester, with the second test performed within 3 days of the first. The intraclass correlation coefficient test (model 2, 1) was used for statistical analysis.[14] Intraclass correlation coefficients for the cognitive, motor, and whole-test DAs were, respectively, 0.96, 0.97, and 0.97 ($p < 0.01$) for test-retest reliability and 0.93, 0.99, and 0.99 ($p < 0.01$) for inter-rater reliability. Intraclass correlation coefficients of the cognitive, motor, and whole-test DQs were, respectively, 0.76, 0.86, and 0.78 for test-retest reliability, and 0.62, 0.91, and 0.95 for inter-rater reliability.

Subjects

This study was reviewed and approved by the Ethics Committee of National Taiwan University Hospital. We recruited 160 preterm infants with a corrected age of 6-18 months from 2 medical centers in Taipei. The inclusion criteria were birth at a gestational age of 37 weeks or younger and written informed consent from a parent. Infants with a serious medical illness and those who were hospitalized were excluded.

Anatasi and Urbina subdivided the first 5 years of life into infancy (approximately 0-18 months) and a preschool period (18-60 months).[8] Infants change more rapidly and more dramatically in development than preschoolers. Early development scores may be poor predictors of later scores because of environmental factors and the discontinuity hypothesis.[15] To investigate the concurrent validity of the CDIIT in the preschool age group, therefore, we had the same study procedures for these children at the age of 21-40 months.

A total of 141 children (88%) completed the secondary tests at 21-40 months; most of them (135) were aged of 24-35 months. The corrected age of the infants was 9.3 ± 3.5 months (mean ± SD; range, 6-18 months; 84 boys, 76 girls). Their gestational age was 33.3 ± 3.2 weeks (range, 22-37 weeks), birth weight was 1952.0 ± 621.7 g (range, 478-3558 g), and birth length was 42.3 ± 5.2 cm. During the preschool period, the infants were aged 27.6 ± 2.9 months (range, 21-40 months; 72 boys, 69 girls). Clinical diagnoses at birth were retinopathy of prematurity (n = 2), patent ductus arteriosus (n = 1), tetralogy of Fallot (n = 1), atrial septal defect (n = 1), congenital scoliosis (n = 1), cleft palate (n = 1), and inguinal hernia (n = 1).

Procedure

After the children were allowed to familiarize themselves with the testing environment, their developmental history and demographic data were recorded. One tester administered the CDIIT with the BSID-II. This tester had been trained by a licensed clinical psychologist and had achieved 95% agreement with the psychologist on BSID-II testing before data collection. Using the CDIIT, the tester had achieved 98% agreement with an experienced physical therapist, who had also trained the tester. Most of the testing was done at the pediatric assessment laboratory of the School and Graduate Institute of Physical Therapy of National Taiwan University and took about 40-80 min when the subjects were tested at 6-18 months and 60-90 min at 21-40 months. The subjects' mother or a familiar adult was present during testing.

Instruments

The CDIIT assessed cognitive, motor, language, self-help, and social development, providing DQs and DAs derived from the 5 subtest scores and whole-test scores.[16] The item pool of the CDIIT was generated from developmental theories or modified from items of various measures. Items such as chopsticks use were designed to be culturally sensitive. After a pilot study, 343 best-fit items were selected on the basis of Rasch analysis results, item difficulty, and clinical judgment. The standardization sample of 3703 infants (1832 boys, 1871 girls; aged 3-72 months) was randomly selected by age, gender, and geographic regions in Taiwan. Urban versus rural residence was also considered.[9]

One tester administered all items in the cognitive subtest, motor subtest, and parts of the language subtest of the CDIIT. The social and self-help subtests appeared as self-report items and were completed by the child's main caregiver.[16] Test items were scored 0 or 1, where 1 indicated success during the test or observation by the caregiver.[16] DQs of the CDIIT were calculated for 1-month intervals in children younger than 24 months.[16] After 24 months, DQs were obtained according to the norm of 3-month increments. The mean DQ was 100 ± 15. According the CDIIT manual, a DQ of 85 or above was within normal limits, 70-84 was borderline, and < 70 indicated delayed performance. Our previous research showed that the best cut-off point for identifying children with motor disabilities on the motor subtest was a DQ of 70, which had 87% sensitivity and 97% specificity.[17]

