

# Evaluating the Reliability of Short-Form Berg Balance Scales and Short-Form Postural Assessment Scales in Chronic Stroke Survivors

Seung-Heon An<sup>a</sup>, Dae-Sung Park<sup>b\*</sup>

<sup>a</sup>Department of Gait Lab of National Rehabilitation Center, Republic of Korea

<sup>b</sup>Department of Physical Therapy, Konyang University, Daejeon, Republic of Korea

**Objective:** This study aims to assess the test-retest reproducibility of the Short Form Berg Balance Scale (SF-BBS) and the Short Form Postural Assessment Scale for Stroke (SF-PASS) among chronic stroke survivors, focusing on their reliability for consistent measurements over time.

**Design:** A cross-sectional study design

**Methods:** Thirty chronic stroke survivors participated in this study, undergoing evaluations with SF-BBS and SF-PASS scales at two different points, separated by a seven-day interval. The analysis focused on test-retest reliability, employing statistical measures such as the Intra-Class Coefficient (ICC<sub>2,1</sub>), Standard Error of Measurement (SEM), Minimal Detectable Change (MDC), and MDC%, the Bland-Altman plot to assess the limits of agreement and the extent of random measurement error.

**Results:** The study found notable test-retest reproducibility for both SF-BBS and SF-PASS, with ICC values demonstrating strong reliability (0.932 to 0.941, with a confidence interval of 0.889 to 0.973). SEM values for SF-BBS and SF-PASS were reported as 1.34 and 0.61, respectively, indicating low measurement error. MDC values of 3.71 for SF-BBS and 1.69 for SF-PASS suggest that the scales have an acceptable level of sensitivity to change, with reliability metrics falling below 20% of the maximum possible score.

**Conclusions:** The findings suggest that both SF-BBS and SF-PASS exhibit high intra-class correlation coefficients, indicating strong test-retest reliability. The SEM and MDC values further support the scales' reproducibility and reliability as tools for evaluating mobility and dynamic balance in chronic stroke survivors. Therefore, these scales are recommended for clinical use in this population, providing reliable measures for assessing progress in rehabilitation.

**Key Words:** Stroke, Balance, Reproducibility, Minimal detectable change, Postural change

## 서론

뇌졸중 후 균형 장애는 일반적인 현상으로 뇌졸중 환자의 균형 결핍은 기능 회복에 부정적인 결과를 초래하는 주요인으로 작용한다[1, 2] 치료사는 환자들의 균형 능력을 평가하고 치료 중재 후 효과를 파악하거나 기능적인 변화와 특성을 재현할 수 있는 심리측정학적인 특성이 입증된 평가 도구를 사용하여야 한다[3]. 게다가 임상평가는 평가 결과를 간결하게 정량화 할 수 있어야 하며, 검사 재검사의 재현성이 충분히 검증되어야 한다.

재현성(Reproducibility)이란 첫 평가 후 재평가에서도 이전 평가 결과와 유사한 값이 얼마나 일관성 있게 유지되고 있는 지로 정의되며, 신뢰도와 일치율의 개념으로서 포괄적인 용어로 사용되고 있다[4, 5]. 일반적으로 신뢰도는 측정값의 일치율과 측정 오차가 포함되어 있다[6]. 검사-재검사의 재현성은 대부분 급간내상관계수(Intraclass correlation coefficient, ICC)에 의해 결정되는데[5] 이는 반복 측정 시 검사 결과값들의 일관성을 나타낼 수 있으나 검사자 간의 신뢰도를 비교하는데 사용되는 상대적 신뢰도의 한 형태일 뿐 단위 없이 숫자

Received: Mar 19, 2024 Revised: Apr 24, 2024 Accepted: Apr 29, 2024

Corresponding author: Dae-Sung Park (ORCID <https://orcid.org/0000-0003-4258-0878>)

Department of Physical Therapy, Konyang University, Daejeon, Republic of Korea

Tel: +82-42-600-6419 E-mail: daeric@konyang.ac.kr

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright © 2024 Korean Academy of Physical Therapy Rehabilitation Science

로만 표현된다[4, 5]. 게다가 높은 ICC값은 반복적으로 측정하였을 때 측정 오차에 관한 정보가 없고 검사자 간 불일치율의 정도를 알 수 없으므로 재현성의 포괄적인 특성을 반영하는데 한계가 있다[5]. 모든 평가는 재 평가 시 무작위 측정 오차가 존재하므로 이를 정량화 할 수 없으면 검사 결과를 분석하는데 어려움이 있다[7, 8].

절대적 신뢰도 지수인 표준 오차 측정(Standard error measurement, SEM)과 최소 식별 변화(Minimal detectable change, MDC)는 무작위 측정 오차의 크기를 결정하는데 사용된다[9]. SEM은 평가 시 무작위 변수(chance variation)에 의해 나타나는 측정 오차의 크기를 의미하며[9, 10], MDC는 개개환자들의 평가 결과값이 95% 신뢰 구간에서 얼마나 일관성 있게 유지되고 있는지 표현할 수 있으며, 치료 중재 후 효과 크기에 대한 임상 의사 결정을 반영하는 중요한 지표가 된다[11, 12]. 전향적 코호트 연구에서 사용되는 SEM과 MDC는 실질적으로 서로 비교할 수 있는 단위가 있는 통계수치로 측정된 값의 변화를 서로 비교하거나 신뢰수준을 판단하기 위하여 사용된다[12]. 이 두 지수는 임상 치료와 연구 분석 결과에서 실제 측정된 값을 수치로 정량화하여 피실험자들의 시간 추이 경과에 따라 얼마나 기능적인 변화가 있었는지 또는 효과가 있을지 규명하는데 사용된다[9, 10]. 그러므로 균형 평가의 재현성을 검증하기 위해서는 신뢰도와 일치도로 평가하여야 한다.

