Journal of The Korean Society of Integrative Medicine, 2017, 5(1), 43~53
https://doi.org/10.15268/ksim.2017.5.1.043
ISSN 2288-1174(Print)
ISSN 2383-9651(Online)

만성 뇌졸중 환자들의 Step Test의 상대적·절대적 신뢰도와 타당도

이병권[1] · 최현수[2≠] · 안승헌[3]
[1]건양대학교 의과학대학 물리치료학과, [2≠]건양대학교 융합IT학과,
[3]국립재활병원 물리치료실

The Relative·Absolute Reliability and Validity of Step Test in Patients with Chronic Stroke

Lee Byoungkwon, PT, Ph.D[1] · Choi Hyunsoo, Ph.D[2≠] · An Seungheon, PT, Ph.D[3]
[1]Dept. of Physical Therapy, Konyang University of Medicine and Science
[2≠]Dept. of Convergence Information Technology, Konyang University
[3]Dept. of Physical Therapy, National Rehabilitation Center

Abstract

Purpose : To examine the relative·absolute reliability and validity of step test (ST) scores in subjects with chronic stroke.

Method : A total of 27 stroke patients, participated in the study. A relative reliability index (intraclass correlation coefficient, ICC) was used to examine the level of agreement of inter-rater·test-retest reliability for ST score. Absolute reliability indices, including the standard error of measurement(SEM) and the minimal detectable change (MDC), and limits of agreement by Bland and Altman analysis. The validity was demonstrated by spearman correlation of ST score with 10 m Walk Test (10mWT), Fugl-Meyer Assessment-Lower/Extremity (FMA-L/E)-total score, Berg Balance Scale (BBS)-total score.

Result : An excellent inter-rater reliability in ST scores was found (paretic, ICC=0.993～0.996; nonparetic, ICC=0.982～0.991). In addition, excellent test-retest reliability was found (paretic, ICC=0.992; nonparetic, ICC=0.967). It all showed acceptable SEM of the ST score as paretic and nonparetic were 0.22 and 0.46 respectively (average score <10 %), and the MDC of the paretic and nonparetic were 0.61 and 1.27 respectively (possible highest score <20 %). indicating that measures had a small and acceptable measurement error. The ST score of paretic and nonparetic were also found to be significantly associated with 10MWT (r=0.77～0.79), FMA-LE scores (r=0.73～0.81) and BBS scores (r=0.72～0.76).

Conclusion : The ST showed highly sufficient Inter-rater·test-retest agreement and validity and acceptable measurement errors caused by due to chance variation in measurement. It also can be used by clinicians and researchers to assess the balance and mobility performance and monitor functional change in chronic stroke patients.

Key Words : balance, mobility performance, absolute reliability, step test, stroke

≠교신저자: 안승헌 ptlove1@hanmail.net

논문접수일 : 2017년 02월 6일 | 수정일 : 2017년 03월 6일 | 게재승인일 : 2017년 03월 8일

Ⅰ. 서 론

1. 연구의 배경 및 필요성

뇌졸중 환자의 독립적인 일상생활동작 수행 능력은 동적 균형, 마비측 상·하지 운동 조절 및 하지 근력에 영향을 받는다(Desrosiers 등, 2003; Patterson 등, 2007). 특히 뇌졸중 환자들의 지역 사회 참여와 활동 영역의 제한은 상지 보다 하지 운동 기능과 매우 밀접한 관련이 있다(Patterson 등, 2007). 이 중 뇌졸중 환자의 동적 균형 소실은 예측치 못한 자세 조절, 외적인 동요 유발, 자가 움직임을 조절 할 수 없는 상태에서 발생한다(Geurts 등, 2005). 동적 균형 능력은 평지 보행과 Rivermead 이동 지수(Rivermead Mobility Index, RMI)평가에서 기능적인 이동성에 중요한 요소로 알려져 있다(Tason 등, 2004; Tason 등, 2006). 게다가 동적 균형과 마비측 하지 운동 조절 능력은 방향 전환 하기(회전), 계단 오르내리기, 테이블과 길모퉁이 돌기, 주변을 탐색하면서 걷기와 같은 독립보행에서 필수 요소이다(Goh 등, 2013; Wong 등, 2013). 따라서 뇌졸중 환자의 독립 보행은 지역 사회 참여 활동과 여가 생활을 누리는 등 삶의 만족도에 기여하므로 물리치료영역에서 가장 중요한 목표 중 하나 이다(Blennerhassett 등, 2012; Mayo 등, 2002).