The BSID-II, with its mental and motor scales, was designed to assess infants aged 0-42 months. The norm of the BSID-II was established on developmental data from 1700 American children randomly selected for race/ethnicity, major geographic region, and parental education. Less than 3.7% of the children in the BSID-II sample were Asian.[18] High internal consistency, stability coefficients, and inter-scorer reliabilities were also reported. The BSID-II also provides information about content, construct, and discriminate validity.[18] Reliabilities of the BSID-II were high when it was used in Taiwanese children with developmental delays.[19]

The mental and motor scales of the BSID-II yield standardized scores: the Mental Developmental Index (MDI) and the Psychomotor Developmental Index (PDI), respectively. On the BSID-II, a child received credit only for items performed correctly. The MDI or PDI of children aged 36 months or younger was scored in their appropriate age categories of 1-month increments. After 36 months, the MDI or PDI was obtained according to the norm of 3-month increments. The mean MDI or PDI was 100 ± 15. An MDI or PDI score of 115 or above indicated accelerated performance, 85-114 within normal limits, 70-84 mildly delayed performance, and 69 or below significantly delayed performance.[18] Table 1 compares the CDIIT and the BSID-II. Both tests provide motor and mental DQs for children aged 3-42 months, and their years of standardization are near. Their sampling areas are different.[9,16,18]

Data analysis

For each infant, DA and DQ were calculated for the whole test and for the CDIIT and the BSID-II subtests. To obtain the DQ, the corrected age was used for preterm infants with a chronologic age younger than 24 months, and the uncorrected age was used for

Table 1. Comparison of the Comprehensive Developmental Inventory of Infants and Toddlers (CDIIT) and the Bayley Scales of Infant Development-II (BSID-II).

	CDIIT*	BSID-II†
Sample age	3-71 months	1-42 months
Subtests	Cognitive	Mental
	Motor–gross motor, fine motor	Motor
	Language	
	Self-help	
	Social	
Behavior rating scale	Yes	Yes
Number of standardized samples	3703	1700
Percentage of preterm babies	4%	0%
Sampling area	Taiwan	USA
Year of standardization	1995	1993
Number of samples by age range	3-71 months: 3703	
	3-42 months: 2815	1-42 months: 1700
	6-18 months: 1098	6-18 months: 600
	21-40 months: 804	21-40 months: 500
Number of items in total test/mental scale/	3-71 months: 343/81/97	
motor scale by age range	3-42 months: 229/36/71	1-42 months: 289/178/111
	6-18 months: 84/12/36	6-18 months: 95/58/37
	21-40 months: 96/15/21	21-40 months: 41/24/17
Scale scores	Developmental Quotient, developmental age, percentile	Developmental Index (95% CI, 90% CI), developmental age

* Summarized from Wang et al[9] and the CDIIT manual.[16]
† Summarized from the Bayley II manual.[18]
CI = confidence interval.

Table 2. Developmental ages and developmental quotients estimated by the Comprehensive Developmental Inventory for Infants and Toddlers (CDIIT) and the Bayley Scales of Infant Development-II (BSID-II) for preterm infants (mean ± standard deviation).

	Developmental age		Developmental quotient	
	6-18 months (n = 160)	21-40 months (n = 141)	6-18 months (n = 160)	21-40 months (n = 141)
CDIIT				
Cognitive subtest	9.4 ± 3.8	22.1 ± 4.4	98.3* ± 13.2	80.7‡ ± 11.7
Motor subtest	9.6 ± 4.1	22.4 ± 4.2	101.3† ± 16.5	77.8§ ± 11.9
Gross motor	9.5 ± 4.1	21.5 ± 4.7	100.7 ± 17.1	77.5 ± 14.4
Fine motor	9.8 ± 4.6	24.4 ± 6.1	100.5 ± 16.7	85.4 ± 13.7
Whole test	9.5 ± 3.6	23.6 ± 4.7	102.4 ± 17.4	86.3 ± 13.4
BSID-II				
Mental scale	8.6 ± 3.6	23.3 ± 3.9	94.0 ± 13.6	74.3 ± 19.5
Motor scale	8.8 ± 3.6	23.4 ± 4.2	94.5 ± 16.5	74.0 ± 18.7

* CDIIT cognitive DQ *vs* BSID-II MDI at 6-18 months, $t = 4.44$, $p < 0.001$, by paired *t* test.
† CDIIT cognitive DQ *vs* BSID-II MDI at 21-40 months, $t = 5.26$, $p < 0.001$, by paired *t* test.
‡ CDIIT motor DQ *vs* BSID-II PDI at 6-18 months, $t = 7.71$, $p < 0.001$, by paired *t* test.
§ CDIIT motor DQ *vs* BSID-II PDI at 21-40 months, $t = 2.34$, $p < 0.05$, by paired *t* test.
DQ = developmental quotient; MDI = Mental Developmental Index; PDI = Psychomotor Developmental Index.

those older than 24 months.[7] In the BSID-II manual, DQs are 50-150, whereas CDIIT DQs are 54-146. For infants with test scores outside these ranges, either the lowest or highest score was assigned.