뇌졸중 환자들의 균형 평가 도구들은 많이 개발되고 수정 보완 되었으며, 이 중 Berg Balance Scale(BBS)(14개 항목: 5점 척도, 0~56점)[13], Postural Assessment Scale for Stroke(PASS)(12개 항목: 4점 척도, 0~36점)[14]이 가장 널리 사용되고 있다. 비록 BBS와 PASS는 신뢰도와 타당도가 입증되었으나[14, 15, 16] 이 두 평가는 평가를 완료하는데 소요되는 시간이 약 10~20분이다[14, 16]. 심리측정학적인 특성을 잘 반영할 수 있는 평가도구는 실험 설계와 연구과정에서 정보 수집이 비교적 용이하고 데이터 관리와 평가 결과를 해석하는데 최소한의 시간이 요구되며 임상에서 바로 사용이 가능하여야 한다[18, 19]. 평가에 많은 시간이 소비되거나 측정 오차에 의해 데이터 수집 과정에서 체계적인 오차가 나타날 수 있으며 평가도구의 구조적인 문제로 인하여 측정값이 실제 값과 차이가 발생하여 부정확한 평가 결과를 초래할 수 있다[20, 21].

대부분의 연구자들은 4~5단계로 구성된 평가 척도가 이보다 낮은 3단계 척도 보다 환자의 기능적인 변화와 심리측정학적인 특성을 잘 반영할 수 있다고 믿고 있으나 이를 입증할 수 있는 수 근거 자료는 매우 부족하다[22]. 항목의 난이도를 높히거나 평가 척도의 점수를 단계적으로 증가시키는 것은 오히려 심리측정학적인

특성을 개선시키는데 영향을 주지 못하고 불필요한 단계를 제거함으로써 변별력의 효과를 높일 수 있다[19]. 이러한 이유로 평가 결과가 평가도구의 심리측정학적인 특성에 영향을 주지 않고 데이터 정보의 손실을 최소화하는 범위내에서 척도의 단계를 줄여 평가할 수 있는 단순화하는 작업이 필요하다[20-22]. 이를 고려하여 서로 영향을 줄수 있는 문항을 분리하거나 축소하여 동일한 척도로 구성된 축소형 BBS(7개 항목: 3점 척도, 0~28점)[21]와 PASS(5개 항목: 3점 척도, 0~15점)가 개발되었으며[20], BBS와 PASS에 대한 축소형 평가는 원본 평가와 심리측정학적인 특성이 유사하고 높은 반응률과 일치도 및 예측 타당도가 입증되었다[22].

BBS와 PASS 평가도구의 축소형 평가는 임상과 연구설계에 적용이 가능하며 축소형 평가의 재현성은 검증되었다[20, 21]. 그러나 국내는 축소형 BBS와 PASS 평가의 측정자간, 검사-재검사 신뢰도 검증[23, 24]은 있었으나 검사 재검사의 재현성에 관하여 보고된 적이 없다. 따라서 본 연구의 목적은 만성 뇌졸중 환자들의 축소형 BBS와 PASS의 검사 재검사 재현성과 일치율 및 MDC를 조사하고자 하였다.

## 연구 방법

### 연구 대상자

본 연구에 참여한 피실험자들은 뇌졸중으로 인하여 편마비로 진단받고 ○○병원에서 의학적인 치료를 받고 있는 뇌졸중 환자로 연구기간은 2023년 5월부터 2024년 3월까지 시행하였다. 본 연구 대상자 선정 기준은 첫째, 뇌경색, 뇌출혈로 인하여 편마비 진단을 받은 6개월이 지난 만성 뇌졸중 환자, 둘째, 인지 능력이 한국판 정신 상태 검사(Mini Mental State Examination - Korean version, MMSE-K)에서 24점 이상인 자, 셋째, 구두지시에 따라 균형능력을 평가할 수 있는 자, 제외기준은 첫째, 전정계 이상 또는 균형에 영향을 주는 약물을 복용하거나 무시, 복시가 있는 자, 둘째, 하지의 정형 외과적인 질환으로 균형 평가를 수행할 수 없는 환자는 제외하였다. 본 연구 대상자의 연구 윤리를 지키기 위하여 피실험자들에게 연구의 목적, 절차, 위험 등을 투명하게 설명을 하여 이해할 수 있도록 하였으며 피실험자들의 자발적인 동의하에 본 연구에 참여하도록 하였다. 피실험자들의 개인정보를 보호하기 위해 기능수행 평가 점수 이외에 사적인 정보를 수집하지 않았으며 적절한 보안 조치를 시행하고, 데이터를 익명화하여 식별이 불가능하도록 하였다. 피실험자들의 낙상과 안전을 최우선으로 생각하고, 연구 절차나 실험 과정에서 발생할 수