임상에서 뇌졸중 환자의 동적 균형 능력은 버그 균형 척도(Berg Balance Scale, BBS)(Berg 등, 1992). 마비측 하지 운동 조절은 푸글 마이어 하지 평가(Fugl Meyer Assessment-Lower/Extremity)(Fugl-Meyer 등, 1975)를 가장 많이 사용되고 있으며, 입원 기간과 유병 기간에 따른 기능 변화(민감도, 특이도), 반응률과 같은 심리학적인 특성이 잘 검증되었다(Chen 등, 2014; Chen 등, 2015). 그럼에도 불구하고 BBS를 완전히 평가하는데 평균 20분 이상(Chou 등, 2006; Stevenson, 2001), FMA의 세부 항목인 상·하지 운동 기능, 균형, 감각, 협응, 반사 항목을 평가하는데 대략 58분 이 소요된다(Hsieh 등, 2007; Malouin 등, 1994). 이러한 구조적인 문제로 인하여 평가 항목의 복잡함과 번거로움은 치료사와 환자에게 심리적인 부담감과 피로감을 주게 된다(Salter 등, 2005; Chou 등, 2006; Chen 등, 2014; Hsieh 등, 2007). 그러므로 임상 현장에서 동적 균형과 하지 운동 조절 능력을 조합한 평가 방법이 필요하다. 아울러 평가에 소요되는 시간이 짧고 쉽게 적용이 가능하며, 평가 결과에 대한 해석이 명확해야 한다(Mercer 등, 2009).

따라서 임상가들은 환자들의 기능 수행을 정확히 평가할 수 있고 향후 기능적인 변화를 예측하거나 관찰이 용이한 신뢰할만하고 타당성이 있는 평가 도구를 사용하는 것이 필요하다(Hong 등, 2012). Hill 등(1996)에 의해 개발된 Step 검사(Step test, ST)는 뇌졸중 환자의 동적 서기 균형을 평가하기 위해 고안되었고 검사에 소요되는 시간은 5분 미만이다. 이는 7.5 ㎝높이의 발판을 앞에 두고 한쪽 하지(마비측)는 지면에 두고 다른 한쪽 하지(비마비측)는 15초 동안 가능한 빠르게 반복적으로 발판 위에 발을 올렸다 내리도록 하는 검사이다. ST는 노인의 검사-재검사 신뢰도(ICC=0.90)(Hill 등, 1996), 뇌졸중 환자의 검사-재검사 신뢰도(ICC=0.93)(Hill 등, 1996), 측정자간 신뢰도(ICC=0.996～0.999)와 측정자내 신뢰도(ICC=0.981～0.995)(Hong 등, 2012)는 보고되었다. Hong 등(2012)의 연구에서 ST의 타당도는 마비측 하지 근력(고관절 굴곡, 신전, 슬관절 신전, 족관절 배측 굴곡근), 균형(Berg Balance Scael, BBS), 하지 협응(lower-extremity motor coordination test, LEMOCOT), 5 m 보행 검사와의 유의미한 관련성을 증명하는데 연구에 제한점(표본수 15명)이 있다. 부가적으로 ST 스텝 점수의 표준 평균응답(Standardized response mean, SRM)으로 본 효과크기는 평균 0.92～0.95로 높아 천장 효과는 없다(Bernhardt 등, 1998). 이는 뇌졸중 발병 후 회복기 동안 기능 변화를 감지하는데 충분히 민감도가 높다고 할 수 있다.

