

물리치료과 학생들을 대상으로 아동균형척도의 항목별 측정자 신뢰도에 관한 연구

황룡 · 김기원†

동남보건대학교 물리치료학과, ¹수원여자대학교 물리치료과

The Study on the Rater Reliabilities of Physical Therapy Students Pediatric Balance Scales Items

Ryoung Hwang, PT, PhD, Gi-Won Kim, PT, PhD†

Department of Physical Therapy, Dongnam Health College

¹Department of Physical Therapy, Suwon Women's Collegey

Received: February 8, 2013 / Revised: February 18, 2013 / Accepted: February 19, 2013

© 2013 Journal of the Korean Society of Physical Medicine

| Abstract |

PURPOSE: The purpose of this study is to examine assessment of test-retest reliability for 13 items of pediatric balance scales, then the intrarater reliability among the raters.

METHODS: Study participants included 6 children with spastic cerebral palsy who could walk. Raters were consist of seventy-four, 2nd year physical therapy major students. The children's ability to achieve physical balance was wideotaped for PBS items. The raters watched the tapes and evlauated each child twice. Test-retest reliability was analyzed using the Spearman correlation, and interrater reliability was analyzed using the Kendall's coefficient of concordance for ranks.

RESULTS: The total PBS scores averaged 49.22 and 50.06 for first and second tests. Test-retest reliability of PBS individual items were between low and high. The low itmes

were 3 items, and high were 4 itmes. Interrater reliability were low agreement.

CONCLUSION: Based on the study, the pediatric balance scale is an effective evaluating tool for measuring functional balance of school age child. However, the agreement of students or beginner's interpretation is necessary in conjunction to this study results. On top of that, it is important to provide detailed explanation and repetitive training, which improves the analysis reliability. Also, each category subscores level for the pediatric balance scale must be examined carefully in conjunction to performing test, only after improving match quality via preceding practice.

Key Words: Pediatric balance scale(PBS), Rater reliability, Cerebral palsy

†Corresponding Author : rldnjs44@hanmail.net

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

I. 서론

균형은 평형상태에서 몸을 유지하는 능력을 말한다 (Shumway-Cook와 Woollacott, 2001). 자세조절은 수의

적 움직임의 기본이며 신체의 안정성을 마련하고 팔 다리를 이용하여 환경과 상호작용이 가능하도록 한다(Bertoti, 2003). 자세 안정성과 균형은 같은 의미로 쓰여지며 지지기저면을 지나 중력중심점을 유지하는 능력을 필요로 한다. 이들의 문제는 일상생활동작을 수행하는 기능적 과제의 어려움을 일으키며 물리치료사가 경험하는 가장 대표적인 문제 중의 하나이다(Kembhavi 등, 2002).

균형을 평가하는 방법은 여러 가지이다. 그 중 기능적 측면을 평가하기 위해 개발된 평가도구로는 일어나 걸어가기 검사(Timed up and go test; TUG)와 기능적 뻗기 검사(Functional reach test), Tinetti의 운동성 검사(Tinetti's performance oriented mobility assessment; POMA), Berg 균형척도(Berg balance scale; BBS) 등이 있다(Shumway-Cook과 Woollacott, 2001).

이 검사들은 주로 노인의 균형 능력을 평가하기 위해 개발된 검사들이다. 아동의 균형능력 평가는 표준화된 아동발달 평가 중 균형 측정 항목만을 사용하거나 정량적 평가가 가능한 측정 기구를 사용하는 것이다(Kim 등, 2010). 그러나 이러한 방법들은 균형의 기능적 측정의 평가가 불가능하다는 단점이 있다(Gan 등, 2008).

아동균형척도(Pediatric balance scale; PBS)는 약합에서 중등도의 운동장애를 가진 학령기 아동의 기능적 균형 능력 평가를 위해 Franjoine 등(2003)이 Berg 균형척도를 수정하여 만든 평가도구이다. 이 평가 도구는 앉은 자세에서 일어나기(1), 선 자세에서 앉기(2), 의자에서 의자로 이동하기(3), 잡지 않고 서있기(4), 등받이 없는 의자에 발을 바닥에 지지하여 앉기(5), 두 눈을 감고 잡지 않고 서있기(6), 두발을 모으고 잡지 않고 서있기(7), 한발 앞에 다른 발을 일자로 두고 서있기(8), 한 다리로 서 있(9), 제자리에서 360° 회전하기(10), 뒤돌아보기(11), 바닥에 있는 물건 집어 올리(12), 일정한 높이의 발판 위에 발을 교대로 놓기(13), 선 자세에서 앞으로 팔을 뻗쳐 내밀기(14)의 14개 항목으로 구성되어 있다

