PDMS-2 대동작 척도의 신뢰도

김채경 · 황병용 · 김지선*
용인대학교 재활보건과학대학원 · 인하대학 부속병원 재활의학과*

Reliability of the Gross Motor Scale of the Peabody Developmental Motor Scales-2

Chae-Kyoung Kim · Byong-Yong Hwang · Ji-sun Kim*
Graduate School of Rehabilitation Health Science, Yong-In University
*Department of Rehabilitation Medicine, In Ha University Hospital**

Abstract

The purpose of this study was to describe the interrater and test-rest reliability of the gross motor of the Peabody Development Motor Scales-2(PDMS-2). Subject were 33 children without delay development between 26 months and 35 months. Three physical therapists viewed videotapes of the children and scored each child's performance on the gross motor scale of the test. 33 children were tested on one occasion, with each children scored independently and simultaneously by three physical therapists to examination interrater reliability. 30 children were each by two therapists to determine test-retest reliability within two weeks. The ANOVA statistics allows examination of the magnitude of the different sources of variability. The interrater reliabilities were analyzed for all three testers with ICC. To calculated significant differences between test and retest, paired t-test was used. The test-retest reliabilities were calculated with the ICC. The intraclass correlation coefficient(ICC) for interrater and test-retest reliability of measurements obtained with the PDMS-2 (GM) were high(ICC=0.89~0.97). The ICCs for inter-rater reliability varied from 0.92 to 0.97 for three sub-test scores and 0.93 for gross motor quotient. The ICCs for test-retest reliability varied from 0.89 to 0.94 for three sub-test scores and 0.92 for gross motor quotient. The finding suggest the PDMS-2 to be a useful and reliable instrument for assessing gross motor function and identifying children with delayed or aberrant motor development in Korea.

Key Words : Reliability, PDMS-2 (GM)

I. 서 론

최근 의학의 발달로 인하여 조산아 등의 고 위험아와 외상성 뇌손상 아동들의 생존율이 높아졌으며(Kurtz and Scull, 1993), 발달장애의 위험을 가진 아동의 수도 증가하는 추세에 있다(Allen, 1993). 또한 신생아집중치료실의 치료로 인하여 미숙아와 극소체중아의 생존율이 향상되었고(Sinclair et al, 1981), 여러 가지 신경학적인 이상 및 발달장애를 가진 아동의 이상을 조기 발견할

수 있는 발달평가 방법 및 추적관리가 가능하게 되었다(Kumar et al, 1980; Kitchen et al, 1982).

소아의 발달지연은 객관적이고 정확한 발달검사를 통하여 조기발견이 가능하고 이에 따른 적절하고 체계적인 소아재활치료 등으로 발육지연 정도를 완화시킬 수 있으며, 부모나 가족이 유아의 문제를 이해하고 협조하여 효과적인 교육을 할 수 있게 된다(Illingswort and Birch, 1982; 이근, 1987).

선별도구(screening tools)는 유아의 정상적인 성장과 발달을 파악하고 조기 중재가 필요한 유아를 조기 선별하여 특별한 서비스를 제공함으로서 유아에게 나타날 수 있는 심각한 장애들을 예방하는 것을 목적으로 하고 있다. 발달검사들을 진단목적으로 사용하는 것은 선별절차를 거친 유아의 발달상의 문제점의 원인을 밝히며 그의 적합한 치료책을 제시하는 것이다. 또한 발달검사들은 치료프로그램의 계획을 세우는데 있어서 아동의 기능적인 수준과 발달단계를 객관적으로 나타낸다. 따라서 치료를 어떻게 시작할 것인지, 치료의 효과가 어떻게 나타날 것인지에 대한 정보들을 제공해 준다. 검사의 결과는 그들의 부모에게 아동이 갖고 있는 능력과 제한점을 설명해주며, 아동의 잠재력과 능력은 미래의 일반적인 목표와 계획을 세우는데 기초가 될 것이다. 검사의 결과에 따라 세부적인 평가를 실시하고 발달지연의 근본적인 원인이 되는 어느 특정 영역의 지체를 알 수 있게 해준다(Brenneman, 1992).