있는 위험을 최소화하기 위한 조치를 취하였다. 연구 결과를 공유함으로써 피실험자들에 대한 보답을 제공하고, 연구 과정에 대한 투명성을 유지하였다. 본 연구는 교차 분석(cross sectional study) 연구로 투명성은 처리 그룹에 대한 정보를 실험자와 연구자가 알지 못하도록 하였고 이를 위해 피실험자와 관련된 정보를 제한하여 결과를 분석하였다. 본 연구의 표본수 산출은 G-power 분석(3.1.9.7)에서 신뢰도 ICC = 0.90 검증(유의수준 0.05, 효과 크기 0.5)을 위해서 95%의 파워 검정력에 필요한 최소한의 표본 크기가 42명이었다. 축소형 BBS와 PASS의 검사-재검사 신뢰도(ICC3,1)는 이 평가 도구를 사용해 본 경험이 있는 물리치료사에 의해 첫 평가 후 7일 간격으로 재평가하여 일치율을 비교하였다[25]. 평가자의 기억 효과를 최소화하는 것은 검사-재검사의 신뢰도를 증가시키므로[25] 한 세션에 최소 5명의 환자를 평가하였다. 본 연구에서 일반의학적인 특성은 입·퇴원기록지와 환자와의 면담을 통해 나이, 유명기간, 진단명, 마비 부위, MMSE-K를 수집하였다. 3개의 평가는 1일 또는 2일에 걸쳐 2회 연습 과정을 거친 후 3회 평균값을 사용하였고, 피로도를 최소화하기 위하여 평가 시 1~3분의 휴식을 갖도록 하였으며, 피실험자들의 낙상과 예측지 못한 위험 예방을 위하여 평가자 2명 외 보조 1인이 참여하였다. 최종 데이터 수집과정에서 응급 퇴원(2명)을 제외하고 최종 40명의 데이터를 수집하여 분석하였다.

## 측정 도구

### 버그 균형 척도(Berg Balance Scale, BBS)

BBS는 기능적인 기립 균형을 평가하는 것으로 앉기, 서기 자세, 자세 변화 등의 크게 3개 영역으로 구성되어 있다. 최소 0점~최대 4점을 적용하여 총 14개 항목에 대한 총점은 56점이다[13]. 축소형 BBS(7개 항목: 3점 척도)는 원본 BBS(14개 항목: 5점 척도, 0-4)중 3, 4단계(2, 3점)체계를 사용하지 않으며, BBS의 1-2-3을 단일 등급(2)으로 바꾸어 0-2-4로 기록하게 되어 있다[21, 22] 본 연구에서 사용된 축소형 BBS-3점 척도는 선 자세에서 앞으로 팔을 뻗어 내밀기, 두 눈을 감고 서있기, 한 발을 다른 발 앞에 일자로 두고 서 있기, 좌·우측으로 뒤돌아보기, 바닥에 있는 물건을 집어 들기, 한 다리로 서기, 앉은 자세에서 서기 7개 항목으로, 만점은 28점이며, 측정자간 신뢰도 ICC = 0.99이다[21].

### 뇌졸중 자세 평가 척도(Postural Assessment Scale for Stroke, PASS)

PASS는 3가지의 기본적인 자세인 눕기, 앉기, 서기

로 이루어져 있고 자세유지 5항목과 자세 변환 7항목으로 총 12항목으로 구성되어 있으며, 최소 0점에서 최고 3점을 적용하여 총 36점이 만점이다. 축소형 PASS(5개 항목: 3점 척도)는 원본 PASS(0-1-2-3)의 1-2를 단일 등급으로 바꾸어 축소형 PASS-3점 척도(0-1.5-3)로 구성되어 있다[20, 22]. 축소형 PASS의 평가 항목은 테이블 가장자리에서 앉은 자세에서 똑바로 눕기, 누운 자세에서 테이블 가장자리로 앉기, 앉기 자세에서 서기, 서기 자세에서 앉기, 비마비측으로 서기로 총 5개 항목으로 이루어져 있다. 4개 항목의 점수 기준은 0점은 수행할 수 없다. 1.5점은 외부도움을 통하여 수행할 수 있다. 3점은 독립적으로 수행이 가능하다. 비마비측으로 서기는 0점은 수행할 수 없다, 1.5점은 10초 미만 가능하다. 3점은 10초 이상 가능하다 로되어 있다. 5개 항목의 총점은 15점으로 측정자간 신뢰도 ICC = 0.97이다[20].

### 수정된 바델 지수(Modified Barthel Index, MBI)

MBI는 일상생활 동작 수행여부를 평가하는 것으로 환자의 기능적인 독립수준의 변화를 반영한다. 10가지의 구체적인 일상생활 활동으로 구성되어 있으며, 검사-재검사 신뢰도  $r = 0.89$ , 검사자간 신뢰도  $r = 0.95$  로 알려져 있다[26].

