The Reliability of a Pediatric Balance Scale Based on the Raters' Clinical Work Experience and Test Experience

The Journal Korean Society of Physical Therapy

Gi-Won Kim, PT, PhD; Joo-Yeon Ko, PT, PhD[1], Soon-Gi Baek[2]

Department of Physical Therapy, Dong Nam Health College; [1]Department of Physical Therapy, Bundang CHA General Hospital; [2]Department of Leisure & Sports, Chongju University

Purpose: To investigate the rater reliability of a Pediatric Balance Scale (PBS) for children with cerebral palsy, and to investigate possible differences among raters according to their clinical work experience and testing experience.

Methods: Study participants included 18 children with spastic cerebral palsy who could walk. They were instructed by pediatric physical therapists, two of whom had ten years of clinical work experience and two who had less than one year of experience. The children's ability to achieve physical balance was videotaped for PBS items. The raters watched the tapes and evaluated each child twice. Rater reliability was analyzed using the intraclass correlation coefficient (ICC). Differences between experienced and novice raters were analyzed using a paired t-test. The statistical significance level was set to 0.05.

Results: The total PBS scores averaged 45.78~48.00 and 45.72~47.67 for first and second tests. Intra-rater reliability was very high (ICC=0.89~0.99), and the repeated measurement coincidence was high ($p>0.05$). Inter-rater reliability was high (ICC=0.83~0.84), but there was a bit of a difference in the coincidence ($p<0.05$). The experienced raters' reliability and coincidence were higher than those of the novices, and there were differences in reliance and coincidence between experienced and novice raters ($p<0.05$).

Conclusion: Inter-rater and intra-rater reliability is very high. However, rater reliability showed defferences depending on clinical work experience and testing experience. When testing pediatric patients with the PBS, the rater's clinical experience and test experience may affect the test results.

Keywords: Balance, Postural control, Pediatric balance scale (PBS), Rater reliability

Received: September 17, 2010
Revised: November 30, 2010
Accepted: December 6, 2010
Corresponding author: Joo-Yeon Ko, 7806218@hanmail.net

I. Introduction

Cerebral palsy refers to a non-progressive developmental disorder of motor skills that affects movement and posture, and is caused either by brain damage, or immaturity of the brain.[1] A child with cerebral palsy experiences muscle coordination problems, difficulties in integrating sensory data, and functional limitations,[2] and this causes problems in muscle tone, which affect functional balance abilities, and also causes abnormal postural control.[3] Children with developmental disorders in motor skills lack both the compensatory postural reaction, and anticipatory postural control, and this delays the children's motor skill gathering or development, and also delays motor development.[4,5] Thus children with cerebral palsy have balance strategies that are different from typically developing children.[3]

Postural control or balance refers to the ability to maintain the body in equilibrium,[6] and postural stability is defined as one's innate ability to maintain a specific balance position.[7] Maintaining postural control is essential in carrying out skillful movement, and requires the ability to pass a base of support and maintain the center of gravity in order to complete simple or complex gross and fine motor tasks.[8] Balance is an important part of gross motor function, and problems in balance create problems in functional tasks required for carrying out activities of daily living.[2] A deficit in balance is the most representative problem that physical therapists experience, and it is an important issue in rehabilitation

and intervention.[2,3,9]

In order to evaluate balance ability, one normally would use a posture stability test which measures the movement of the center of pressure within the base of support, and other various standardized tests. In order to test a child, a physical therapist or occupational therapist mostly uses tests that measure balance time and orientating response, such as a one-leg standing test, or a tilt-board balance test, and tests that examine postural stability in different sensory environments, such as the Pediatric Clinical Test of Sensory Interaction for Balance (P-CTSIB), or those that test the ability to maintain balance dynamically.[4,8] Moreover, some exams use items that test balance ability among standardized child development tests, such as the Bruininks-Oseretscky Test of Motor Proficiency (BOTMP), the Peabody Developmental Motor Scale (PDMS), and the Gross Motor Function Measure (GMFM).[3,10] One may also use methods that use tools such as an Active Balance System or a Smart Balance Master System,[7] in order to quantify balance control abilities. Such methods have to be carried out in a controlled laboratory condition, and sometimes requires expensive equipment.[3] The latter are not adequate for a normal clinical environment, and they cannot test functional aspects of balance. Therefore, these methods may be difficult to use with children with cerebral palsy when testing child developmental status.[10]