한편 우리나라의 소아 물리치료 분야에서는 이충휘 등(1995)이 대동작 운동기능 평가를 국문으로 번안하여 검사자간의 신뢰도를 검증한 연구 이외에는 물리치료사들이 외국의 평가도구를 대상으로 국내에서 신뢰도 및 타당도를 검증하거나 표준화시켜 사용한 것은 아직 보고 된 바 없다(박혜정, 2000). 아동학 분야에서는 몇몇의 표준화 발달도구를 임상에서 적용하고 있는데, 영유아용 발달검사로 미국의 덴버발달선별 검사를 번안하여 표준화를 시도하였으며(오가실, 1976; 이근, 1987), 또한 제경숙(1982), 정문자(1993), 배윤희 등(2002)은 베일리 유아 발달척도를 번안하여 한국표준화를 시도하였다.

박혜정(2000)의 연구에서 발달지연 아동 및 뇌성마비 아동의 평가 실태를 보면 치료사의 평가도구 사용에 대한 인식도가 높음에도 불구하고 발달지연 아동의 있어서 주관적인 서술이 121명(50.6%), 대동작 운동기능 평가(Gross Motor Function Measure)가 43명(18%), 기관 내에서 자체 제작한 도구가 41명(17.2%), 덴버 발달 판별 검사(Denver Developmental Screening Test)가 14명(5.9%)이었다. 기타에는 Test of Infant Motor Performance, Peabody Developmental Motor Scales (PDMS), Test of Motor Impairment (TIMP), Bayley Motor Scale과 Alberta Infant Motor Scale(AIMS)을 사용하고 있다.

대부분이 표준화 검사들이지만, 우리나라의 아동들에게 적합하게 적용하기 위한 신뢰도 및 타당도, 표준화에 대한 연구는 미비한 실정이다. 또한 대다수의 치료사들이 사용하고 있는 운동발달평가는 운동발달단계이며 대부분이 수치화된 평가도구의 사용이 적고, 정형화된 사용이 적으며 사용하는 평가도구들에 대한 신뢰도나 타당도, 적용 가능한 대상들인지 확인하지 않고 외국의 평가도구를 그대로 번역하여 사용하고 있는 실정이다(김지선, 2002). 이는 제한된 치료시간에 많은 환자들을 치료해야하는 현재의

임상적 여건이 평가도구를 사용하기에 불리하다는 점과 각각의 평가도구로 평가한 것에 대한 의료보험수가가 따로 책정되어 있지 않으며, 평가에 관련된 교육을 받은 적이 없거나 평가도구의 선정이 어렵다는 것이 그런 원인으로 작용한다(박혜정, 2000). 그러나 물리치료의 타당성 및 필요성을 객관적으로 제시하기 위해서는 반드시 신뢰도와 타당도가 뒷받침되는 평가도구의 사용이 필요하다(김지선, 2002).

이에 본 연구자는 현재 미국에서 이런 목적으로 물리치료사 작업치료사들이 가장 많이 사용하고 있는 Peabody Developmental Motor Scales-2(PDMS-2)를 사용하여 우리나라의 만 2세 정상아동에게 운동발달평가를 하여 검사자간의 신뢰도와 검사-재검사의 신뢰도를 측정하고, 우리나라 정상아동과 장애아동에게 적용가능한지 알아본다. 또 국내의 물리치료사들에게 소개하여 운동발달지연 및 운동발달상의 문제점들을 판별하고 장애 아동의 운동기능검사로서 임상에서 적절히 활용하고자한다.

II. 연구방법

1. 연구대상 및 연구기간

본 연구의 대상은 경기도 부천시에 소재하고 있는 국·공립어린이집 7곳을 선정하였다. 의학적인 문제없이 태어난 만기(full-term) 출산아이며 최근에 정신적, 신체적, 행동적인 문제가 없는 정상 아동을 대상으로 하였으며, 연령은 26개월부터 35개월까지의 남·여 아동 33명을 대상으로 하였다.

연구에 참여한 검사자들은 임상에서 소아 물리치료 경력 5년 이상이며, 보바스 소아 기본 강좌를 이수한 치료사들로 구성하였다.

연구의 기간은 2003년 2월 14일부터 2월 28일까지 아동의 발달평가를 실시하고 비디오 촬영을 하였으며, 검사자간의 신뢰도 측정은 3월 12일부터 3월 14일 까지 3일 동안 실시하였으며, 검사-재검사는 1차 평가는 3월 12일, 3월 14일에 시행하였고, 2차 평가는 3월 28일에 실시하였다.

2. 평가방법

1) 아동의 운동발달 평가

대동작 척도에 대한 평가는 아동이 자연스럽고 편안한 환경을 조성하여 아동의 능력을 최대한 발휘 할 수 있도록 어린이집에서 실시하였다.