## 통계 방법

본 연구에서 윈도우 7 SPSS 18.0을 이용하여 통계적 분석을 시행하였다. 모든 자료는 Shapiro-wilk 검정 방법을 통해 정규성 검정을 하였고, 대상자들의 일반적인 특성은 빈도분석과 기술통계를 하였다. 3개의 체간 조절 평가의 검사-재검사 신뢰도(test-retest reliability)는 급간 내상관계수(Intra Class Coefficient, ICC2,1)를 이용하였는데  $ICC \geq 0.80$ 이면 검사 재검사에서 재현성이 우수한 것으로 간주된다[27]. 절대적 신뢰도인 Standard Error of Measurement(SEM)(standard deviation of all test-retest score) $\times \sqrt{1-ICC}$ 으로 계산하여 측정 오차를 정량화 하였고, Minimal Detectable Change(MDC)( $1.96 \times SEM \times \sqrt{2}$ )는 95% 신뢰구간에서 개개 피실험자들이 실제로 변화될 수(개선 또는 악화)있는 점수를 결정하기 위하여 사용하였다[28].  $MDC\% [MDC / \text{means of measurements taken} \times 100\%]$ 는 독립적인 측정 단위로 검사-재검사 재현성을 비교하기 위하여 사용하였다[28, 29]. SEM은 모든 검사-재검사 평균 점수의 20%미만, MDC는 획득 가능한 가장 높은 점수의 <20%인 경우 수용할만하며 그 값이 작을수록 신뢰할 만하다[30]. MDC%는 평균값이 <30%인 경우 수용할 만하다[10, 30]. 게다가 검사-재검사에 체계적인 오차(systematic error)가 발생되었는지

검증하기 위하여 두 검사 값의 평균에 차이를 t-검정을 이용하여 분석하였다. 또한 측정값들의 불일치의 양상을 알아보기 위하여 Bland and Altman 산점도 분석을 이용하였다[8]. 이는 검사 재검사 평가에서 값에 차이가 있는지 산점도로 표기되며, 95% 일치 한계 값을 추정하였다. 모든 통계학적 유의수준  $\alpha = 0.05$ 로 하였다.

## 연구 결과

### 1. 피실험자들의 일반적인 특성

본 연구의 피실험자 40명 중 남자는 18명(45%), 여

자는 22명(55%), 나이는 평균 62.69세이었다. 유병기간은 평균 8.82개월, 발병 원인으로 뇌경색 23명(57.5%), 뇌출혈 17명(42.5%), 마비측 부위는 좌측 17명(42.5%), 우측 23명(57.5%), MMSE-K는 평균 26.80점, 바델 지수는 평균 78.45점이었다(Table 1).

### 2. SF-BBS와 SF-PASS의 검사-재검사 신뢰도

SF-BBS 개별 항목 중 2번의 가중치 카파 계수는 0.77로 일치율은 우수, 1, 3~7항목은 .82~.89로 매우 우수하였다. SF-PASS 개별 항목 중 1, 5번의 가중치 카파 계수는 0.77~0.79로 일치율은 우수하였고, 2~5 항목은 .80~.89로 매우 우수한 것으로 확인되었다(Table 2).

**Table 1.** Characteristics of the stroke patients

(n=40)

Characteristic	n (%) or M ± SD	min~max
Gender (male/female)	18 (45) / 22 (55)	
Age (y)	62.68 ± 13.79	26~1
Onset (months)	8.82 ± 2.88	6~7
Diagnosis (infarction/hemorrhage)	23 (57.5) / 17 (42.5)	
Side of hemiplegia (left/right)	17 (42.5) / 23 (57.5)	
MMSE-K (score)	26.80 ± 1.43	24~0
Barthel Index (score)	78.45 ± 13.66	53~8

MMSE-K: Mini mental state examination-Korean version

**Table 2.** Test-retest reliability of the Short-Form Berg Balance Scale and the Postural Assessment Scale for Stroke patients.

Items of SF-BBS	% Agreement	wKappa	p-value	Interpretation
1 Sitting to standing	90	0.81	0.001	Very good
2 Standing with eyes closed	87	0.77	0.001	Good
3 Reaching forward with outstretched arm	88	0.85	0.001	Very good
4 Retrieving object from floor	90	0.82	0.001	Very good
5 Turning to look behind	92	0.88	0.001	Very good
6 Standing with one foot in front	89	0.82	0.001	Very good
7 Standing on one foot	91	0.85	0.001	Very good

Items of SF-PASS	% Agreement	wKappa	p-value	Interpretation
1 Standing on nonparetic leg	95	0.77	0.001	Good
2 Supine to sitting up on the edge of the table	92	0.80	0.001	Very good
3 Sitting on the edge of the table to supine	92	0.80	0.001	Very good
4 Sitting to standing up	95	0.89	0.001	Very good
5 Standing up to sitting down	90	0.78	0.001	Good

SF-BBS: Short form Berg Balance Scale, wKappa: weighted kappa

SF-PASS: Short form Postural Assessment Scale for Stroke Patients

### 3. SF-BBS와 SF-PASS의 스피어만 상관계수와 문항 내적 일치도

SF-BBS의 개별항목과 총점 간의 상관계수는  $r = 0.78 \sim 0.87$ 로 양의 높은 유의한 관련성이 있었고, SF-BBS의 크론바 알파는 0.83, 개별항목을 제거한 경우에도 0.79~0.84로 우수하였다. SF-PASS의 개별 항목과 총점 간의 상관계수는  $r = 0.75 \sim 0.83$ 으로 양의 높은 유의한 관련성이 있었고, SF-PASS의 크론바 알파는 0.82,

개별항목을 제거한 경우 0.76~0.82로 우수한 것으로 확인되었다(Table 3).