An assessment tool that was invented to evaluate functional aspects of balance control abilities includes the Pediatric Balance Scale (PBS), which modified the Timed Up and Go (TUG), Functional Reach Test (FRT), and Berg Balance Scale (BBS) to befit children.[3,4,10] The Balance Evaluation Systems Test (BEST) was recently developed to fit the system model of motor control.[9] Such evaluation tools are for testing functional aspects of balance control with standardized tests, and were mainly developed for adults. PBS, especially, was developed as a modified BBS, which is for adults, to test functional balance abilities of children with motor disabilities, simply and quickly. Having a simple and easy measure of functional balance abilities, along with reliability and validity, is an essential condition for physical therapists who treat children with motor disorders and determine the results of interventions.[3,4] Franjoine and others, who invented the PBS, reported that the test-retest reliability was 0.87～1.0, and that the inter-rater reliability was 0.99.[10] In Korea, there only exists a study that was conducted using a Korean translation. They studied children with developmental disabilities and reported an inter-rater reliability range of 0.78～0.97;[11] for children with cerebral palsy test-retest reliability was 0.89～0.93, with >0.90 intra- and inter-rater reliabilities.[12]

The purpose of the current study was to compare rater reliability in the use of the Pediatric Balance Scale, which was developed to evaluate balance abilities of children with cerebral palsy. We analyzed reliability with respect to the tester's clinical work experience and testing experience.

II. Methods

1. Participants and testers

The participants were selected from among children being treated for cerebral palsy in two general hospitals in Gyeonggi-do and one Rehabilitation Medical Treatment hospital in Cheonan. These children had spastic cerebral palsy and could walk. With the video records as criteria, a sample of 19 children were selected from among the 24 children who had participated in prior research studies conducted by these investigators. Parental consent was obtained.

The testers included 4 pediatrics physical therapists who treat children with disorders at general hospitals and child development centers, 2 of whom had 10 years of clinical work experience, and 2 of whom had less than 1 year of experience. Therapists with a lot of clinical work experience had previous experience in testing children using the Pediatric Balance Scale; those who had less work experience had no prior test experience. The tests were conducted between April 1st and May 30th, 2010.

The research participants' general characteristics, the characteristics of their disorders, and the general characteristics of the raters are shown in Table 1.

2. Methods

1) Instruments

In order to select the participants, the researchers used the Gross Motor Function Classification System (GMFCS), which was first conceptualized by Russell and others, and developed by Palisano and others.[13] The researchers chose children at level 1 and 2, who could stand erect and walk alone.

To test balance abilities, the testers used the Pediatric Balance Scale, invented by Franjoine and others,[10] and translated into

Table 1. The research participants' general characteristics and the characteristics of their disorder

Characteristics		Test-retest and Inter-rater reliability (n=24)	
		Frequency	Mean±SD
Sex	Male (n)	10	
	Female (n)	8	
Age (y)			10.50±2.92
Height (cm)			137.73±12.99
Weight (kg)			35.23±10.13
GMFCS level	Level Ⅰ	11	
	Level Ⅱ	7	
Type of CP	hemiplegia	7	
	diplegia	11	
Hand function	Level Ⅰ/Ⅱ/Ⅲ/IV/V	9/5/4/0/0/	
Communication	Level Ⅰ/Ⅱ/Ⅲ/IV	18/0/0/0	
GMFM total score			95.75±2.60

Korean by Ko et al,[11] after revising and supplementing the Korean version. Among the total of 14 items, the 14th item, "reaching forward with outstretched arm" was difficult to evaluate by videotaping it with a 2D camcorder, and thus was excluded, as was done in Franjoine's study.

2) Procedure

By interviewing the participants' parents, the researchers found the participants' general characteristics. The therapists who treat the participants implemented the Gross Motor Function Measure, and evaluated hand function and communication abilities.