또한 ST는 균형 장애가 있는 노인을 선별할 수 있고 낙상군과 비낙상군을 구별할 수 있다(Hill 등, 1996), 뇌졸중 환자의 낙상 위험을 예측할 수 있는 ST 스텝 점수의 선별 기준값은 7점 이하(Mackintosh 등, 2006). <10점(≥10점 인 환자들보다 낙상할 확률이 2.1배 높음)으로 보고되었다(Blennerhassett 등, 2012). 건강한 성인과 뇌졸중 환자들의 양측 하지 ST 수행의 차이를 구별 할 수 있는 스텝 수는 마비측 13점, 비마비측 11점으로 알려져 있다(Hong 등, 2012). 그러나 ST 스텝 점수의 절대적 신뢰도인 측정의 표준 오차(Standard Error Measurement, SEM)와 최소 감지 변화(Minimal Detectable Change, MDC)는 알려져 있지 않다. 또한 국내에선 ST에 대한 임상 적용과 그 사례 연구는 보

고된 적이 없다. 따라서 본 연구의 목적은 첫째, ST 검사의 상대적 신뢰도와 절대적 신뢰도를 구하고 둘째, 뇌졸중 특이성 장애를 반영하는 보행 속도와 마비측 하지 운동 조절 및 균형과의 관련성을 알아보고자 하였다.

Ⅱ. 연구방법

1. 연구대상 및 기간

본 연구의 피 실험자들은 뇌졸중으로 인하여 편마비 진단을 받고 연구에 동의한 만성 뇌졸중 환자 들을 대상으로 하였다. 피 실험자들은 G병원에서 통상적인 의료서비스(의학, 간호, 물리, 작업치료 등)를 받고 있는 환자들로 본 연구 기간은 2016년 10월~2016년 12월 까지 시행되었다. 연구 대상자의 선정 기준은 보행 보조 도구 없이 10 m 이상 독립 보행이 가능하고, 인지 기능이 한국판 간이 정신 상태 검사(Mini Mental State Examination-Korean version, MMSE-K)에서 >24점, 뇌졸중 발병 후 6개월이 지난 만성 뇌졸중 환자를 대상으로 하였다. 족관절 저측 굴곡근의 강직이 Composite Spasticity Score(CSS)에서 >10인 자(Levin & Hui-Chan, 1992)와 수동 움직임 시 족관절 배측 굴곡이 ≤10° 인자, 하지의 정형외과적인 질환으로 본 연구 실험과정에서 평가를 수행 할 수 없는 환자는 제외하였다.

2. 연구 방법 및 평가

본 연구에서 피실험자들의 일반의학적인 특성은 입원 기록지와 1:1면담을 통해 연령, 유병기간, 진단명, 마비 부위, MMSE-K를 수집하였다. 표본 수 산출은 이전 연구에 근거하였다. 뇌졸중 환자의 ST 검사-재검사 신뢰도 ICC= 0.93이다(Hill 등, 1996). 검사-재검사 일치율 검증(신뢰 수준 0.05)에서 ICC=0.93을 검증하기 위해서는 파워 분석에서 96 % 파워 검정력에 필요한 표본 수의 크기는 14명인 것으로 분석되었다. 그러나 변수들간의 유의한 관련성(타당도)을 검증하기 위해서는 표본 수의 크기는 최소 20명 이상이어야 한다(Curtin & Schulz, 1998). 본 연구의 표본수 산출은 G Power 분석에서 효과 크기 f=0.3, 93 % 파워 검정력(유의수준 0.05)의 반복 측정(3회)분산 검증에서 표본 수는 28명 이었다. 일반적으로 피 실험자들의 탈락률(20 %)을 감안하여 35명을 대상으로 하였고, 검사-재검사 일치율에서 피실험자들의 동적인 균형 능력의 잠재적인 회복의 효과를 최소화하기 위하여 뇌졸중 발병 후 6개월이 지난 만성 뇌졸중 환자들을 대상으로 하였다(Liaw 등, 2008). ST 검사에서 피 실험자들은 보행 보조 도구 없이 수행이 가능하고, 안전과 낙상의 위험에 대비하여 보조 치료사 1인의 관찰과 감독 하에 실시하였다. 연구 대상자 선정 기준에 부합하지 않은 3명과 최종 자료 수집 과정에서 3명은 건강 악화로 인한 응급 퇴원과 연구 불참, 신뢰할 수 없는 인지 장애로 인하여 평가 결과 누락 2명으로 최종 27명의 데이터가 수집되었다. ST의 측정자간 신뢰도($ICC_{2,1}$)검증은 검사자(A, B)를 무작위로 할당하여 2(X, Y)일에 걸쳐 피 실험자들의 ST 스텝 점수를 제 각기 평가하였다. 검사-재검사 신뢰도($ICC_{3,1}$)는 7일 간격으로 두 차례 동일한 피 실험자들을 대상으로 검사자 1인(A)에 의해 재평가 되었으며(Liaw 등, 2008), 모든 값은 3회 시도한 평균값을 이용하여 일치율을 비교하였다. ST 스텝 점수의 타당도는 10 m 보행 속도, FMA-L/E총점, BBS 총점와의 상관계수를 구하였다. 피 실험자들의 기능 수행 평가는 임상 경력이 15년 차인 물리치료사 2인에 의해 평가되었다.