### 4. SF-BBS와 SF-PASS의 검사-재검사 재현성 지수

본 연구에서 SF-BBS와 SF-PASS의 ICC는 각각 0.941(95% 신뢰 구간 0.901~0.973), 0.932(95% 신뢰 구간, 0.889~0.962)으로 일치율은 높은 것으로 확인되었다. SF-BBS의 SEM은 1.34(3.03, 검사-재검사간 두

**Table 3.** The spearman correlation coefficient and internal consistency of the SF-BBS and SF-PASS

Items of SF-BBS	Spearman coefficient (r)	Cronbach's alpha if item deleted
1 Sitting to standing	0.79	0.84
2 Standing with eyes closed	0.85	0.81
3 Reaching forward with outstretched arm	0.84	0.82
4 Retrieving object from floor	0.87	0.81
5 Turning to look behind	0.84	0.78
6 Standing with one foot in front	0.78	0.78
7 Standing on one foot	0.87	0.80
Items of SF-PASS		
1 Standing on nonparetic leg	0.79	0.81
2 Supine to sitting up on the edge of the table	0.77	0.74
3 Sitting on the edge of the table to supine	0.77	0.75
4 Sitting to standing up	0.83	0.82
5 Standing up to sitting down	0.75	0.76

SF-BBS: Short form Berg Balance Scale, SF-PASS: Short form Postural Assessment Scale for Stroke Patients, The spearman coefficient between the SF-BBS & SF-PASS-total score and individual's items

**Table 4.** Test-retest reproducibility indices of the SF-BBS and SF-PASS

Measures	Mean (SD)		Mean difference (SD)	ICC <sub>2,1</sub> (95% CI)	SEM	MDC (%)	p
	Median (Q1~3)						
	Range (Min~max)						
1st test	2nd test						
SF-BBS	15.58 (5.96)	14.73 (5.86)	0.85 (1.91)	0.941 (0.901~.973)	1.34	3.71 (24.48)	0.385
	16 (10.25~0) 4~8	14 (10~8) 4~6					
SF-PASS	11.06 (2.39)	10.84 (2.43)	0.22 (0.86)	0.932 (0.889~.962)	0.61	1.69 (15.43)	0.310
	12 (9.37~3.5) 6~5	11.25 (9~2) 6~5					

SD: standard deviation, Q1: first quartile, Q3: third quartile, SF-BBS: Short form Berg Balance Scale, SF-PASS: Short form Postural Assessment Scale for Stroke Patients, ICC: Intraclass correlation coefficient, CI: confidence interval, SEM: standard error of measurement = standard deviation of all the test-retests score  $\times \sqrt{1-ICC}$ , MDC: minimal detectable change =  $1.96 \times SEM \times \sqrt{2}$ , MDC % = (MDC/ mean of measurements taken)  $\times 100\%$ ,

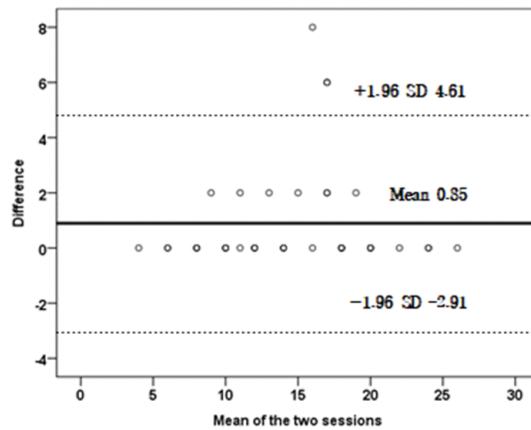


Figure 1(a). Short Form - Berg Balance Scale

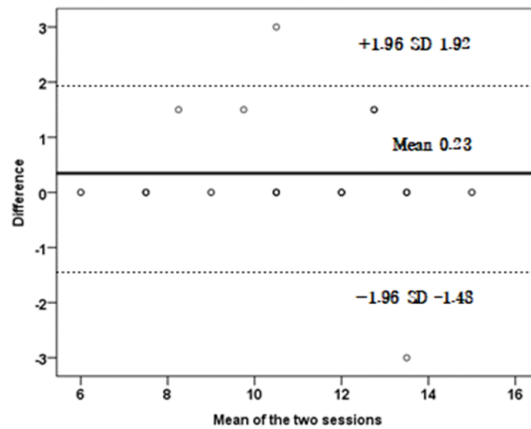


Figure 1(b). Short Form - Postural Assessment Scale for Stoke

**Figure 1.** Bland-Altman plot for the difference between measures from the two test sessions against the mean of the two test sessions for each subjects The LOA for the SF-BBS and the SF-PASS ranged from  $-2.91$  to  $4.61$  and  $-1.48$  to  $1.92$ , respectively. (a) Short form Berg Balance Scale (SF-BBS) (b) Short form Postural Assessment Scale for Stroke Patients (SF-PASS). ——— : mean difference, - - - - - the 95% limits of agreement (LOA). The 2 dashed lines define the 95% LOA (mean of difference  $\pm 1.96 \times$  SD of the difference).