For the Pediatric Balance Scale, two researchers videotaped each item in order. When videotaping, they gave the participants enough explanations, demonstrations and the opportunity to practice. The fourteen items of the Pediatric Balance Scale were graded on a scale of 0 to 4. Only thirteen items were measured for a maximum possible total score of 52. Videotaping was carried out for all thirteen items in order. When the participants received a score of 4 on the first performance, they moved on to the next item, and the participants were allowed to try for a maximum of three times. If participants were not able to try or succeed by the third attempt, he was graded as a 0. Assessments were done by the four testers, each watching the videotaped data. This was conducted as 1st and 2nd assessments with a one week interval in between. Before grading the items, the 2 testers with more work experience discussed grading standards, explained the standards to the testers with less test experience, and carried out a training session once.

3. Data analysis

Research data was organized using Excel, and the researchers used SPSS for Windows version 12.0 for statistical analysis. The level of significance for the statistical test was set at $p<0.05$. Inter-rater reliability among the four testers, and intra-rater reliability between the first and second assessment by the same tester were analyzed using the Intraclass Correlation Coefficient (ICC3,1). Paired t-tests were used to compare those with more clinical work experience and those with less experience.

III. Results

The general characteristics and disorder characteristics of children with cerebral palsy that participated in this study on the reliability of the Pediatric Balance Scale are as follows (also see Table 1).

We calculated the mean and standard deviation for each item from the results of tests 1 and 2 conducted by the 4 raters. We also calculated the totalscore for the 13 items. The total score on the PBS for the 18 participants was 45.78～48.00 for the first evaluation and between 45.72 and 47.67 for the second evaluation (Table 2).

When we analyzed rater reliability for the PBS using the Interclass correlation coefficient, the intra-rater reliability was shown to be 0.89～0.99. Because reliability based on the ICC does not show correspondence, we did a repeated analysis of variance and found that differences between the first and second assessment by all four raters weren't statistically significant ($p>0.05$). For inter-rater reliability, the ICC for the first

Table 2. The participants' total score on the PBS, and descriptive statistics of scores for each item

Items	1st Measure				2nd Measure			
Rater	Rater A	Rater B	Rater C	Rater D	Rater A	Rater B	RaterC	Rater D
	Experience		Novice		Exprerience		Novice	
	M (SD)							
1. Sitting to standing	3.67 (0.59)	4.00 (0.00)	3.67 (0.49)	3.94 (0.24)	3.67 (0.59)	4.00 (0.00)	3.67 (0.49)	3.94 (0.24)
2. Standing to sitting	3.83 (0.71)	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)	3.83 (0.71)	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)
3. Transfers	3.89 (0.32)	3.89 (0.32)	4.00 (0.00)	3.89 (0.32)	3.89 (0.32)	3.83 (0.38)	4.00 (0.00)	4.00 (0.00)
4. Standing unsupported	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)	3.94 (0.24)
5. Sitting unsupported	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)
6. Standing with eye closed	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)
7. Standing with feet together	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)
8. Standing with one foot in front	1.89 (1.23)	1.83 (1.25)	2.22 (1.26)	2.28 (1.27)	1.89 (1.23)	1.72 (1.07)	2.06 (1.31)	1.89 (1.08)
9. Standing on one foot	2.50 (1.04)	2.89 (1.18)	2.89 (1.28)	2.56 (1.20)	2.50 (1.04)	2.83 (1.20)	2.89 (1.28)	2.72 (1.18)
10. Turning 360 degree	3.39 (0.78)	3.39 (1.04)	3.78 (0.65)	3.61 (0.78)	3.39 (0.98)	3.22 (1.00)	3.67 (1.03)	3.33 (0.97)
11. Turning to look behind	3.56 (0.51)	3.44 (0.78)	3.94 (0.24)	3.78 (0.43)	3.50 (0.62)	3.50 (0.71)	3.94 (0.24)	3.61 (0.78)
12. Retrieving object from floor	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)	4.00 (0.00)
13. Placing alternate foot on stool	3.06 (0.94)	3.17 (0.92)	3.50 (0.79)	3.17 (0.92)	3.06 (0.94)	3.17 (0.92)	3.44 (0.98)	3.11 (0.96)
PBS total score	45.78 (4.49)	46.61 (3.66)	48.00 (3.07)	47.22 (3.28)	45.72 (4.79)	46.28 (3.56)	47.67 (3.33)	46.56 (3.78)