균형 조절은 운동기능의 기본 요소이며, 뇌성마비 같은 운동장애를 가지고 있는 아동의 운동기능 향상을 위해 균형 훈련은 중요하다. 또한 간단하고 단순한 기능적 균형 능력 측정은 운동장애 아동을 치료하고 중재

의 결과를 판단하는 물리치료사에게 필수적이다(Gan 등, 2008; Laffel과 Blumenthal, 1989). 따라서 뇌성마비 아동의 균형 평가를 위해 아동균형척도를 사용하였을 때 이 척도의 신뢰도와 타당도는 매우 중요하다(Palitano, 1989). 현재 아동균형척도에 대해 우리나라에서 이루어진 선행 연구는 3편이며 모두 총점(total test score; TTS)에 대한 신뢰도를 연구하였다(Kim 등, 2010; Ko 등, 2008; Ko와 Kim, 2010). Ko 등(2008)의 경우 뇌성마비 아동의 아동균형척도의 측정자간 신뢰도를 .97로 보고하였고, Ko와 Kim (2010)은 .93의 측정자간 신뢰도와 .97이상의 측정자내 신뢰도를 보고하였다. Kim 등(2010)은 경력자와 초심자를 비교하여 .83이상의 측정자간 신뢰도와 .89이상의 측정자내 신뢰도를 보고하였다. 이들은 측정자 신뢰도를 평가하면서 8번 항목을 비롯한 2~3개 항목에 대한 문제점을 지적한 바 있다(Kim 등, 2010; Ko와 Kim, 2010). 또한 측정자의 수가 적고 총점 외에 항목별 신뢰도를 평가하지 못했음을 제한점으로 보고하였고 앞으로 이를 보완한 연구가 필요하다고 하였다(Kim 등, 2010).

따라서 본 연구에서는 측정에 익숙치 않은 물리치료학과 재학생을 대상으로 이전 연구들 보다 측정자의 수를 확대하여 항목별 신뢰도를 연구하고자 하였다. 본 연구의 목적은 아동균형척도의 13개 항목에 대한 물리치료과 재학생들의 검사-재검사 신뢰도와 측정자간 신뢰도를 알아보고자 하였다.

II. 연구 방법

1. 연구 대상과 측정자

본 연구의 대상은 운동장애를 가진 뇌성마비 아동 6명이 참여하였다. 대상자는 대동작기능분류체계(GMFM)의 1단계에 분류되어 독립적으로 서기와 걷기가 가능하였고 충청도 소재 재활병원에서 물리치료를 받고 있는 아동이었다. 부모의 동의를 얻어 대상자가 아동균형척도의 각 항목을 수행하는 것을 약 20분가량 비디오로 촬영하였다.

측정자는 경기도 소재 한 대학에 재학 중인 물리치료

Table 1. General characteristics of participants

subject	sex	Age (years)	Height (cm)	Weight (kg)	GMFCS level	paralysis	hand function	communication	GMFM Total score
1	Female	13	154.1	40.1	1	No	1	1	92
2	Female	11	138	32.9	1	Yes	2	1	92
3	Male	13	155	45	1	Yes	3	1	97
4	Male	8	129.9	26.5	1	Yes	1	1	98
5	Female	12	145	43	1	Yes	2	1	96
6	Male	9	136	32	2	No	3	1	95

과 2학년 학생 74명이 참여하였다. 이들은 측정과 평가에 대한 교과목을 이수하였고 아동균형척도에 대해 2시간의 교육을 받았다.

연구대상자의 일반적인 특성과 장애특성, 측정자의 일반적 특성은 표 1과 같다.

2. 실험방법

1) 측정도구

측정도구는 Franjoine 등(2003)이 개발하고 Ko 등(2008)이 번역한 한글판 아동균형척도를 수정 보완하여 사용하였다. 아동균형척도의 14개 항목은 0~4점 척도로 되어 있다. 14번째 항목인 선 자세에서 앞으로 팔을 뻗쳐 내밀기”는 2차원 비디오로 촬영하여 평가하는데 제한점이 있어 이전 선행연구들과 같이 제외하였다. 본 연구에서는 아동균형척도의 13개 항목을 평가하여 총점을 52점으로 하였다.

또한 대상자 선정을 위해 대동작기능분류체계와 손기능측정도구, 의사소통능력평가도구를 사용하였다.