평균 점수의 20%미만), SF-PASS의 SEM은 0.61(2.19)로 신뢰할 수 있었다. SF-BBS와 SF-PASS의 MDC는 각각 3.71, 1.69로 획득 가능한 최고 평가 점수의 20%미만으로 수용할만하였다. SF-BBS와 SF-PASS의 MDC%는 각각 24.48%, 15.43%로 두 평가 모두 30%미만으로 측정 오차 수준은 신뢰할만한 수준이었다. 또한 두 평가 모두 검사-재검사 간에 평균차이에 대한 체계적인 오차는 없었다(각각  $p = 0.310, 0.385$ ). SF-BBS의 LOA범위는  $-2.91 \sim 4.61$ , SF-PASS는  $-1.48 \sim 1.92$ 로 두 검사간 오차의 범위는 수용할 만하였다(Table 4)(Figure 1).

### 논의

본 연구의 목적은 만성 뇌졸중 환자들의 축소형 BBS와 PASS의 검사 재검사 재현성을 검증하고자 하였다. 축소형 BBS와 PASS의 검사-재검사 신뢰도는 이전 원본 평가와 유사하였다[18, 20-22]. 게다가 두 평가의 개별 항목에서도 가중치 카파계수와 일치율 모두 ‘ 좋음’에서 ‘매우 좋음’이었다. 축소형 BBS 총점(0.941)과 PASS 총점(0.932)의 검사 재검사 신뢰도가 높았는데 이는 만성 뇌졸중 환자에게도 충분히 신뢰성이 있음을 입증하는 것이다. 이전 원본 평가 (BBS: 0.98, PASS: 0.97)와 비교하여도 유사함을 확인할 수 있었다[17]. 비

록 ICC로 검사-재검사 측정 사이에 재현성의 크기를 대표할 수 있으나[10]. 높은 ICC값은 반복 측정사이 측정 오류에 대한 정보를 제공할 수 없을 뿐만 아니라 작은 측정 오류도 해석할 수 없으므로 진정한 재현성 평가로 활용하는데 제한이 있다[11, 12].

임상에서 환자의 기능적인 결과를 반영할 수 있는 평가 방법으로 사용하기 위해서는 실제 반복적으로 측정된 결과값이 무작위 측정 오차 없이 일관성 있게 유지되고 있는지 또는 실제 계측된 결과 값의 변화가 무작위 오차 때문에 발생하는지 이를 규명할 수 있어야 한다. SEM과 MDC는 동일한 검사를 재 평가하였을 때 결과 값의 변화가 무작위 오차 없이 95% 신뢰 구간에서 체계적으로 일정하게 유지하고 있는지 또는 변화하고 있는지 판단할 수 있는 지수이다[5, 6]. SF-BBS의 개별 항목과 총점 간의 상관계수는  $r=0.78\sim 0.87$ 로 양의 높은 유의한 관련성이 있었고, SF-BBS의 크론바 알파는 0.83, 개별항목을 제거한 경우에도 0.79~0.84로 우수하였다. SF-PASS의 개별 항목과 총점 간의 상관계수는  $r=0.75\sim 0.83$ 로 양의 높은 유의한 관련성이 있었고, SF-PASS의 크론바 알파는 0.82, 개별항목을 제거한 경우 0.76~0.82로 우수한 것으로 확인되었다. 이는 두 축소형 평가 모두 개별 항목만으로도 뇌졸중 환자의 동적 균형과 자세조절 능력을 평가하는데 특성을 충분히 반영하고 있다. 또한 SF-BBS와 SF-PASS척도내에서 개별 항목들은 매우 밀접한 관련성이 있고 이동성을 평가하는데 유사하도록 구성되어 있다.

본 연구에서 SF-BBS와 SF-PASS의 SEM은 각각 1.34, 0.61로 평균 점수의 20%미만, SF-BBS와 SF-PASS의 MDC%는 각각 24.48%, 15.43%로 두 평가 모두 30% 미만으로 측정 오차 수준은 신뢰할만한 수준이었다. 또한 SF-BBS와 SF-PASS의 MDC는 각각 3.71, 1.69로 최고 획득 가능한 점수의 20% 미만으로 수용할 만하다. MDC는 치료사의 관점에서 치료 후 환자의 기능이 얼마나 좋아졌는지 또는 향후 얼마만큼 개선될 수 있는지 예측이 가능한 역치 값이다. 즉 치료 중재 후 효과 크기의 정도를 규명할 수 있는 기준 값으로 사용되기 때문에 임상 의사 결정 시 중요한 지표로 활용이 된다[11, 12].