평가항목은 아동의 연령이 26개월 이상이므로 반사 항목은 실시하지 않았으며 고정, 이동, 사물조작의 수행을 평가하였다. 아동이 관심을 보이는 하위검사부터 선택하여, 아동의 교정연령에 따른 출발점(entry level)에서 시작하여, 수행의 천정점(ceiling level)에 이르면 검사를 종료하였다.

대동작 검사를 시행하는데 소요되는 시간은 아동 1명 당 30분이였으며, 전 과정을 비디오로 촬영하여 편집하였다. 신뢰도 측정을 위한 편집된 아동의 항목의 수행시간은 15분 이였다. 비디오 촬영의 기술적인 문제로 아동의 자세와 움직임이 정확하게 보이지 않는 항목들은 제외하였다.

신뢰도 검사 전에 검사도구의 매뉴얼을 가지고 평가에 참여할 치료사들에게 평가방법과 평가 시에 숙지해야할 사항에 대하여 오리엔테이션을 실시하였고, 평가 전에 비디오 훈련을 2회 실시하였다.

2) 검사자간 신뢰도 측정

검사자간 신뢰도 측정은 3월 12일부터 3월 14일까지 3일간 33명의 아동을 대상으로 3명의 검사자가 실시하였다. 편집된 비디오 화면을 보면서 동시에 동일한 아동을 안정, 이동, 사물조작의 하부 검사를 평가하였으며, 각 항목의 규준에 맞게 채점하여 실제점수, 표준점수, 대동작 지수를 산출하였다.

3) 검사-재검사 신뢰도 측정

검사-재검사는 33명의 아동 중 30명을 대상으로 2명의 치료사가 각각 15명씩 나누어 1차 평가를 실시한 후 2주 간격으로 재평가하였다.

아동의 수행을 비디오로 촬영을 하였기 때문에 본 연구에서는 아동의 수행이 시간이나 환경에 따라서 동일한 조건이 된다. 2차 평가는 기억효과와 반복효과를 통제하기 위하여 검사-재검사의 시간 간격을 2주로 하였다.

4) 예비 검사

예비 검사는 부천시에 거주하고 있는 정상아동 15명과 발달장애를 가진 15명의 아동을 대상으로 실시하였다.

연구기간은 2002년 7월 15일부터 8월 30일까지 어린이집, 물리치료실, 가정에서 각각 실시한 후 비디오 촬영을 하여 9월 2일부터 9월 27일까지 3명의 물리치료사가 검사자간의 신뢰도를 측정하였다.

대동작 척도의 전체 실제점수에 대한 검사자간 신뢰도는 ICC(intraclass correlation coefficient)값 0.99였다. 각 하부 검사에 따른 신뢰도는 반사 0.99, 고정 0.98, 이동 0.99, 사물조작에서 0.98로 높게 나타났다($p<0.001$).

3. 분석방법

본 연구에서 대상자들의 특성을 알아보기 위해서 빈도 분석을 실시하였으며 모든 측정치에 대한 평균값과 표준편차를 구하였다.

검사자간 신뢰도는 동일한 아동을 대상으로 3명의 검사자가 각각 측정한 점수에 대한 차이가 있는지를 추정하기 위하여 ANOVA (one-way analysis)를 하였고, 신뢰도 계수는 급간내 상관계수(ICC)를 구하였다. 검사점수의 상관관계는 검사자를 2인씩 짝지어 Pearson 적률 상관계수(product-moment correlation coefficient)를 구하였다.

검사-재검사 신뢰도는 2주 간격으로 측정한 점수의 차이가 있는지 알아보기 위하여 Paired t-test를 하였고, 두 검사의 신뢰도를 알아보기 위해 신뢰도 계수 ICC를 구하였다. 또한 각 검사점수의 상관관계를 알아보기 위해서 Pearson 적률 상관계수를 구하였다. 자료의 통계처리는 SAS(Ver. 8.0)를 이용하여 분석하였다.

III. 결 과

1. 연구 대상자의 특성

연구 대상자는 정상아동으로 남아 16명(48.5%), 여아는 17명(51.5%)으로 모두 33명이다. 모든 아동은 만기 출산아로 재태 연령이 동일하였으며, 교정 평균연령이 31.39±2.41개월이었다. 32개월에서 34개월까지의 아동이 15명(45.4%), 35개월 아동이 2명((6.1%)으로 가장 적었다.