2) 실험 절차

연구 대상자가 편안하고 정확하게 기능을 발휘할 수 있도록 낯선 사람과 환경에 대한 변수를 고려하여 대상자들을 직접 치료하는 치료사가 대동작운동기능과 손기능, 의사소통 능력을 평가하고 아동균형척도의 항목을 비디오 촬영하였다. 비디오 촬영 시 치료사는 충분한 설명을 하였고 필요하다면 시범과 연습기회를 제공하였다. 아동균형척도의 13개 항목을 순서대로 수

행하였고 항목별로 최대 3회의 수행기회를 제공하였다. 첫 번째 수행에서 4점에 해당하는 능력을 발휘한 경우 바로 다음 항목으로 넘어갔으며 3회의 기회에도 시도하지 못하거나 실패한 경우 다음 항목을 수행하도록 하였다.

아동균형척도의 평가는 74명의 측정자를 반으로 나누어 32명씩 같은 장소에서 동시에 같은 화면을 보면서 진행하였다. 측정자들은 점수 매기는 방법에 대해 아동균형척도 평가 경험이 있는 치료사로부터 2시간의 교육을 받았다. 교육은 아동균형척도의 한글판을 프린트하여 일반적인 내용과 평가 항목과 점수 판정 방법에 대해 설명하였고 대상자가 아닌 다른 아동이 아동균형척도 평가를 수행하는 비디오를 함께 보면서 시험 측정하였다. 시험 측정은 2회 실시하였고 질문과 대답을 통해 정확한 기준을 수립하는데 도움을 주었다. 본 실험에 해당하는 평가는 1주일 간격으로 1차 평가와 2차 평가를 실시하였다. 학습효과를 배제하기 위하여 2차 평가 전에는 따로 교육을 실시하지 않았다.

3. 자료 분석

통계분석은 SPSS for Windows 12.0 ver을 사용하였고 유의수준은 .05로 하였다. 아동균형척도 항목별 점수는 평균과 표준편차를 계산하였고 13개 항목의 총점을 계산하였다. 아동균형척도는 서열척도이므로 동일 측정자의 검사-재검사 신뢰도는 Spearman 순위 상관계수(Spearman correlation)로 분석하였고, 측정자간 신뢰도는 Kendall의 일치계수(Kendall's coefficient of concordance for ranks)로 분석하였다. 다른 연구와의 비교를 위해

Table 2. Test-retest reliability of the PBS total test score and individual items score

Individual item	First	Second	Spearman	p
	M±SD			
1. Sitting to standing	3.83±.33	3.91±.17	.81*	.05
2. Standing to sitting	3.93±.09	3.96±.09	.90*	.02
3. Transfers	3.87±.16	3.92±.11	.94**	.01
4. Standing unsupported	3.97±.04	4.00±.00	N/A	N/A
5. Sitting unsupported	3.93±.04	3.99±.01	-.39	.44
6. Standing with eye closed	3.98±.02	4.00±.00	N/A	N/A
7. Standing with feet together	3.96±.02	4.00±.01	.88*	.02
8. Standing with one foot in front	3.10±.51	3.42±.53	.26	.62
9. Standing on one foot	3.63±.37	3.74±.28	.84*	.04
10. Turning 360 degree	3.80±.30	3.86±.17	.93**	.01
11. Turning to look behind	3.77±.26	3.81±.17	.66	.16
12. Retrieving object from floor	3.98±.01	4.00±.01	-.61	.20
13. Placing alternate foot on stool	3.44±.29	3.45±.28	1.00**	.00
Total test score (TTS)	49.22±.97	50.06±.67	.37	.47

N/A: Non Arithmetic

측정자 신뢰도를 분석하는 대표적인 방법인 급간내상관계수(Intraclass correlation coefficient; ICC_{3,1})를 사용하였다.

III. 결 과

1. 아동균형척도의 검사-재검사 신뢰도

74명의 측정자가 2번에 걸쳐 아동균형척도를 평가한 결과 총점은 Spearman 순위 상관계수가 .37로 낮은 일치도를 나타내었고 통계적으로 유의하지 않았다. 급간내상관계수는 .53으로 중간정도의 일치도를 나타내었다.