본 연구 결과 SEM과 MDC는 95% 신뢰 구간에서 측정 오차 없이 체계적으로 일정하게 유지되고 있으므로 만성 뇌졸중 환자들에게 적용할 수 있는 충분히 민감한 평가 도구임을 확인할 수 있었다. 임상적인 의의로 본 연구에 참여한 피실험자들은 현재 SF-BBS(검사-재검사 평균 15.16점)와 SF-PASS(검사-재검사 평균 11.63점) 수행 점수 보다 추후 각각 약 3.71점, 1.69점으로 증가하여 기능적으로 유의미하게 개선될 수 있음을 의미하는 것으로 이는 치료 효과의 크기를 말한다. SF-BBS

와 SF-PASS의 SEM, MDC, MDC(%)를 신뢰할 수 있는 이유는 첫째, ICC 값이 상대적으로 평균 0.932~0.941로 높고 둘째, 검사-재검사 간에 평균값에 대한 T-검증에서( $p$ 값 0.31 ~0.38) 유의한 차이가 없으며 정규 분포하였기 때문에 체계적인 오차는 발생되지 않았다. 셋째, 이전 선행 연구의 SEM, MDC를 추정하는 공식에 근거하여 정량적인 수치를 도출한 결과 ICC값이 매우 높을 수록 SEM과 MDC 값은 낮아지는 역의 관계에 있으므로 오차 범위는 신뢰할 수 있는 것으로 규명할 수 있다[12]. 이는 본 연구 결과를 지지한 것으로 명확하게 입증되었다.

SF-BBS와 SF-PASS의 검사자 간 일치율의 크기를 추정하거나 측정 차이의 정규성을 검증하는 방법으로 Bland and Altman 방법을 이용하여 시각화 하였다[8]. 이는 측정지 간의 평균 차이에 관한 산점도로 95% 일치 한계 수준을 의미하며, 임상 연구의 신뢰성(ICC)을 더욱 더 보강할 수 있는 대안을 제시한다. 95% 일치 한계(점선) 범위에서 SF-BBS는 -2.91~4.61, SF-PASS는 -1.48~1.92 측정된 값 간의 차이들의 표기 값은 95% 일치 한계 수준이 상한과 하한 사이에 분포하고 있다. 이러한 결과는 검사 재검사의 평균값의 크기에 따라 측정 차이 값이 달라지지 않으며 획득 가능한 전체 점수 범위안에 존재하므로 정규분포하고 있다는 것이다. 따라서 SF-BBS와 SF-PASS 검사 모두 검사 재 검사간의 측정된 값은 실제 피실험자들의 예측 가능한 값에 분포하고 있으므로 신뢰할 수 있다[8]. 또한 측정된 값이 95% 신뢰 구간 내에 분포하므로 평균값이 체계적인 오차로 발생되고 있지 않음을 의미한다. 이는 이전 원본 평가와 동일한 것으로 95% 내에서 허용할 수 있는 변동성 범위에 존재하며 Bland and Altman 산점도에는 체계적인 오류가 없음을 나타낸다[8]. 본 연구 결과에서 만성 뇌졸중 환자들의 SF-BBS는 동적 균형, SF-PASS는 자세조절 능력을 측정하는데 매우 재현성이 높은 평가 방법이라고 할 수 있는데 신뢰할 수 있는 검사 방법은 ICC 값이 매우 높고 SEM과 MDC 값은 낮아야 한다[9-11]. 따라서 본 연구에서 SF-BBS, SF-PASS의 ICC 값은 높고, SEM과 MDC 값은 낮은 수준으로 충분히 재현성이 검증되었다. SF-BBS와 SF-PASS는 임상 현장에서 평가 소요 시간의 단축, 데이터 수집이 용이하며, 뇌졸중 환자의 균형, 자세조절을 평가하는데 쉽게 적용이 가능하다.

본 연구의 제한점으로 첫째, 한정된 지역에서 생활하고 있는 뇌졸중환자들을 대상으로 하였기에 모든 연구 결과를 일반화할 수 없고, 둘째, 피실험자들의 MBI 평균 점수는  $78.45\pm 13.66$ 인 경도의존(75점~90점)[26]이 있는 환자들로 SF-BBS와 PASS평가가 가능한 환자를

대상으로 하였다. 따라서 추후 발병 초기 뇌졸중 환자와 중증 또는 장애 정도에 따른 대규모 표본을 대상으로 한 SF-BBS, SF-PASS의 임상적 유용성을 검증할 수 있는 연구가 필요하다.

## 결론

만성 뇌졸중 환자를 대상으로 한 SF-BBS, SF-PASS의 검사 재검사간의 재현성은 입증되었고, 두 평가 모두 체계적인 오류 없이 MDC는 매우 신뢰할만한 수치였다. 따라서 SF-BBS와 SF-PASS는 임상 연구에서 동적 균형과 자세 조절 능력의 기능적인 변화를 감지할 수 있는 민감하고 유용한 평가 도구가 될 수 있을 것이다.