아동의 평균 체중은 14.62±1.48kg이였고, 11kg 미만이 1명(3.0%), 13.01kg에서 15kg까지 15명(45.5%)이였으며, 평균 신장은 91.2±3.4cm

표 1. Characteristics of subjects

Characteristics	Group	N(%)	M(SD)
Gender	male	16 (48.50%)	
	female	17 (51.50%)	
Gestational age	full-term	33 (100%)	
Corrected age	26 ~ 28 months	4 (12.10%)	
	29 ~ 31 months	12 (36.40%)	31.39 (2.41)
	32 ~ 34 months	15 (45.50%)	
	35 ≤	2 (6.00%)	
Weight	11 kg	1 (3.00%)	
	11.01 ~ 13.00 kg	7 (21.20%)	14.62 (1.48)
	13.01 ~ 15.00 kg	15 (45.50%)	
	15.01 ~ 17.00 kg	10 (30.30%)	
Heihgt	85 cm	1 (3.00%)	
	86 ~ 90 cm	15 (45.50%)	91.20 (3.40)
	91 ~ 95 cm	11 (33.30%)	
	96 cm≤	6 (18.20%)	
	3.00 kg	6 (18.20%)	
Birth weight	3.01 ~ 3.20 kg	11 (33.30%)	3.28 (0.35)
	3.21 ~ 3.40 kg	6 (18.20%)	
	3.41 kg≤	10 (30.30%)	

표 2. Mean and standard deviation of gross motor scores on PDMS-2 (n=33)

	Tester A	Tester B	Tester C
Standard scores			
Stationary	11.33±1.68	11.33±1.38	11.54±1.54
Locomotion	10.00±1.06	10.00±1.11	10.00±0.96
Object manipulation	10.21±1.26	10.09±1.01	10.21±1.13
Quotient			
Gross motor	103.30±5.96	103.21±4.94	103.66±5.17

Values are mean ± standard deviation.

표 3. ANOVA table for gross motor scores of three testers on PDMS-2 (n=33)

Source of variance	SS	df	MS	F	p
Stationary					
Between groups	3.818	2	1.909	0.07	0.93
Within groups	2777.818	96	28.936		
Total	2781.636	98			
Locomotion					
Between groups	0.990	2	0.495	0.21	0.81
Within groups	228.848	96	2.384		
Total	229.838	98			
Object manipulation					
Between groups	0.000	2	0.000	0.000	1.00
Within groups	122.000	96	1.271		
Total	122.000	98			
Gross motor quotient					
Between groups	0.323	2	0.162	0.12	0.88
Within groups	125.758	96	1.310		
Total	126.081	98			

SS : sum of sequres, df : degree of freedom, MS : mean squres

±3.4cm로 85cm미만이 1명(3.0%), 86cm에서 90cm사이는 15명(45.5%)으로 나타났다.

출생 시의 평균 체중은 3.28±0.35kg이였다. 3.0kg미만이 6명(18.2%)이고, 3.01kg에서 3.20 kg 사이는 11명(33.3%)으로 나타났다(표 1).

2. 대동작 척도의 검사자간 신뢰도

PDMS-2 평균 대동작 운동지수는 검사자 A가 103.30±5.96, 검사자 B가 103.21±4.94, 검사자 C가 103.66±5.17였다. 고정의 평균 표준점수는 검사자 A가 11.33±1.68, 검사자 B가 11.33±1.38, 검사자 C가 11.54±1.54였다. 이동의 평균 표준점수는 검사자 A가 10.00±1.06, 검사자 B가 10.00±1.11, 검사자 C가 10.00±0.94였다. 사물조작의 평균 표준점수는 검사자 A가 10.21±1.26, 검사자 B가 10.09 ±1.01, 검사자 C는 10.21±1.13이였다(표 2).

3명의 검사자가 각각 측정한 대동작 척도의 점수는 안정, 이동, 사물조작, 전체 대동작 지수에서 유의한 차이가 없었다(표 3).

대동작 척도에서 2인씩 짝지은 검사자간의 상관관계를 보면, 전체 대동작 지수는 높은 상관관계를 보였으며(r=0.78~0.88), 고정에서 비교적 상관관계가 높게 나타났다(r=0.73~0.92). 검사자간의 상관관계가 가장 높은 하부검사는 이동(r=0.90~0.93)이였고, 사물조작도 상관관계는 높게 나타났다(r=0.78~0.88). 또한 예비연구에 참여하였던 검사자 A와 검사자 C의 대동작 척도는 대동작 운동지수(r=0.88), 고정(r=0.92), 이동(r=0.93), 그리고

표 4. Interrater reliability coefficients for gross motor scale on PDMS-2 (n=33)