1) Step 검사(ST)

뇌졸중 환자의 동적 서기 균형을 평가하기 위해 ST를 이용하였다(Hill 등, 1996). ST 수행 과제는 피 실험자 전방에 7.5 ㎝높이의 발판을 두고 한쪽 하지(마비측)는 지면에 놓고 반대측 하지(비마비측)는 15초 동안 가능한 빠르게 반복적으로 발판 위에 발을 올리고 내리도록 한다. 모든 검사는 마비측과 비마비측 모두 15초 이내에 완전히 수행하는데 소요된 스텝 수를 기록하여 점수화 하였다. ST는 마비측 하지 운동 기능과 협응 조절 뿐만 아니라 단하지 입각기에 있는 하지는 반대측 하지의 디딤 동작에 필요한 체중 이동과 움직임 시 동적인 균형 능력을 유지해야 한다. 마비측 발의 디딤 동작에서 마비측 하지는 빠르게 굴곡을 해야 하는 반면 비마비측 발의 디딤 동작에서는 마비측 하지는 완전한 신전과 동시에 체중 부하와

이동이 필요하다. 즉 ST는 한 쪽 하지의 발 디딤 동작을 하는 동시에 반대측 입각기에 있는 하지 위로 무게 중심을 안정적으로 유지하기 위하여 적절한 하지 근력이 필요하다(Hill 등, 1996). ST는 측정자간 신뢰도(ICC=0.996～0.999)와 측정자내 신뢰도(ICC=0.981～0.995)는 보고되었다(Hong 등, 2012).

2) 10 m 보행 검사

10 m 보행 검사는 뇌졸중 환자의 보행 속도를 측정하는데 많이 사용하고 있으며, 12 m 보행 시 2 m 가속구간과 감속구간을 뺀 10 m 구간을 보행하는데 소요된 시간을 측정하는 검사이다. 검사-재검사 신뢰도 ICC=0.93으로 보고되었다(Salbach 등, 2001).

3) 푸글 마이어 하지 운동 평가(Fugl Meyer Motor Assessment for the lower extremities, FMA-L/E)

FMA-L/E는 뇌졸중 환자의 주동근과 협력근의 움직임, 협응 및 반사를 포함하여 뇌졸중 환자의 하지 운동 기능 장애를 평가하는데 사용되고 있다. 이는 17개 항목(0～2점)으로 총 만점은 34점이다(Fugl-Meyer 등, 1975). FMA-L/E의 측정자간 신뢰도 ICC=0.83～0.95(Duncan 등, 1983), 측정자내 신뢰도 ICC=0.92(0.81～0.96)로 보고되었다(Sanford 등, 1993).