Concurrent validities were examined using classification agreement analysis or correlation analysis (with the Pearson product-moment correlation coefficient).[8] The magnitude of each correlation coefficient was interpreted using the criteria for examining validity proposed by Domholdt.[20] The strength of the correlation between 2 test scores was considered very high if the correlation coefficient (r) was 0.90 or higher, high at 0.70-0.89, moderate at 0.50-0.69, low at 0.26-0.49, and very low at < 0.26.[20] To examine the classification agreement between the CDIIT and the BSID-II, the cross-tab of developmental delay (DQ < 70, < 2 SDs below the mean) and no obvious developmental delay (DQ ≥ 70, ≥ 2 SDs below the mean) for each test and the κ coefficient statistics were used. The strength of agreement was slight for κ = 0.00-0.20, fair for κ = 0.21-0.40, moderate for κ = 0.41-0.60, substantial for κ = 0.61-0.80, and almost perfect for κ = 0.81-1.0.[20] McNemar test was used to examine the symmetry of the classification distribution.

All statistical analyses were performed by using the Statistical Package for Social Sciences (SPSS) version 10.0 (SPSS, Chicago, IL, USA). The criterion used for statistical significance was $\alpha = 0.05$ for 2-tailed tests.

Results

CDIIT and BSID-II results of DA and DQ for preterm infants

Table 2 shows CDIIT and BSID-II estimates of the DAs and DQs for the sample. The mean DQ estimated by the CDIIT whole-test was 102.4 at 6-18 months, and 86.3 at 21-40 months. The mean DQs of the CDIIT were higher than those of the BSID-II at both ages ($t = 2.34$-7.71, $p < 0.05$).

In addition, children were classified as delayed (DQ < 70), borderline (70-84) or normal (> 85) according to their DQ results. At 6-18 months, 4 infants (3%) had whole-test CDIIT DQs in the delayed range, 20 (13%) had borderline DQs, and 136 (84%) had normal DQs. MDI scores were in the delayed range in 3 infants (2%), borderline in 13 (8%), and normal in 142 (90%). PDIs were delayed in 4 infants (3%), borderline in 32 (20%), and normal in 124 (77%).

At 21-40 months, whole-test CDIIT DQs were delayed in 10 infants (7%), borderline in 59 (42%), and normal in 72 (51%). MDIs were delayed in 35 infants (25%), borderline in 64 (45%), and normal in 42 (30%). PDIs were delayed in 42 infants (30%), borderline in 58 (41%), and normal in 41 (29%).

Correlations of DAs and DQs between the CDIIT and the BSID-II

Table 3 shows the correlation coefficients between DAs and DQs of the CDIIT and BSID-II. The cognitive, motor, and whole-test DAs for the CDIIT were highly correlated with DAs for the mental and motor scales of the BSID-II at 6-18 months ($r = 0.91$-0.97) and moderately to highly correlated at 21-40 months ($r = 0.58$-0.83). DAs between the 2 motor scales and between the 2 mental scales were highly correlated.

Motor and mental DQs were moderately to highly correlated at 6-18 and 21-40 months. In addition, cognitive DQs on the CDIIT had a high correlation with MDI on the BSID-II but a lower correlation with PDI in both age ranges. Motor DQs on the CDIIT

Table 3. Correlation coefficients between developmental ages (DAs) and developmental quotients (DQs) of the Comprehensive Developmental Inventory for Infants and Toddlers (CDIIT) and the Bayley Scales of Infant Development-II (BSID-II).