PDMS-2	A vs B	A vs C	B vs C	ICC
Standard scores				
Stationary	0.78	0.92	0.73	0.93
Locomotion	0.90	0.93	0.91	0.97
Object manipulation	0.74	0.88	0.80	0.92
Quotient				
Gross motor	0.78	0.88	0.77	0.93

ICC : intraclass correlation coefficients
A : Tester A, B : Tester B, C : Tester C

표 5. Comparison of tests on gross motor scale (n=30)

PDMS-2	First test	Second test	t	p
Standard scores				
Stationary	11.60 ± 1.58	11.76 ± 1.45	1.00	0.33
Locomotion	10.30 ± 1.26	10.20 ± 1.03	-1.00	0.33
Object manipulaion	10.20 ± 1.09	10.20 ± 1.03	0.00	1.00
Quotient				
Gross motor	104.23 ± 5.62	104.66 ± 4.87	0.87	0.39

values are mean ± standard deviation

표 6. Test-retest reliability coefficients for gross motor scale on PDMS-2 (n=30)

PDMS-2	1st test vs 2nd test	ICC
Standard scores		
Stationary	0.82	0.90
Locomotion	0.90	0.94
Object manipulation	0.81	0.89
Quotient		
Gross motor	0.87	0.92

Value are correlation coefficients(r)
ICC : intraclass correlation coefficient

사물조작(r=0.88)에서 강한 상관관계를 보였다. 평가도구의 사용경험이 많았던 검사자간의 상관관계가 더 높게 나타났다. 3명의 검사자간 신뢰도는 하부검사 고정에서 ICC값 0.93, 이동 0.97, 사물조작 0.92로 나타났다. 대동작 지수는 ICC값 0.93으로 대동작 척도의 검사자간 신뢰도는 높게 나타났다(표 4).

3. 대동작 척도의 검사-재검사 신뢰도

대동작 척도에 대한 1차 검사의 평균 운동지수는 104.23±5.62였고, 2차 검사는 104.66±4.87였다. 1차 검사와 2차 검사의 평균점수를 비교하면 안정, 이동, 사물조작에서 유의한 차이가 없었으며, 대동작 지수 역시 검사간의 점수에서 통계적으로 유의한 차이가 없었다(표 5).

검사-재검사간의 상관관계는 운동지수(r=0.87, $p<0.01$)에서 높은 상관관계를 보였다. 또한 각 하위검사에서 고정(r=0.82), 이동(r=0.90), 사물조작 (r=0.81) 모두 강한 상관관계를 보였다. 검사-재검사 신뢰도는 안정에서 ICC값 0.90, 이동 0.94, 사물조작 0.89, 대동작 지수는 0.92로 높게 나타났다(표 6).

Ⅳ. 고 찰

물리치료에서 표준화된 평가도구는 아동의 판별, 진단, 치료의 적합성, 치료 계획 그리고 치료의 효과를 입증하기 위한 목적으로 사용한다(Palisano, 1995). 물리치료사나 작업치료사가 아동의 발달을 평가하고자 할 때 제일 먼저 생각하는 것은 사용하고자 하는 평가도구의 신뢰도이다. 신뢰도는 검사시에 얻어진 측정점수를 기초로 하여 임상적인 판단을 할 때 필수적이기 때문에 중요하다(Waltz et al, 1991).

검사자간 신뢰도는 동일한 집단을 대상으로 다른 검사자들이 검사를 시행할 때 검사점수의 일관성을 보는 것이다. 이런 검사자간 신뢰도에서는 검사자의 판단이 가장 중요하다. 검사자간 신뢰도를 연구할 때 높은 신뢰도를 가지기 위해서는 검사시행 과정과 채점 과정, 표준점수 산출 과정을 검사 메뉴얼을 통해 숙지해야만 하고 연구를 설계한 사람은 검사자들에게 직접 사용할 검사시행 방법과 객관적으로 채점할 수 있도록 자세한 설명을 해주며 비디오 훈련을 통하여 검사 규준에 일치하여 정확하게 판단할 수 있도록 돕는 것이 신뢰도를 높이는 것이다. 또한 검사에 참여한 치료사들의 풍부한 임상적인 경험과 평가도구를 얼마나 잘 숙달하게 사용할 수 있느냐에 대한 문제도 신뢰도에 많은 영향을 미치게 된다(Deitz, 1989).