4) 버그 균형 척도(Berg Balance Scale, BBS)

BBS는 노인의 낙상과 균형 능력을 평가하기 위해 고안되었다. 일상생활에서 흔히 접하게 되는 앉기 균형 1개 항목과 13개의 서기 항목을 평가하도록 구성되어 있으며, 최근에는 뇌졸중 환자의 균형 평가에도 가장 널리 사용되고 있다. BBS는 14개의 과제로 구성되어 있으며, 각 과제는 5점 척도(0～4점)로 만점은 56점이다. 뇌졸중 환자의 측정자간 신뢰도(ICC=.97～.98)는 높은 것으로 보고되었다(Berg 등, 1992).

3. 분석 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. ST 스텝 점수의 측정자간·검사-재검사 신뢰도(Inter-rater reliability)는 상대적 신뢰도 지수인 급간내상관계수(Intra Class Coefficient, ICC)를 이용하였다. 또한 절대적 신뢰도 지수인 무작위 측정 오류를 정량화 할 수 있는 표준 측정 오차(Standard Error Measurement, SEM= 검사-재검사 점수의 표준 편차 ×[√1-ICC])와 최소 감지 변화(Minimal Detectable Change, MDC=1.96×SEM×√2)를 이용하였다. SEM은 평균 점수의 <10 %로 그 값이 작을수록(Atkinson과 Nevill, 1998) MDC는 측정한 값 중 최고점수의 <20 %인 경우 신뢰할 수 있다(Lu 등, 2008). 또한 검사-재검사의 두 검사 사이 일치율의 크기를 추정하기 위하여 Bland and Altman plot을 이용하였다(Bland & Altman, 1986). 이는 평균에 대한 검사-재검사 두 검사 사이의 차이에 대한 산점도로 95 % 일치 한계 수준을 의미한다. 95 % 일치 한계(95 % limits of agreement, LOA= 측정자내 두 검사 측정값의 평균 차이 ± 1.96 × 차이의 표준편차)는 두 검사 사이의 자연적인 변이(natural variation)를 결정하기 위해 사용되는데 이 간격이 좁을수록 매우 신뢰 할 수 있다(Altman & Bland, 1983). ST 스텝 점수의 타당도는 10 m 보행 속도, FMA-L/E총점, BBS총점간의 관련성을 스피어만 상관 계수(Spearman Correlation Coefficient)로 구하였으며, 모든 통계학적 유의수준은 α=.05로 하였다.

Ⅲ. 결 과

1. 연구 대상자들의 일반적인 특성과 기능 수행 평가

본 연구에 참여한 피 실험자들의 성별은 남자 14명(51.8 %), 여자 13명(48.1 %), 평균 연령은 59.04세 로 발병 원인은 뇌경색 17명(63 %), 뇌출혈 10명(37 %), 마비측 부위는 좌측 12명(44.4 %), 우측 15명(55.6 %), 유병기간은 8.96개월 이었으며, MMSE-K는 26.93점 이었다. 기능 수행 평가에서 ST 스텝점수는 마비측 6.75점 비마비측 7.94점, 10 m 보행 검사는 0.81 ㎧이었다. FMA-L/E는 평균 20.52점, BBS는 44.73점 이었다(표 1).

표 1. 연구 대상자들의 일반적인 특성과 기능 수행 평가 (N=27)

구분	n(%) 또는 평균±표준편차(최소～최대)
성별: 남자/여자	14(51.8 %)/13(48.1 %)
연령(세)	59.04±10.04(34～80)
발병 원인: 뇌경색/뇌출혈	17(63 %)/10(37 %)
마비측 부위: 좌측/우측	12(44.4 %)/15(55.6 %)
유병기간(개월)	8.96±3.06(6～18)
MMSE-K(점)	26.93±1.73(24~30)
ST 마비측 스텝 수(점)	6.75±2.58(2～11)
ST 비마비측 스텝 수(점)	7.94±2.68(2～12)
10m 보행 검사(m/s)	0.81±0.15(0.53～1.08)
FMA-L/E(점)	20.52±3.78(13～26)
BBS(점)	44.73±3.56(38～51)

MMSE-K: Mini Mental State Examination-Korean, ST: Step Test, FMA-L/E: Fugl Meyer Motor Assessment for the lower extremities, BBS: Berg Balance Scale

2. ST 스텝 점수의 측정자간 신뢰도

마비측 ST 스텝 점수의 측정자간 신뢰도 ICC=0.993～0.996(0.984～0.998), 비마비측은 ICC=0.982～0.991(0.961～0.996)로 일치율은 매우 높은 것으로 확인되었다(표 2).