BSID-II	CDIIT				
	Cognitive	Motor	Gross motor	Fine motor	Whole test
6-18 months DAs					
Mental	0.93*	0.94*	0.90*	0.93*	0.93*
Motor	0.91*	0.97*	0.95*	0.91*	0.93*
21-40 months DAs					
Mental	0.83*	0.73*	0.53*	0.62*	0.79*
Motor	0.58*	0.80*	0.72*	0.58*	0.69*
6-18 months DQs					
MDI	0.60*	0.52*	0.42*	0.53*	0.55*
PDI	0.56*	0.77*	0.73*	0.50*	0.64*
21-40 months DQs					
MDI	0.73*	0.54*	0.34*	0.51*	0.66*
PDI	0.48*	0.67*	0.60*	0.43*	0.57*

* $p < 0.001$ (2-tailed), by Pearson's product moment test.
MDI = Mental Developmental Index; PDI = Psychomotor Developmental Index.

had a high correlation with PDI on the BSID-II but a lower correlation with MDI. In both age ranges, whole-test DQs for the CDIIT were moderately correlated with MDI or PDI on the BSID-II. Correlation coefficients for DAs and DQs on the BSID-II motor scale were higher with the CDIIT gross motor test than with the CDIIT fine motor test. The fine motor subdomain of the CDIIT was better correlated with the mental scale than with the motor scale of the BSID-II.

Classification agreement and test of symmetry

Table 4 shows the classification agreements between DQs from the CDIIT and the BSID-II. At 6-18 months, agreement was 99%, with $\kappa = 0.80$-0.85. At 21-40 months, agreement between CDIIT and BSID-II was 81-84%, with $\kappa = 0.44$-0.57. Classification agreement between the 2 motor scales or the 2 mental scales was generally high at 6-18 months and moderate at 21-40 months. McNemar analysis revealed a significantly asymmetric classification between only the motor DQ on the CDIIT and the PDI on the BSID-II at 21-40 months. Eighteen children (13%) had a higher motor level on the CDIIT than on the BSID-II, and 5 (4%) had a lower motor level on the CDIIT than on the BSID-II. That is, higher classification was found on the CDIIT motor DQ than on the PDI of BSID-II at 21-40 months.

Discussion

This study supports the concurrent validity of the motor and cognitive subtests and whole CDIIT and BSID-II for preterm children aged 6-40 months. The significant correlation between these 2 scales suggests that the new test (CDIIT) has the same traits as the reference test (BSID-II) in the investigated area. The CDIIT now has information for content, concurrent, and construct validity that fulfills the standard recommended by the American Psychological Association.[10] These findings indicate that the CDIIT is an appropriate instrument for the early identification of developmental delay in infants and toddlers in clinics. This study also found that the CDIIT DQs were significantly higher than the BSID-II DQs, and the CDIIT motor DQs tended to classify children in higher developmental levels than the BSID-II. This tendency should be considered when comparing DQs obtained from these 2 scales in clinical or in research settings.

The different correlation strength of the fine and gross motor DQs with the MDI and PDI of the BSID-II observed in this study has been previously reported.[21,22] The content of the items of the subtests may explain this finding. The BSID-II has more items on gross motor skills and fewer on fine motor skills in its motor scale,[18] and items on the mental scale of the BSID-II and the fine motor subdomain of the CDIIT overlap.[16,18] Items such as "reaches for the cube," "transfers object from hand to hand," "imitates crayon stroke," "places pegs," and "builds towers of cubes" are on the mental scale of the BSID-II but on the fine motor subdomain of the CDIIT. Washington et al[23] reported that motor skills in a 9-month-old preterm infant had been overestimated as being within normal limits of the PDI because of his relatively strong fine motor skills. A motor test with separate fine and gross motor subtests may be better for measuring motor disability than one which includes these items in a single test. Furthermore, the motor subtest of the CDIIT was completed in 10-15 min compared with 20-40 min for the BSID-II motor scale. This may provide the former some advantage over the latter in clinical use.

Table 4. Classification agreement between the developmental quotient (DQ) of the mental subtest/motor subtest of the Comprehensive Developmental Inventory for Infants and Toddlers (CDIIT) and the Mental Developmental Index (MDI)/ Psychomotor Developmental Index (PDI) of the Bayley Scale of Infant Development-II (BSID-II).