M: Mean
SD: Standard deviation

Table 3. Intra-rater and inter-rater reliability of PBS

Intra-rater reliability		1st Measure M (SD)	2nd Measure M (SD)	ICC	95% C.I.	F-value	p-value
Experience	Rater A	45.78(4.49)	45.72(4.79)	0.99	0.98~0.99	0.19	0.67
	Rater B	46.61(3.66)	46.28(3.56)	0.91	0.79~0.97	0.90	0.36
Novice	Rater C	48.00(3.07)	47.67(3.33)	0.98	0.94~0.99	4.25	0.06
	Rater D	47.22(3.28)	46.56(3.78)	0.89	0.74~0.96	2.96	0.10

Inter-rater reliability	ICC	95% C.I.	F-value	p-value
1st Measure	0.83	0.70~0.92	7.04	0.00*
2nd Measure	0.84	0.70~0.93	4.84	0.01*

M: Mean
SD: Standard deviation
ICC: Intraclass correlation coefficient
95% C.I.: 95% confidence interval
*p<0.05

evaluation was 0.83; for the second evaluation it was 0.84. The results of repeated analysis of variance showed that all differences were statistically significant. There were differences in the coincidence between raters (Table 3).

After dividing the four raters into experienced and novice groups according to their clinical work experience and test experience, we compared between group differences in intra-rater reliability. On average, the reliability of the experienced raters was higher; the difference in average scores between the 1st and 2nd assessments weren't statistically significant, but the difference between the novices' first and second assessments were significant (p=0.04). Moreover, when we compared the inter-rater reliability of the experienced and novice raters, and the difference in average scores, we found that reliability increased as the assessments were repeated, and the differences in the average scores rated by the experienced and the novice raters were statistically significant for both first and second assessments (p<0.05)(Table 4).

IV. Discussion

A physical therapist should evaluate if a child has adequate functional balance capabilities for responding safely to daily necessities at home, school, and even in the local community.[10]

The Pediatric Balance Scale is a standardized tool for assessing functional balance that was developed by revising the Berg Balance Scale for evaluation of motor disorders of school-aged children. The rating standard of the Pediatric Balance Scale assesses both qualitative and quantitative evaluations, considering a variety of performances. This scale is adequate for evaluating functional balance capabilities of children with motor disorders because such variety is a characteristic of motor development.[12] Accurate and adequate assessment and evaluation are necessary procedures for constructing a specific and effective plan for treatment, and the reliability and validity of the tools for assessment evaluation are very important factors in all evaluations.

In this research, we determined the intra and inter-rater reliability of evaluating balance capabilities of children with cerebral palsy using the Pediatric Balance Scale, and compared differences in reliability according to the raters'clinical and test experience. The study by Franjoine et al, who developed the Pediatric Balance Scale and studied school-age children with light to medium motor disorders, reported test-retest reliability and inter-rater reliability.[10] The inter-rater reliability of the 10 pediatric physical therapists for 10 child participants showed a high Interclass Correlation Coefficient, one that was >0.99, and the results of repeated analysis of variance, which defines absolute reliability, did not show a statistically significant difference, which revealed that the correspondence between raters was very high. For studies conducted in Korea, Ko et. al. reported an inter-rater reliability for total scores on the PBS for three groups with different motor disorders; the coefficient was 0.97 for children with cerebral palsy, which was a statistically significant difference.[11] Prior research by Ko and Kim, in which some of the

Table 4. Comparison of rater reliability according to raters' clinical and test experience and the average difference in total PBS score

1st vs 2nd Measure	1st Measure M (SD)	2nd Measure M (SD)	Mean difference	t-value	p-value	ICC (95% C.I.)
Experience	46.19(3.97)	46.00(4.01)	0.19	1.24	0.23	0.99 (0.96~0.99)
novice	47.61(3.11)	47.11(3.45)	0.50	2.19	0.04*	0.96 (0.87~0.98)
Experience vs novice	Mean difference		t-value	p-value	ICC	95% C.I.
1st Measure	-1.42		-3.04	0.01*	0.85	0.64~0.94
2nd Measure	-1.11		-2.97	0.01*	0.91	0.78-0.87