표 2. 뇌졸중 환자들의 ST 스텝 점수의 측정자간 신뢰도

평가자	일	마비측 스텝 점수 (평균±표준편차)	$ICC_{2,1}$(95 % CI)	비마비측 스텝 점수 (평균±표준편차)	ICC2,1(95 % CI)
A-B	X	6.77±2.69	0.996(0.991～0.998)	7.90±2.89	0.982(0.961～0.992)
	Y	6.79±2.67	0.993(0.984～0.997)	7.87±2.95	0.991(0.980～0.996)

ICC(95 % CI): Intra Class Coefficient(95 % Confidence Interval)

3. ST 스텝 점수의 검사-재검사 신뢰도와 절대적 신뢰도

마비측 ST 스텝 점수의 검사-재검사 신뢰도 ICC=.992(.983～.997), 비마비측은 ICC=.967(.929～.985)로 매우 높은 것으로 확인되었다. ST 스텝 점수의 SEM은 마비측 0.22(평균 6.78회), 비마비측 0.46(평균 7.93회)으로 검사 1과 2의 평균 점수의 <10 %미만이었고, MDC는 마비측 0.61(11회), 비마비측은 1.27(12회)로 획득 가능한 최고 스텝 점수의 <20 %로 신뢰 할만 수준인 것으로 나타났다(표 3). Bland & Altman plot에서 마비측 스텝 점수의 일치 한계(limit of agreement, LOA; 실선)의 범위는 -0.62~+0.63, 비마비측 Step 점수의 일치 한계는 -1.29~+1.35로 검사-재검사에서 추정된 두 값은 실제 피 실험자들의 계측 가능한 값에 분포하고 있으므로 신뢰할 수 있는 수준인 것으로 확인되었다(표 3)(그림 1, 2).

표 3. ST 스텝 점수의 검사-재검사 신뢰도와 절대적 신뢰도

구분	평균±표준편차(최소～최대)		$ICC_{3,1}$(95 % CI)	SEM	MDC
	검사 1	검사 2			
마비측 스텝 점수	6.74±2.62(2～11)	6.73±2.47(3～10)	.992(.983～.997)	0.22	0.61
비마비측 스텝 점수	7.92±2.58(3～12)	7.95±2.59(3～11)	.967(.929～.985)	0.46	1.27

ICC(95 % CI): Intra Class Coefficient(95 % Confidence Interval), SEM: standard error measurement, MDC: minimal detectable change