## References

- Berg K, Norman KE. Functional assessment of balance and gait. *Clin Geriatr Med.* 1996;12(4):705-723.
- Tyson SF, Hanley M, Chillala J, Selley A, Tallis RC. Balance disability after stroke. *Phys Ther.* 2006;86(1):30-38.
- Carr J, Shepherd R. *Stroke Rehabilitation: guidelines for exercise and training to optimize motor skill.* London; Butterworth Heinemann; 2003.
- De Vet HC, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. *J Clin Epidemiol.* 2006;59(10):1033-1039.
- De Vet HC, Terwee CB, Ostelo RW, Beckerman H, Knol DL, Bouter LM. Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. *Health Qual Life Outcomes.* 2006; 4:54.
- Atkinson G, Nevill AM. Statistical methods for assessing measurement error(reliability) in variables relevant to sports medicine. *Sports Med.* 1998;26(4):217-238.
- Statistical methods for assessing measurement error(reliability) in variables relevant to sports medicine. *Sports Med.* 1998;26(4):217-238.
- Bland JM, Altman DG. A note on the use of the intraclass correlation coefficient in the evaluation of agreement between two methods of measurement. *Comput Biol Med.* 1990;20(5):337-340.
- Kovacs FM, Abaira V, Royuela A, Corcoll J, Alegre L, Tomas M, et al. Minimum detectable and minimal clinically important changes for pain in patients with nonspecific neck pain. *BMC Musculoskelet Disord.* 2008;9:43.
- Beckerman H, Roebroeck ME, Lankhorst GJ, Becher JG, Bezemer PD, Verbeek AL. Smallest real difference, a link between reproducibility and responsiveness. *Qual Life Res.* 2001;10(7):571-578.
- Flansbjer UB, Blom J, Brogardh C. The reproducibility of berg balance scale and the single-leg stance in chronic stroke and the relationship between the two tests. *PM R.* 2012;4(3):165-170.
- Schuck P, Zwingmann C. The 'smallest real difference' as a measure of sensitivity to change: a critical analysis. *Int J Rehabil Res.* 2003;26(2):85-91.
- Berg K, Wood-Dauphinee S, Williams JI. The Balance Scale: reliability assessment with elderly residents and patients with an acute stroke. *Scand J Rehabil Med.* 1995;27(1):27-36.
- Brennan P, Silman A. Statistical methods for assessing observer variability in clinical measures. *BMJ.* 1992;304(6840):1491-1494.
- Mao HF, Hsueh IP, Tang PF, Sheu CF, Hsieh CL. Analysis and comparison of the psychometric properties of three balance measures for stroke patients. *Stroke.* 2002; 33(4):1022-1027.
- Blum L, Korner-Bitensky N. Usefulness of the Berg Balance Scale in stroke rehabilitation: a systematic review. *Phys Ther.* 2008;88(5):559-566.
- Liaw LJ, Hsieh CL, Lo SK, Chen HM, Lee S, Lin JH. The relative and absolute reliability of two balance performance measures in chronic stroke patients. *Disabil Rehabil.* 2008;30(9):656-661.
- Liaw LJ, Hsieh CL, Hsu MJ, Chen HM, Lin JH, Lo SK. Test-retest reproducibility of two short-form balance measures used in individuals with stroke. *Int J Rehabil Res.* 2012;35(3):256-262.
- Hobart JC, & Thompson AJ. The five item Barthel index. *Journal of Neurology Neurosurgery, and Psychiatry.* 2001;71(2):225-230.
- Chien CW, Lin JH, Wang CH, Hsueh IP, Sheu CF, Hsieh CL. Developing a Short Form of the Postural Assessment Scale for People With Stroke. *Neurorehabil Neural Repair.* 2007;21(1):81-90.
- Chou CY, Chien CW, Hsueh IP, Sheu CF, Wang



- CH, Hsieh CL. Developing a short form of the Berg Balance Scale for people with stroke. *Phys Ther*. 2006;86(2):195-204.
22. Wang CH, Hsueh IP, Sheu CF, Yao G, Hsieh CL. Psychometric Properties of 2 Simplified 3-Level Balance Scales Used for Patients With Stroke. *Phys Ther*, 2004;84(5):430-438.
23. An SH, Kim JH, Song CH. The Comparison of Postural Assessment Scale for Stroke. (PASS : 5items-3Level) and Berg Balance Scale(BBS : 7items-3Level) Used for Patients with Stroke. *Journal of the Korean Society of Physical Medicine*. 2010;5(1):89-99.
24. Park CS, Choi YI, An SH. The Comparison of Simplified Postural Assessment Scale for Stroke and Berg Balance Scale Used for Stroke Patients. *The Journal of Korean Society of Occupational Therapy*. 2010;18(1):65-77.
25. Chen HM, Hsieh CL, Lo SK, Liaw LJ, Chen SM, Lin JH. The test-retest reliability of 2 mobility performance tests in patients with chronic stroke. *Neurorehabil Neural Repair*. 2007;21(4):347-352.
26. Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. *Md State Med J*. 1965;14:61-65.
27. Prince B, Makrides L, Richman J. Research methodology and applied statistics. Part 2: The literature search. *Physiother Can*. 1980;32(4):201-206.
28. Beckerman H, Roebroeck ME, Lankhorst GJ, Becher JG, Bezemer PD, Verbeek AL. Smallest real difference, a link between reproducibility and responsiveness. *Qual Life Res*. 2001;10(7):571-578.
29. Smidt N, van der Windt DA, Assendelft WJ, Mourits AJ, Deville WL, de Winter AF, et al. Inter-observer reproducibility of the assessment of severity of complaints, grip strength, and pressure pain threshold in patients with lateral epicondylitis. *Arch Phys Med Rehabil*. 2002;83(8):1145-1150.
30. Schuck P, Zwingmann C. The 'smallest real difference' as a measure of sensitivity to change: a critical analysis. *Int J Rehabil Res*. 2003;26(2):85-91.