아동균형척도의 항목별 순위상관계수는 일정한 높이의 발판 위에 발을 교대로 놓기, 의자에서 의자로 이동하기, 제자리에서 360° 회전하기, 선 자세에서 앉기 순으로 4개 항목이 .90이상으로 매우 높은 신뢰도를 나타내었다($p < .05$). 상관계수가 가장 낮은 항목은 바닥에 있는 물건 집을 올리기로 .061이었고 그 다음은 등받이 없는 의자에 발을 바닥에 지지하여 앉기, 한 발 앞에 다른 발을 일자로 두고 서있기 항목이 통계적으로 낮은 신뢰

Table 3. Intraclass correlation coefficient of PBS individual items between test and retest

Individual items	ICC	95% C.I.
1. Sitting to standing	.81	.14-.97
2. Standing to sitting	.95	.66-.99
3. Transfers	.99	.55-.99
4. Standing unsupported	N/A	N/A
5. Sitting unsupported	-.37	-.88-.55
6. Standing with eye closed	N/A	N/A
7. Standing with feet together	.47	-.44-.90
8. Standing with one foot in front	.48	-.43-.91
9. Standing on one foot	.84	.24-.98
10. Turning 360 degree	.85	.28-.98
11. Turning to look behind	.83	.21-.98
12. Retrieving object from floor	-.60	-.93-.28
13. Placing alternate foot on stool	.98	.87-1.00
Total test score (TTS)	.53	-.38-.92

ICC: Intraclass correlation coefficient

95% C.I.: 95% confidence interval

N/A: Non Arithmetic

도를 나타내었다(Table 2). 잡지 않고 서있기와 두 눈을 감고 잡지 않고 서있기 항목은 2차 측정 시 74명의 측정자 모두 4점을 주어 측정값의 일치로 인해 통계량 계산이 불가하였다. 뒤돌아보기 항목은 중간 정도의 신뢰도를, 그 외 나머지 3항목은 높은 신뢰도를 나타내었다.

검사-재검사 신뢰도를 급간내상관계수로 분석한 결과 Spearman 순위 상관계수와 약간의 차이는 있었으나 매우 유사한 결과를 나타내었다(Table 3).

2. 아동균형척도의 측정자간 신뢰도

측정자간 신뢰도는 Kendall의 일치계수로 분석한 결과 1차 측정에서 .46, 2차 측정에서 .42로 낮은 신뢰도를 나타내었고 통계적으로 유의하였다($p < .05$)(Table 4). 급간내상관계수는 1차 측정 시 .27, 2차 측정 시 .29로 낮은 신뢰도를 나타내었다(Table 5).

Table 4. reliability of the PBS total test score and individual items score

Individual items	First		Second	
	Kendall's W	p	Kendall's W	p
1. Sitting to standing	.19	.17	.24**	.01
2. Standing to sitting	.17	.47	.17	.39
3. Transfers	.23**	.01	.25**	.00
4. Standing unsupported	.18	.25	N/A	N/A
5. Sitting unsupported	.40**	.00	.41**	.00
6. Standing with eye closed	.35**	.00	N/A	N/A
7. Standing with feet together	.32**	.00	.33**	.00
8. Standing with one foot in front	.47**	.01	.24**	.01
9. Standing on one foot	.24*	.02	.22*	.02
10. Turning 360 degree	.34**	.00	.34**	.00
11. Turning to look behind	.25**	.00	.35**	.00
12. Retrieving object from floor	.34**	.00	.17	.51
13. Placing alternate foot on stool	.48**	.00	.42**	.00
Total test score (TTS)	.46**	.00	.42**	.00

N/A: Non Arithmetic

각 항목별로 보면 1차 측정에서는 일정한 높이의 발판 위에 발을 교대로 놓기가 가장 높았고 한발 앞에 다른 발을 일자로 두고 서있기, 등받이 없는 의자에 발을 바닥에 지지하여 앉기 순이었다. 2차 측정에서는 1차 측정과 동일하였으나 한발 앞에 다른 발을 일자로 두고 서있기 항목은 낮은 일치도를 나타내었다. 가장 낮은 일치도를 나타낸 항목은 1차 측정 시 선 자세에서 앉기와 잡지 않고 서있기 항목이었고, 2차 측정 시 선 자세에서 앉기와 바닥에 있는 물건 집어 올리기였다. 한발 앞에 다른 발을 일자로 두고 서있기 항목과, 바닥에 있는 물건 집어 올리기 항목은 2차 측정 시에 일치계수가 낮아졌고 뒤돌아보기 항목은 1차 측정보다 2차 측정에서 일치계수가 높아졌다.