본 연구에 참여한 3명의 치료사들은 임상에서 PDMS-2를 처음 소개받아서, 본 실험 전에 충분한 검사자 훈련과 비디오 훈련을 실시하였다. 이중 2명의 치료사는 예비실험에 참여하여 장애아동과 정상아동 30명을 평가한 경험이 있으며, 1명의 치료사는 예비실험에 참여하지 않았다. 본 연구의 결과에서는 예비실험에 참여한 2명의 치료사는 대동작 운동지수에서 높은 상관관계를 보였다(r=0.88). 또한 하위검사는 고정, 이동, 사물조작에서 강한 상관관계를 보였다(r=0.88~0.93). 따라서 검사자의 훈련수준과 평가도구의 사용 경험이 신뢰도에 영향을 주는 요인이 된다.

GMFM의 신뢰도 연구를 보면 뇌성마비 아동을 임상에서 2년 이상 치료한 경험이 있고, GMFM을 사용한 경험이 없는 치료사 15명이 검사자 훈련도 거치지 않고 3명의

뇌성마비아동을 평가하였다. 검사자간 신뢰도를 6개월 간격으로 2번 측정하였는데, 1차 검사자간 신뢰도는 0.77, 2차 검사자간 신뢰도는 0.88이 나왔다. 이는 검사를 시행할 때 검사자의 임상 경험이 아동의 수행판단이 매우 중요하며, 검사도구가 사용하기에 쉽다는 것을 보여준다. 또한 2차 평가에서 검사자간의 신뢰도가 높게 나온 결과는 검사의 사용 경험이 증가하였고, 치료사들이 검사를 시행할 때 1차 평가보다는 좀더 적극적인 태도로 참여한 결과이다(Nordmak et al, 1997).

Schmidt 등(1993)의 연구에서 4세와 5세 사이의 아동 중 11명은 발달지연이 없는 아동, 22명은 발달지연이 있는 아동들을 각각 세 그룹으로 나누어 PDMS 대동작을 평가하고 검사자간 신뢰도를 측정하였다. 발달지연이 있는 두 그룹에서는 ICC값이 각각 0.84, 0.93이였고, 발달지연이 없는 아동은 0.94로 가장 높았다. 전체의 신뢰도 수준은 0.84에서 0.94로 높은 편이다. 발달지연이 없는 아동에서 신뢰도가 높게 나타났던 이유는 수행기준에 대한 2점이 발달지연이 있는 아동보다는 검사자들이 판단하기가 쉬웠기 때문에 점수의 일치도가 높았다. 이는 검사자간 신뢰도의 연구에서는 검사항목의 정확한 판단기준이 채점과정에 중요하다는 것을 보여준다(Schmidt et al, 1993).

개정된 PDMS-2로 실시한 검사자간의 신뢰도는 전체 신뢰도 ICC값은 0.93였고 각 하부검사의 신뢰도는 고정 0.93, 이동 0.97, 사물조작에서 0.92로 나타났다. PDMS-2의 검사자간의 신뢰도는 대동작 운동지수가 0.97, 고정 0.97, 이동 0.99, 사물조작에서 0.98로 본 연구의 신뢰도 수준과 비교하면 PDMS-2가 약간 높지만, 본 연구의 검사자간 신뢰도의 수준은 ICCs〉 0.90이므로 높은 수준이다. 또한 PDMS와 PDMS-2의 신뢰도를 비교하면 PDMS-2가 더 높게 나타났다. 이는 수행판단의 1점 기준이 더 명확하게 정의되었고 표준화 과정에서 운동장애를 가진 아동이 10% 포함되었기 때문이다(Folio and Fewell, 2000).

검사-재검사 신뢰도는 반복적인 측정했을 때 안정적인 점수를 갖는 범위를 말한다. 검사-재검사 신뢰도를 측정하기 위해서는 동일한 검사를 동일한 검사자가 2회에 걸쳐 시행하는 것이다. 점수의 안정성이란 임상에서 사용하는 검사에 매우 중요하며, 측정이 반복되더라도 영향을 받지 않는다.

본 연구에서는 검사-재검사의 두 검사의 대동작 운동지수에 대한 상관관계는 높게 나타났다(r=0.87). 신뢰도는 ICC값 0.92로 높게 나타났고, 하부검사에서 고정 0.90, 이동 0.94, 사물조작에서 0.89로 높게 나타났다. 이는 PDMS-2의 저자가 실시한 검사-재검사의 신뢰도의 대동작 운동지수 0.89, 고정 0.92, 이동 0.94, 사물조작 0.89와 비교하면 본 연구의 신뢰도가 조금 높게 나타났다. 저자가 실시한 검사와 재검사의 시간 간격은 1주일이며, 본 연구에서는 재검사시에는 검사자의 기억과 반복의 효과를 배제하기 위하여 첫 번째 검사를 실시하고 2주 후에 재검사를 실시하였다.