A. (6-18 months)	BSID-II MDI		
	< 70	≥ 70	Total
CDIIT cognitive DQ			
< 70	2	0	2
≥ 70	1	157	158
Total	3	157	160

Agreement 99%, κ = 0.80 ($p < 0.001$), McNemar test $p = 1$

B. (6-18 months)	BSID-II PDI		
	< 70	≥ 70	Total
CDIIT motor DQ			
< 70	3	0	3
≥ 70	1	156	157
Total	4	156	160

Agreement 99%, κ = 0.85 ($p < 0.001$), McNemar test $p = 1$

C. (21-40 months)	BSID-II MDI		
	< 70	≥ 70	Total
CDIIT cognitive DQ			
< 70	17	9	26
≥ 70	18	97	115
Total	35	106	141

Agreement 81%, κ = 0 .44 ($p < 0.001$), McNemar test $p = 0.12$

D. (21-40 months)	BSID-II PDI		
	< 70	≥ 70	Total
CDIIT motor DQ			
< 70	24	5	29
≥ 70	18	94	112
Total	42	99	141

Agreement 84%, κ = 0.57 ($p < 0.001$), McNemar test $p = 0.01$

DQ classification for the CDIIT motor subtest tended to be higher than that for BSID-II PDI at 21-40 months. Previous study found that the classification between the PDMS and the motor scale of the BSID-II differed.[12] A couple of reasons may explain this asymmetry. First, the tests are administered and scored differently. For the CDIIT, scores can accumulate because of information reported by the caregiver, and the child could easily achieve a higher score on the motor subtest of the CDIIT than on the BSID-II. On the cognitive subtest, special testing materials, such as the specially designed picture book, needed for administering the CDIIT are usually not available at home; therefore, caregiver information is not likely to increase the score. Second, 4% of the norm sample for the CDIIT were preterm babies, but the norm sample for the BSID-II were only healthy babies. Preterm infants are usually at increased risk for poor motor developmental outcomes,[5,6,24] and the relatively high norm score of the BSID-II for preterm children may cause a lower standard score. Eighteen children with motor delay on the BSID-II were normal on the CDIIT at 21-40 months. The BSID-II PDIs of these children were 50-69, but their CDIIT motor DQs were 71-89. Because of measurement error, the BSID-II provides the 95% confidence intervals (CIs) of its developmental index for both the mental and motor scales to assess the precision of any estimate of diagnostic accuracy.[18,25] When the PDI CIs were considered, the upper limit of the 95% CI of these 18 children was 68-83, and most (n = 16) were > 70. Careful clinical judgment is required to prevent mislabeling or inappropriate referral when children have a DQ near the cut-off point for developmental delay or when the 95% CI covers the cut-off point.

Studies have also shown a high to moderate correlation for various developmental tests.[12,18,21,22,26] Variation in correlation coefficients depends on the similarities and differences in the content of the criterion tests[18] and on the variability of the sampled group.[8] Therefore, the concurrent validity coefficients are not comparable.

One limitation of this study was the small number of children with developmental delay at 6-18 months. Only 3 preterm infants (2%) had an MDI of < 70. This small sample may have had little effect on the classification agreement percentage, but a greater effect on the κ coefficient. For example, if the number of children that were classified as delayed in both scales changed from 2 to 1, then the κ coefficient might have changed from 0.80 to 0.68. Therefore, a larger sample is needed to estimate the κ coefficient between the 2 scales. Another limitation of this study was that DQs of the BSID-II were derived from the US norm instead of the Taiwanese norm. Rates of child development are different in various cultures.[27,28] Further studies are needed to investigate the diagnostic accuracy and best cut-off point of the BSID-II for use in Taiwan.

In conclusion, this study has demonstrated that the CDIIT has acceptable concurrent validity compared with the BSID-II in preterm infants. Validity is specific in that a test may be valid for one group (or age range) of children but not in another.[13] Further studies are needed to determine the concurrent validity of the CDIIT for children with disabilities. The diagnostic accuracy of the whole CDIIT, such as its sensitivity and specificity, should be also investigated.

ACKNOWLEDGMENT: This study was supported by a grant from the Department of Health, Executive Yuan of the Republic of China (DOH90-TD-1141).