Table 5. Intraclass correlation coefficient of PBS individual items between test and retest

Individual item	First		Second	
	ICC	95% C.I.	ICC	95% C.I.
1. Sitting to standing	.78	.57-.95	.83	.65-.96
2. Standing to sitting	.32	.13-.75	.75	.53-.95
3. Transfers	.36	.17-.78	.50	.26-.86
4. Standing unsupported	.18	.01-.66	N/A	N/A
5. Sitting unsupported	.09	-.01-.55	.41	-.50-.89
6. Standing with eye closed	.21	-.12-.75	N/A	N/A
7. Standing with feet together	-.05	-.13-.40	N/A	N/A
8. Standing with one foot in front	.49	.20-.98	.41	.21-.81
9. Standing on one foot	.44	.21-.86	.40	.20-.80
10. Turning 360 degree	.28	.09-.78	.42	.21-.82
11. Turning to look behind	.35	.16-.77	.32	.14-.75
12. Retrieving object from floor	-.13	-.15-.03	-.20	-.83-.65
13. Placing alternate foot on stool	.33	.15-.75	.37	.18-.78
Total test score (TTS)	.27	.12-.70	.29	.13-.72

ICC: Intraclass correlation coefficient

95% C.I.: 95% confidence interval

N/A: Non Arithmetic

IV. 고 찰

뇌성마비를 포함한 운동장애를 가진 아동들은 신체 균형을 유지하기가 어려우며 결국 잘 넘어지게 되고 기능적 활동에 제한을 받는다(Ko 등, 2008). 아동의 기능적 균형은 일상생활이나 학교 생활, 놀이 등의 활동을 하는 동안 지지기저면에 대해 질량 중심을 유지하는 능력으로 정의한다(Franjoine 등, 2010). 아동균형척도는 경미하거나 중등도의 운동장애를 가진 학령기 아동의 기능적 균형 능력을 평가하기 위해 만들어진 평가도구이다(Franjoine 등, 2003). 본 연구는 학령기 뇌성마비 아동 6명을 대상으로 물리치료과 학생들의 평가 일치도를 알아보기 위해 아동균형척도의 13개 항목에 대한 검사-재검사 신뢰도와 측정자간 신뢰도를 분석하였다.

평가도구의 신뢰도는 동일한 대상에 대하여 같은 평가도구를 이용해 반복 측정할 경우 얼마나 일정성을 가지고 일치하느냐를 검정하는 것이다(Won 등, 2002). 신뢰도 평가방법에는 검사-재검사 신뢰도와 측정자 신뢰도, 평행검사법, 피어슨 상관계수(Pearson product-moment correlation coefficient), Spearman 순위 상관계수, 내적일치도(internal consistency) 등이 있다(Lee, 2008). 피어슨 상관계수는 두 개의 변수 간 관계만을 나타내며 두 개 이상의 관계를 동시에 표현하지 못한다. 이러한 문제를 해결한 방법은 급간내상관계수이다. 급간내상관계수는 신뢰도 평가에 주로 사용하며 연속형 자료의 분석에 사용하는 방법이다. 본 연구에서 사용한 아동균형척도는 14개의 항목으로 구성되어 있고 이 항목들은 0~4점의 서열척도이다. 서열척도와 같은 범주형 자료는 코헨의 카파 지수(Cohen's kappa)를 주로 사용한다. 본 연구에서는 74명의 측정자가 참여하여 카파지수를 사용할 수 없었으며 서열척도로 되어있는 두 변수의 상관관계를 나타내는 Spearman 순위 상관계수를 사용하였다. Spearman 순위 상관계수를 이용하여 신뢰도를 분석한 연구는 아동균형척도를 개발하고 신뢰도를 평가한 Franjoine 등(2003, 2010)의 연구와 한국형 일상생활활동 측정도구의 신뢰도를 평가한 Won 등(2002)의 연구가 있다. 측정자간 신뢰도를 평가하기 위해 본 연구에서는 Kendall의 일치계수를 사용하였다. Kendall의 일치계수

는 3개 이상의 서열척도로 된 변수들의 일치성을 알아보기 위해 사용하는 방법이다. Kendall의 일치계수를 이용하여 신뢰도를 분석한 연구로는 균형 결손을 위해 사용되는 균형평가시스템검사(BESTest)에 대한 연구가 있다(Horak 등, 2009). 그 외의 연구들은 카파지수를 주로 사용하였고 급간내상관계수를 사용하였다. 따라서 본 연구에서는 다른 선행 연구들과의 비교를 위해 급간내상관계수를 함께 분석하였으며 Spearman 순위 상관계수나 Kendall의 일치계수와 유사한 결과를 나타내었다.