Movement Assessment of Infants의 표준판과 개정판을 1주일 간격으로 30명의 고위험아를 측정한 검사-재검사의 신뢰도를 살펴보면 신뢰도 ICC값이 표준판에서는 0.79, 개정판에서는 0.83으로 나왔다(Brander et al, 1993). 또한 뇌성마비 아동의 움직임의 질을 측정하는 Gross Motor Performance Measure의 신뢰도 연구에서는 25명의 뇌성

마비 아동, 2명이 정상아동, 1명의 외상성 뇌손상 아동을 대상으로 검사를 시행하여 비디오 화면을 보면서 아동을 평가하였다. 이 연구에 참여한 검사자들은 1일의 검사자 훈련을 받았다. 하부검사 영역에서는 0.89~0.96까지의 신뢰도를 가지며 전체점수에서는 0.96으로 높게 나왔다.

그러나 비디오 화면을 보면서 평가하는 것은 검사 시 아동에서 나타날 수 있는 변수를 줄일 수 있으나 측정 점수에 대한 오차의 주요 요인이 될 수도 있다. 비디오 화면이 아동의 수행을 완전하게 보여 줄 수 없으며, 수행 시의 아동의 자세와 움직임이 카메라의 각도에 따라서 다르게 보일 수 있기 때문에 측정 오차가 생길 수 있다. 따라서 아동을 직접 평가하는 것이 더 바람직하다고 한다(Gowland et al, 1995). 본 연구에서는 비디오 화면으로 아동을 평가하였기에 재검사시에 아동에게서 나타날 수 있는 수행의 변수는 제한하였다. 그러나 공을 이용한 사물조작과 같은 경우에는 공이 굴러가는 속도와 카메라의 각도가 일치되어야 하는 기술적인 문제가 있었으므로 아동의 자세 및 공이 굴러간 거리를 측정할 때에는 측정오차가 증가할 수 있는 요인이 되었다.

기존의 PDMS는 신뢰도가 높은 반면에 많은 임상가들에게 지적되었던 측정오류는 점수를 채점하는 수행기준에 있다. 항목에 따른 수행기준은 3점 척도이며 0점은 아동이 수행기준에 성공하지 못 했을 때, 1점은 수행기준에 완전히 일치하지 않지만 수행을 모방했을 때, 2점은 수행기준에 완전히 일치하여 성공적으로 수행했을 때에 얻게 된다. 0점이나 2점에 대한 수행기준은 명확히 규정되어 있는데, 1점에 대한 수행기준은 명확하게 정의되지 않았다. 이것은 각 항목에 대한 판단기준은 아동의 비정상적인 움직임의 패턴으로도 2점을 획득할 수 있기 때문에 운동패턴의 질적인 측정에서는 충분히 설명되지 않았다. 또한 1점은 아동의 수행이 판단기준과 거의 비슷하게 모방되는 수행정도가 어느 정도까지인지에 대한 정의도 명확하지 않았다. 이것은 표준화 검사를 시행할 때 평가 결과에 따라 해석이 상당히 달라질 수 있으며, 검사자간의 신뢰도 수준을 낮추게 된다(Palisano and Lydic, 1984). 또한 정상아동의 평균수행에 기초한 규준참조검사에서는 아동의 운동발달지연 판별과 진단, 치료의 적합성을 결정하는데 혼란을 줄 수 있다. 또한 평가도구를 사용하는 각 기관마다 점수가 다르게 해석될 수 있음을 보여주는 것이다.

이런 문제점들을 보완하여 다시 개정된 PDMS-2는 좀더 명확한 수행기준을 가지고 검사를 시행한다. 본 연구의 검사자간 신뢰도와 검사-재검사 신뢰도의 결과는 ICC값 0.93, 0.92로 높게 나타났다. 이는 PDMS-2가 소아 물리치료를 하고 있는 우리나라 물리치료사들에게 운동발달지연의 위험을 가진 아동과 운동발달장애를 가진 아동에게 어려움 없이 적용할 수 있는 안정적인 평가도구라는 것을 보여 준다.

Kerlinger(1973)의 말을 인용하면 높은 신뢰도는 과학적인 결과를 뒷받침할 수 없지만, 과학적인 결과에는 신뢰도가 있어야 된다는 말은 곧 물리치료분야에서 신뢰도와 타당도가 높은 표준화도구를 사용하여 임상에 적용할 때에는 물리치료의 평가와 치료가 객관적이고 과학적인 결과를 가져올 수 있다는 것이다. 따라서 아동의 운동발달을 평가할 때에는 평가도구의 기여도가 매우 중요한 역할을 한다. 그러므로 우리나라에서

보다 객관적인 도구로 사용되기 위해서는 신뢰도를 기초한 타당도와 인종, 사회적·문화적 특성을 고려한 표준화 연구도 앞으로 진행되어야 할 것이다.