M: Mean
SD: Standard deviation
ICC: Intraclass correlation coefficient
95% C.I.: 95% confidence interval
*p<0.05

researchers of our study participated, measured test-retest reliability and rater reliability for child participants with cerebral palsy.[12] They found that inter-rater reliability was between 0.91 and 0.93, and intra-rater reliability was between 0.97 and 0.99. The inter-rater reliability of this research was 0.83～0.84; when we did repeated analysis of variance, there was a statistically significant difference, which means it was different from Franjoine's findings and similar to Ko's findings. Intra-rater reliability was 0.89～0.99, and the absolute reliability showed no significant difference. For calculations of the Interclass Correlation Coefficient, we followed Munro's[14] instructions and interpreted the inter-rater reliability of this research as high, and intra-rater reliability as very high. We found that the inter-rater reliability of the aforementioned prior studies were very high, while the inter-rater reliability of our study was the lowest among them. We think this can be because of the raters' clinical work experience. In Franjoine et al.'s research, experienced pediatric physical therapists with more than 25 years of clinical work experience participated in the assessment; in Ko et. al.'s research, physical therapists with 13 and 16 years of experience each, participated in the assessment. In our study, we attributed the low inter-rater reliability to having pediatric physical therapists with 10 years of clinical work experience as well as those with less than 1 year of experience participate in the assessment. Also in the research of Ko and Kim, therapists with 10 and 2 years of clinical experience participated in the assessment, and their inter-rater reliability was also low.

There was no prior research that would allow us to compare the total scores for the PBS after categorizing the results into those of experienced raters and those of novice raters (according to their clinical and test experience). So, we compared it with research results that used different types of assessments. In our study, when we compared averages for the first and second assessments, the ICC for the experienced raters was 0.99, and paired t-tests showed no statistically significant difference, which meant a very high correspondence. For novice raters, the ICC was 0.96. But when we compared the differences in average scores, there were statistically significant differences (p=0.04). When we compared experienced and novice raters, the inter-rater reliability for the first assessment was 0.85, and that of the second assessment was 0.91; the differences in the average score were all statistically significant (p=0.01). Thus, the correspondence of the assessments by experienced and novice raters can be considered low. Such results were similar to results of research that reported an increase in reliability with increasing clinical experience and training. That study used the Tardieu Scale which evaluates spasticity of children with cerebral palsy. It was similar to our study in that the difference in the average intra-rater correspondence between experienced and novice raters was statistically significant.[15] Among the studies conducted in Korea, there was one that studied tester reliability of the Modified Ashworth Scale (MAS) and Modified Tardieu Scalde (MTS). It showed that there is almost nodifference in reliability according to clinical experience.[16] Moreover, in research that studied intra-rater reliability for the Tinetti Balance Score (BPOMA), which only evaluated the balance part of the Performance Oriented Mobility Assessment (POMA), there was a small difference in correspondence according to clinical experience.[17] Students without any experience also participated in this research, and they showed higher correspondence than the experienced participants in 4 of 8 items. There were 3 items for which those with 0～2 years of experience showed higher correspondence than those with 3～6 years of experience. Such contradictory results seem to be because of the differences in raters and participants, evaluation periods, and the experimental design. The raters who participated in the study by Gracies et al.[15] and evaluated children with cerebral palsy, consisted of 6 professionals in various fields, including a pediatrician. The research of Choi et al 16 evaluated patients with hemiparalysis who were evaluated by two physical therapists; Lisa et al's research17studied elderly participants who were 60～90 years old, living in nursing homes; 9 raters evaluated them over a two month period. These differences have great implications. Gracies et al[15] suggested that training is effective in increasing reliability; however, in our research, while the average did not change much when we repeated the tests conducted by experienced raters, the average did change as the novice repeated the tests. Thus, we think that one needs to consider the variables of training when conducting the evaluations. Other researches concluded that the rater's experience did not influence the reliability, but this is controversial, and seems to require more research. Another study about assessment and evaluation conducted by one of the researchers of this study suggested after reporting a very low reliability the need for educating and training raters with consistent standards and assessment methods.[18] Therefore, we think there needs to be more research conducted on changes in

reliability for assessments and evaluations according to different conditions such as raters' clinical experience, skills, and training. Limitations of this research include not being able to draw a variety of conclusions because the raters' clinical experience and test experience were tied as the same variable, and being able to measure the reliability for the total scores of PBS items but not for each question. Thus, various studies should be conducted to overcome these shortcomings. In addition, it would be good to expand the number of raters studied.

V. Conclusion

This research studied rater reliability for the Pediatric Balance Scale for children with cerebral palsy, and compared raters according to their clinical experience and test experience.