검사-재검사 신뢰도의 경우 아동균형척도 총점은 Spearman 순위 상관계수가 .37로 낮은 일치도를 나타내었고 급간내상관계는 .53으로 중간 정도의 일치도를 나타내었다. 아동균형척도를 개발하고 신뢰도를 제시한 Franjoine 등(2003)은 .998의 매우 높은 급간내상관계수를 보고하였다. 우리나라의 경우 Ko와 Kim (2010)은 .89~.93의 급간내상관계수를 보고하였고 Kim 등(2010)은 .89~.99의 매우 높은 일치도를 보고하였다. 본 연구의 결과와 상이한 결과가 나타난 이유는 측정자의 지식과 경험에 따른 차이로 판단된다. Franjoine 등(2003)의 연구에서는 측정자가 아동균형척도를 개발한 연구자였으며 13년의 임상경력을 가지고 있었다. 우리나라의 선행연구에서도 경력자는 10년 이상의 임상경력을 가지고 있었다. 본 연구의 측정자들은 물리치료과 학생으로 진단 및 평가 과목을 두 학기 수강하였으며 아동균형척도에 대해서는 2시간의 교육과 실습을 받았다. 따라서 척도와 평가에 대한 개념이 정확하게 세워져 있지 않고 뇌성마비 아동에 대한 지식이 부족하였기 때문에 생각된다. 아동균형척도의 신뢰도를 임상 경력자와 초심자로 나누어 비교한 Kim 등(2010)의 연구에서도 측정자 내 신뢰도는 10년 이상 경력자의 경우 .91과 .99였고, 초심자의 경우 .89와 .98이었다. 초심자들에게 측정도구와 측정방법에 대한 연습과 훈련을 실시한 후에는 .85에서 .91까지 측정자 신뢰도가 증가하였다. 이러한 결과를 볼 때 본 연구의 측정자들이 낮은 검사-재검사 신뢰도를 나타낸 것은 지식과 경험이 부족하다는 것을 뒷받침한다고 할 수 있다. 이러한 결과는 많은 시사점을 가지고 있다고 생각한다. 학교에서 정확한 측정과 검사, 평가에 대한 과목을 배우고 있지만 문서나 표로만 보는 경우가

많고 실제 측정해보고 측정의 오류와 문제점을 경험하는 것은 부족하다고 생각된다. 따라서 학교 교육에서부터 정확한 측정과 검사, 평가에 대한 실습을 하는 것이 절대적으로 필요하다고 사료된다.

아동균형척도의 13개 항목별로 신뢰도를 분석한 선행연구는 Franjoine 등(2003)의 연구뿐이었다. 우리나라의 선행연구들은 모두 아동균형척도의 총점을 가지고 신뢰도를 분석하였다. 본 연구에서 아동균형척도의 각 항목별 분석 결과, Spearman 순위 상관계수는 $-.61 \sim 1.00$ 까지 다양한 일치도를 나타내었다. 이러한 결과는 항목별 신뢰도를 평가하여 카파지수로는 $.87 \sim 1.00$, Spearman 순위 상관계수로는 $.89 \sim 1.00$ 을 나타낸 Franjoine 등(2003)의 결과와는 차이가 있었다. 본 연구에서 항목별 검사-재검사 신뢰도가 $.90$ 이상으로 높은 항목은 일정한 높이의 발판 위에 발을 교대로 놓기, 의자에서 의자로 이동하기, 제자리에서 360° 회전하기, 선 자세에서 앉기 순이었다. 검사-재검사 신뢰도가 낮은 항목은 바닥에 있는 물건 집어 올리기, 등받이 없는 의자에 발을 바닥에 지지하여 앉기, 한발 앞에 다른 발을 일자로 두고 서있기 항목이었다. Franjoine 등(2003)의 연구에서는 대부분의 항목이 1.00 의 매우 높은 일치도를 나타내었고 일치도가 낮은 항목은 한 다리로서 있기($k=.87$)와 잡지 않고 서있기($r=.89$)였다. 이러한 항목별 결과 역시 본 연구와는 일치하지 않는 결과를 나타내었는데 이것은 측정자의 특성 이외에도 평가환경에 따른 차이가 있었던 것으로 생각된다. Franjoine 등(2003)의 연구에서는 측정자가 한 사람이었고 본 연구에서는 74명이었으므로 측정자의 수에 큰 차이가 있었고, 또 하나의 이유는 본 연구에서 74명의 측정자들은 37명씩 같은 공간에서 하나의 화면을 보면서 평가하였기 때문인 것으로 생각된다. 본 연구에서는 커다란 벽걸이 스크린을 사용하였으나 측정자들이 화면을 보는 각도에 따라 차이가 있었을 것으로 생각된다. 따라서 평가환경의 차이가 이러한 결과를 나타낸 것으로 사료되는바 좀 더 정확한 측정과 평가를 위해 한 번에 평가할 수 있는 화면의 크기와 그에 따른 적절한 측정자의 수를 고려할 필요가 있다고 사료된다. 또한 본 연구의 결과에서 낮은 신뢰도를 나타낸 3개의 항목들은 항목의 개념을 분명하게

설명할 필요가 있다고 생각된다. 바닥에 있는 물건 집어 올리기, 등받이 없는 의자에 발을 바닥에 지지하여 앉기의 경우 비디오 촬영 시 여러 방향에서 찍을 수 없어 등받이에 등이 닿았는지 아닌지 알 수 없는 경우가 있어 혼란이 있었고 한발 앞에 다른 발을 일자로 두고 서있기의 경우 점수 등급의 3점과 2점을 매기는 방법을 좀 더 분명하게 수정할 필요가 있다고 생각된다.