V. 결 론

부천시에 거주하고 있는 만 2세의 정상아동 33명을 대상으로 표준화 운동발달 검사인 PDMS-2의 대동작 운동기능을 평가하여 검사자간 신뢰도 및 검사-재검사 신뢰도를 알아보았다. 다음과 같은 결과를 얻었다.

1. 33명의 정상아동의 대동작 척도의 운동지수에 대한 검사자간의 신뢰도는 ICC값 0.93로 나타났다. 하부검사의 표준점수는 고정에서 0.93, 이동에서 0.97, 사물조작에서 0.92로 나타났다. 전체 신뢰도의 수준은 ICCs>0.90 수준으로 높게 나타났다.

2. 정상아동 30명에게 시행한 대동작 척도의 검사-재검사의 신뢰도는 대동작 운동지수에서는 ICC값이 0.92로 나타났으며 각 하부검사에 대한 ICC값은 고정에서 0.90, 이동에서 0.94, 사물조작에서 0.89로 비교적 높게 나타났다.

이 연구는 만 2세의 정상아동을 대상으로 하였기 때문에 모든 검사항목이 포함되지 않으며, 24개월부터 35개월 아동의 수행 항목의 신뢰도를 보여준 것이다. 따라서 검사도구 전체의 항목에 대한 신뢰도를 알아보기 위해서는 다양한 연령범위에 따른 신뢰도 연구가 필요하다. 우리나라의 발달장애아동과 뇌성마비 아동에게 적절하게 사용하기 위해서는 신뢰도를 기초한 타당도의 후속 연구가 진행되어야 할 것이다.

참고문헌

김지선. 발달장애 아동의 대동작 운동기능 평가와 베일리 영유아 발달검사-Ⅱ의 비교. 석사학위논문. 용인대학교 대학원; 2002.

박혜정. 발달지연 아동 및 뇌성마비 아동의 평가 실태와 물리치료사들의 평가에 대한 인식도 조사. 석사학위논문. 연세대학교 대학원; 2000.

이근. 서울 아동의 덴버 발육 선별 검사 및 한국판 DDST. 소아과 1987;30(9) 958-971.

제경숙. 베일리의 유아발달 검사의 표준화를 위한 예비연구. 석사학위논문. 연세대학교 대학원; 1983.

Allen MF. The high-risk infant. Pediatr Clin North Am. 1993;40:479-489.

Brenneman SK. Test of infant and child development. In: Tecklin JS eds. Pediatric Physical Therapy. 3rd ed: Lippincott. 1999.

Deitz JC. Reliability. Phys Occup Ther Ped. 1989;9:125-147.

Folio MR, Fewell PR. Peabody developmental motor scales, Second edition :examiner's manual. 2000.

Gowland C, Boyce WF, Wright V, Russell DJ, Goldsmith CH. Reliablity of the gross motor performance measure. Phys Ther. 1995;75:597-602.

Illingsworth RS, Birch IB. Diagnosis of mental retardation in infancy, a follow-up study. Arch Dis Child.

1974;43:269-275.

Kerlinger FN. Foundations of behavioral research. New York, Holt Rinehart & Winston Inc. 1973.

Kurtz LA, Scull SA. Rehabilitation for developmental disabilities. Pediatr Clin North Am. 1993;40:629-642.

Nordmark E, Hägglund G, Jarnlo GB. Reliability of the gross motor function measure in cereberal palsy. Scand J Rehab Med. 1997;29 :25-28.

Paliano RJ, Lydic JS. Peabody Developmental motor scale: An analysis. Phys Occup Ther Ped. 1984;4(1) :69-75.

Schmidt LS, Westcott SL, Crown TK. Interrater reliability of the gross motor scale of the peabody developmental motor scales with 4-and 5-year-old children. Pediatr Phys Ther. 1993;5:169-175.

Sinclair JC, Tarrrance GW, Boyle MH, Horwood SP, Saigal S, Sackette DL. Evaluation of neonatal intensive care programs. New Engl J Med. 1981;305:489-494.

Waltz CF, Strickland OL, Lenz ER, Measurement in nursing research. 2nd ed: Philadelphia, F.A. Davis. 1991.