We found that intra-rater reliability is very high, and the correspondence of repeated tests is also high. Inter-rater reliability is high, but there is a difference in correspondence. Therefore, results between raters appeared to be non-corresponding. Novice raters showed higher reliability and correspondence than experienced raters; the reliability of the experienced raters and novice raters increased as the assessments were repeated; there were differences in the average scores.

Summing upPBS intra-rater reliability is high. However, differences were associated with clinical experience and measurements.

The raters'clinical and test experience may affect PBS evaluation results. More research should be conducted to compare PBS results with those of other evaluation tools, in order to support the conflicting results.

Author Contributions

Research design: Kim GW
Acquisition of data: Kim GW, Ko JY
Analysis and interpretation of data: Kim GW, Baek SG
Drafting of the manuscript: Kim GW, Baek SG
Research supervision: Ko JY

References

1. Wood E, Rosenbaum P. The gross motor function classification system for cerebral palsy: a study of reliability and stability over time. Dev Med Child Neuro. 2000;42(5):292-6.
2. Kembhavi G, Darrah J, Magill-Evans J et al. Using the berg balance scale to distinguish balance abilities in children with cerebral palsy. Pediatr Phys Ther. 200214(2):92-9.
3. Gan SM, Tung LC, Tang YH et al. Psychometric Properties of Functional Balance Assessment in Children with Cerebral Palsy. Neurorehabil Neural Repair. 2008;22(6):745-53.
4. Liao HF, Mao PJ, Hwang AW. Test-retest reliability of balance tests in children with cerebral palsy. Dev Med Child Neurol. 2001;43(5):180-6.
5. Donker SF, Ledebt A, Roerdink M et al. Children with cerebral palsy exhibit greater and more regular postural sway than typically develpoping children. Exp Brain Res. 2008;184(3): 363-70.
6. Shumway-Cook A, Woollacott MH. Motor control: theory and practical applications. 2nd ed. Baltimore, Lippincott Williams & Wilkins,2001:163-90.
7. Geldhof E, Cardon G, De Bourdeaudhuij I et al. Static and dynamic standing balance: test-retest reliability and reference values in 9 to 10 years old children. Eur J Pediatr. 2006; 165(11):779-86.
8. Atwater SW, Crowe TK, Deitz JC et al. Interrater and test-retest reliability of two pediatric balance tests. Phys Ther. 1990;70(2):79-87.
9. Horak FB, Wrisley DM, Frank J. The balance evaluation system test (BESTest) to differentiate balance deficits. Phys Ther. 2009;89(5):484-98.
10. Franjoine MR, Gunther JS, Taylor MJ. Pediatric balance scale: a modified version of the berg balance scale for the school-age child with mild to moderate motor impairment. Pediatr Phys Ther. 2003;15(2):114-20.
11. Ko MS, Lee NH, Lee JA et al. Inter-examiner reliability of the Korean version of the pediatric balance scale. Journal of the Korean Academy of University Trained Physical Therapists. 2008;15(1):86-95.
12. Ko JY, Kim GW. Test-retest, inter-rater, and intra-rater reliability of a pediatric balance scale in children with cerebral palsy. J Kor Soc Phys Ther. 2010;22(4):43-8.
13. Palisano R, Rosenbaum P, Walters S et al. Development and reliability of a system to classifygross motor function in children with cerebral palsy. Dev Med Child Neurol. 1997;39(5): 214-23.
14. Munro BH. Statistical methods for health care research. 3rd ed. New York, Lippincott Williams & Wilkins, 1997.

15. Gracies JM, Burke K, Clegg NJ et al. Reliability of the tardieu scale for assessing spasticity in children with cerebral palsy. Arch Phys Med Rehabil. 201091(3):421-8.
16. Choi YJ, Lee JA, Shin HK. Inter-rater and intra-rater reliability of the modified ashworth scale and modified tardieu scale:a comparison study. J Kor Soc Phys Ther. 2010; 22(4):29-33.
17. Cipriany-Dacko LM, Innerst D, Johannsen J et al. Interrater reliability of the tinetti balance scores in novice and experienced physicaltherapy clinicians. Arch Phys Med Rehabil. 1997; 78(10):1160-4.
18. Kim GW, Hong WS. A study of measurement methods for subtalar joint motion. J Kor Soc Phys Ther. 2010;22(4):57-64.