본 연구에서 아동균형척도의 측정자간 신뢰도를 분석한 결과 Kendall의 일치 계수로는 $.46, .42$ 의 낮은 신뢰도를 나타내었고 급간내상관계수는 $.27, .29$ 로 낮은 신뢰도를 나타내었다. 10명의 소아 물리치료를 대상으로 측정자간 신뢰도를 분석한 Franjoine 등(2003)의 연구에 의하면, 측정자간의 총점에 대한 차이는 통계적으로 유의하지 않았고 $.997$ 의 매우 높은 급간내상관계수를 나타내었다. 우리나라의 경우 Ko 등(2008)은 뇌성마비 아동의 경우 $.97$ 의 매우 높은 측정자간 일치도를 보고하였고 측정자는 임상 경력이 13년, 16년인 물리치료가 수행하였다. 이와 같은 결과는 본 연구의 결과와 차이가 있어 측정자내 신뢰도와 같이 해석할 수 있다. Kim 등(2010)의 연구에 의하면 측정자의 임상경력과 측정경험에 따라 초심자보다 경력자의 일치도가 더 높았고 경력자와 초심자 사이의 일치도에 차이가 있었다고 하였다. 뇌졸중 환자를 대상으로 기능적 보행평가의 신뢰도를 연구한 결과에 의하면, 4명의 측정자 중 2명은 치료사 2명은 학생이었는데 총점은 치료사나 학생에 관계없이 모두 높은 신뢰도($ICC=.91 \sim .95$)를 나타낸 반면 각 항목별로 평가한 결과 학생들의 측정자간 신뢰도가 치료사의 측정자간 신뢰도에 비해 낮은 것으로 보고하였다(Won과 Yu, 2011).

이와 같은 결과들을 토대로 본 연구의 결과와 비교하면 학생이나 경력이 부족한 치료사들의 경우 전공 지식과 환자에 대한 실질적인 이해가 부족하여 측정과 평가에 대한 연습과 훈련이 반드시 필요한 것으로 사료된다. 각 항목별 측정자간 신뢰도는 본 연구에서만 이루어져 다른 선행연구들과 비교가 불가능하였다. 본 연구에서는 Kendall의 일치계수가 $.17 \sim .48$ 사이로 나타났던 낮은 일치도를 나타내었고 급간내상관계수로 분석한 결과 $.13 \sim .83$ 사이로 나타나 다양한 일치도를 나타내

었다. 이 중 가장 낮은 일치도를 나타낸 항목은 바닥에 있는 물건 집어 올리기로 1, 2차 측정에서 13과 20이었다. 이는 검사-재검사 신뢰도와 같은 결과로 아동균형척도의 13개 항목 중 가장 낮은 검사-재검사 신뢰도와 측정자간 신뢰도를 나타낸 항목이었다. 바닥에 있는 물건 집어 올리기 항목은 안전하고 쉽게 물건을 집어들 수 있으면 4점을, 집어들 수는 있으나 옆에서 지켜봐 주는 것이 필요하다면 3점을, 집을 수는 없으나 물건이 있는 지점에서부터 2.5~5cm 거리까지 손이 닿으며 혼자 균형을 잡을 수 있으면 2점을, 물건을 집어들 수 없음 집으려 시도하는 동안에도 옆에서 지켜봐 주는 것이 필요하다면 1점을, 물건을 집으려고 시도해도 할 수 없으며 넘어지지 않게 하려면 도움이 필요하다면 0점을 준다. 이러한 점수 매기기가 비디오 화면상에서 애매하게 보여질 수 있다고 생각되며 보호자가 함께 찍혀있다면 지켜봐 주는 것인지 판단하기가 어렵다고 생각된다. 따라서 다른 항목들처럼 정해진 시간 안에 동작을 수행할 수 있는지와 같은 시간변수가 포함된다면 좀더 정확한 판단을 할 수 있을 것으로 사료된다. 또한 일반적으로 측정자내 신뢰도가 측정자간 신뢰도에 비해 높은 것은 잘 알려져 있으며(Kim과 Hong, 2010) 본 연구의 결과도 이와 유사하게 측정자내 신뢰도가 높은 것으로 나타났다.