References

1. Boyle CA, Decoufle P, Yeargin-Allsopp M: Prevalence and health impact of developmental disabilities in US children. *Pediatrics* 1994;93:399-403.
2. Lavigne JV, Binns HJ, Christoffel KK, et al: Behavioral and emotional problems among preschool children in pediatric primary care: prevalence and pediatricians' recognition. *Pediatrics* 1993;91:649-55.
3. Bierman JM, Connor A, Vagge M, et al: Pediatricians' assessments of the intelligence of two-year-olds and their mental test scores. *Pediatrics* 1964;33:680-90.
4. Villar J, Belizan JM: The relative contribution of prematurity and fetal growth retardation to low birth weight in developing and developed societies. *Am J Obstet Gynecol* 1982;143:792-8.
5. Bhutta AT, Cleves MA, Casey PH, et al: Cognitive and behavioral outcomes of school-aged children who were born preterm: a meta-analysis. *JAMA* 2002;288:728-37.
6. Chen PS, Jeng SF, Tsou KI: Developmental function of very-low-birth-weight infants and full-term infants in early childhood. *J Formos Med Assoc* 2004;103:23-31.
7. Lems W, Hopkins B, Samson JF: Mental and motor development in preterm infants: the issue of corrected age. *Early Hum Dev* 1993;34:113-23.
8. Anatasi A, Urbina S: Psychological Testing. 7th ed. Upper Saddle River, NJ: Prentice Hall; 1997:48-171,234-70.
9. Wang TM, Su CW, Liao HF, et al: The standardization of the comprehensive developmental inventory for infants and toddlers. *Psychol Testing* 1998;45:19-46. [in Chinese; English abstract]
10. American Psychological Association: Standards for Educational and Psychological Tests. Washington, DC: American Psychological Association, 1986.
11. Palisano RJ: Concurrent and predictive validities of the Bayley motor scale and the Peabody developmental motor scales. *Phys Ther* 1986;66:1714-9.
12. Provost B, Crowe TK, McClain C: Concurrent validity of the Bayley Scales of Infant Development II Motor Scale and the Peabody Developmental Motor Scales in two-year-old children. *Phys Occup Ther Pediatr* 2000;20:5-18.
13. Murphy KR, Davidshofer CO: Psychological Testing, Principles and Applications. 5th ed. Upper Saddle River, New Jersey: Prentice Hall; 2001:18- 85,169-91.
14. Shrout PE, Fleiss JL: Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979;86:420-8.
15. Aylward GP: Conceptual issues in developmental screening and assessment. *Dev Behav Pediatr* 1997;18:340-9.
16. Wang TM: The Comprehensive Developmental Inventory for Infants and Toddlers Manual [in Chinese]. Revised version. Taipei: Special Education Division, Ministry of Education, 2003.
17. Wu HY, Liao HF, Yao G, et al: Diagnostic accuracy of the motor subtest of Comprehensive Developmental Inventory for Infants and Toddlers (CDIIT) and the Peabody Developmental Motor Scales-Second Edition (PDMS-2) for Preschool Children. *Formos J Med* 2005;9:312-22. [in Chinese; English abstract]
18. Bayley N: Bayley II. San Antonio: Psychological Corporation, 1993.
19. Huang HL, Shieh YL, Jong YJ, et al: Applicability of BSID-II in diagnosing developmental delay at Kaohsiung area. *Kaohsiung J Med Sci* 2000;16:197-202.
20. Domholdt E: Physical Therapy Research, Principles and Applications. Philadelphia: WB Saunders; 2000:350-78.
21. Mullen EM: Mullen Scales of Early Learning. AGS Edition. Circle Pines, MN: American Guidance Service; 1995:60-4.
22. Folio MR, Fewell RR: Peabody Developmental Motor Scales. Chicago: Riverside; 1983:118-9.
23. Washington K, Scott DT, Johnson KA, et al: The Bayley Scale of Infant Development-II and children with developmental delays: a clinical perspective. *J Dev Behav Pediatr* 1998;19:346-9.
24. Jeng SF, Yau KIT, Liao HF, et al: Prognostic factors for walking attainment in very low-birth weight preterm infants. *Early Hum Dev* 2000;59:159-73.
25. Fritz JM, Wainner RS: Examining diagnostic tests: an evidence-based perspective. *Phys Ther* 2001;81:1546-64.
26. Folio MR, Fewell RR: Peabody Developmental Motor Scales. Examiner's Manual. 2nd ed. Austin: PRO-ED; 2000:53-68.
27. Solomons G, Solomons HC: Motor development in Yucatecan Indians. *Dev Med Child Neurol* 1975;17:41-6.
28. McClain C, Provost B, Crowe TK: Motor development of two-year-old typically developing Native American children on the Bayley Scales of Infant Development II Motor Scale. *Pediatr Phys Ther* 2000;12:108-13.