본 연구의 제한점은 관찰 대상인 뇌성마비 아동의 수가 6명으로 74명의 측정자에 비해 현저하게 작아 적절한 통계방법의 해석이 어려운 점과 짧은 교육기간과 제한된 평가환경이 측정에 영향을 미치는 것을 완전하게 통제하지 못했다는 점이다. 이러한 제한점을 바탕으로 아동균형평가를 사용하여 뇌성마비 아동을 평가하는 치료사들에게 측정오류를 줄일 수 있는 계기가 되기를 바란다.

V. 결론

본 연구는 뇌성마비 아동의 기능적 균형능력을 측정하는 도구인 아동균형척도의 13개 항목에 대해 물리치료과 학생들을 측정자로 하여 검사-재검사 신뢰도와

측정자간 신뢰도를 알아보려고 하였다. 그 결과 검사-재검사 신뢰도와 측정자간 신뢰도는 낮은 일치도를 나타내었다. 또한 항목별 일치도를 측정할 결과 검사-재검사 신뢰도가 높은 항목은 4개였으며 낮은 항목은 매우 낮은 항목은 3개였다. 측정자간 신뢰도는 1차 측정과 2차 측정에서 낮은 일치도를 나타내는 항목이 동일하였다.

이전의 선행 연구들을 바탕으로 아동균형척도는 학령기 아동의 기능적 균형을 평가하기에 적절한 평가 도구라고 할 수 있다. 그러나 본 연구의 결과와 비교하여 학생들이나 초심자의 일치도에 대한 해석이 반드시 필요하며 평가의 신뢰도를 높이기 위해서는 충분한 설명과 반복훈련이 필요하다고 생각된다. 또한 아동균형척도의 각 항목별 점수를 매기는 기준에 대해 충분히 학습할 필요가 있으며 여러 번의 선행 연습을 통해 일치도를 높인 후 평가를 수행해야 할 필요가 있다. 이러한 결과들은 학교 교육에서 진단과 평가에 대한 관심이 높아져야 하며 이론 교육보다는 실습을 위주로 하여 실제 진단과 평가 방법에 익숙해질 필요가 있다고 생각된다. 또한 임상에서도 지식보다는 실제 평가를 수행할 수 있는 능력이 있어야 또 다른 교육 없이 임상 현장에 투입될 수 있을 것으로 생각된다.

Acknowledgement

본 연구는 2012년도 동남보건대학교 학술연구비지원에 의하여 수행된 것임

참고문헌

- Bertoti D. Functional neurorehabilitation through the life span. Philadelphia. F.A Davis. 2003:296.
- Franjoine MR, Darr N, Held SL et al. The performance of children developing typically on the pediatric balance scale. *Pediatr Phys Ther.* 2010;22(4):350-9.
- 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):114-28.
- 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.
- Horak FB, Wrisley DM, Frank J. The balance evaluation system test (BESTest) to differentiate balance deficits. *Phys Ther.* 2009;89(5):484-98.
- 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.* 2002;14(2):92-9.
- Kim GW, Hong WS. A study of measurement methods for subtalar joint motion. *J Kor Soc Phys Ther.* 2010;22(4):57-64.
- Kim GW, Ko JY, Baek SG. The reliability of a pediatric balance scale based on the raters' clinical work experience and test experience. *J Kor Soc Phys Ther.* 2010;22(6):35-42.
- 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.
- 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):86-91.
- Laffel G, Blumenthal D. The case for using industrial quality management science in health care organizations. *JAMA.* 1989;262(20):2869-73.
- Lee CH. Research methodology for physical therapists and occupational therapists. 3rd ed. Seoul. GYe Chuk Mun Wha Sa. 2008:99-108, 275-281, 395-423.
- Palisano RJ. Review of research on reliability and validity of the movement assessment of infants. *Pediatr Phys Ther.* 1989;1(4):157-220.
- Shumway-Cook A, Woollacott MH. Motor control: theory and practical applications. 2nd ed. Baltimore. Lippincott Williams & Wilkins. 2001:163-90, 274-84.
- Won CW, Rho YG, Kim SY et al. The validity and reliability of Korean activities of daily living(K-ADL) scale. *Journal of the Korean Geriatrics Society.* 2002;6(2): 98-106.
- Won JI, Yu KH. Reliability of the functional gait assessment in patients with stroke. *Journal of the Korean Academy of University Trained Physical Therapists.* 2011; 18(1):64-73.