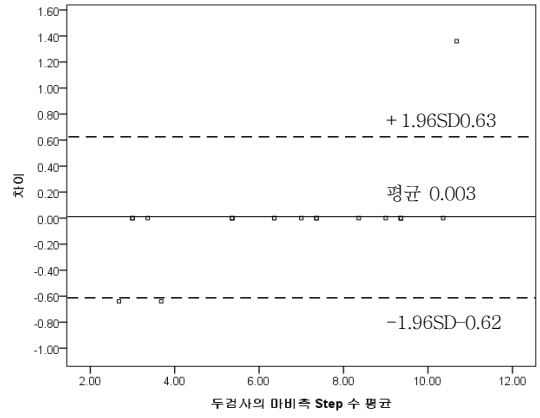

그림 1. 두 검사의 마비측 ST 스텝 점수 평균값과 측정값의 차이에 관한 Bland-Altman plot

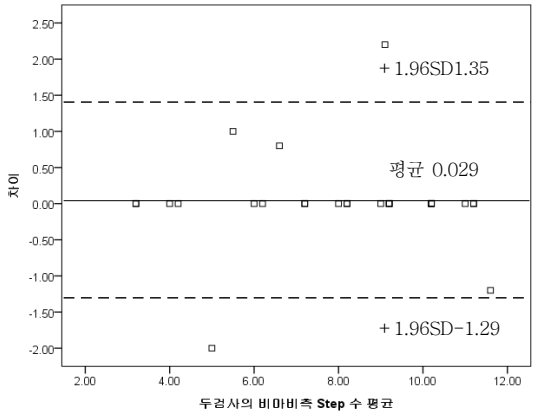

그림 2. 두 검사의 비마비측 ST 스텝 점수 평균값과 측정값의 차이에 관한 Bland-Altman plot

4. ST 스텝 점수, 10 m 보행 속도, FMA-L/E, BBS와의 상관관계

마비측과 비마비측 ST 스텝 점수는 10 m 보행 속도(r=0.77～0.79), FMA-L/E(r=0.73～0.81), BBS(r=0.72～0.76)과는 높은 상관관계가 있는 것으로 확인되었다(p<.01)(표 4).

표 4. ST 스텝 점수, 10m 보행 속도, FMA-L/E, BBS와의 상관관계

구분	10 m 보행속도	FMA-L/E	BBS
마비측 ST 스텝 점수	0.79*	0.81*	0.76*
비마비측 ST 스텝 점수	0.77*	0.73*	0.72*

*p<.01, ST: Step Test, FMA-L/E: Fugl Meyer Motor Assessment for the lower extremities, BBS: Berg Balance Scale

Ⅳ. 고 찰

본 연구의 목적은 만성 뇌졸중 환자들을 대상으로 ST 스텝 점수의 상대적·절대적 신뢰도와 뇌졸중 특이성 장애인 보행 속도, 마비측 하지 운동 기능 및 균형과의 관련성을 조사하고자 하였다. 본 연구 결과 ST 스텝 점수의 측정자간 신뢰도(마비측, ICC=0.993～0.996; 비마비측, ICC=0.982～0.991)와 측정자내 신뢰도(마비측, ICC=0.992; 비마

비측, 0.967)의 일치율은 매우 높은 것으로 나타났다. 이는 이전 연구와 비교하여 검사-재검사 신뢰도 ICC=0.93(Hill 등, 1996), Hong 등(2012)의 측정자간 신뢰도(마비측, ICC=0.997～0.998; 비마비측, 0.996~0.998), 측정자내 신뢰도(마비측, ICC=0.992～0.993; 비마비측, ICC=0.981～0.983)와 일치하는 수준이다. Hong 등(2012)의 연구에서 숙련가와 비숙련가의 ST의 마비·비마비측 스텝 점수의 측정자간 신뢰도 ICC=0.998(0.994～1.000)로 매우 높은 것으로 확인되었다. ST는 측정의 오차를 줄이는 데 표준화된 프로토콜(Protocol)과 검사 지시 사항, 규격화된 실험 장비, 검사의 내용 및 채점 방법이 명확하였다. 이는 신뢰도 검증에서 불필요한 바이어스를 최소화하였기 때문에 이전 연구와 본 연구에서도 높은 신뢰도가 있음을 입증하는데 기여한 것이다. 또한 특별한 교육 없이 비전문가의 ST 평가(Hong 등, 2012)에서도 높은 신뢰도가 검증되었고 검사에 소요되는 시간은 5분미만으로 임상에서 뇌졸중 환자의 동적 균형 능력을 쉽고 빠르게 적용하여 평가 할 수 있다. 본 연구에서 ST의 절대적 신뢰도인 SEM과 MDC를 조사하였는데 이전 연구에선 보고된 적이 없다. 일반적으로 상대적 신뢰도인 ICC값이 높을수록 신뢰도가 높지만 재평가에 따른 학습 효과로 인해 측정 오차가 발생한다(Patten 등, 2003; Weir, 2005).

모든 평가는 측정에서 우연 오차(random error)가 존재하는데 이는 관찰 점수(진점수+오차)를 수집하게 된다. 그러나 임상가와 연구가들은 환자들의 평가를 반복 측정하였을 때 평가 점수가 일정하게 유지되고 있는지 아니면 실제 점수 변화가 측정 오차 때문에 발생한 것인지 결정하기가 어렵다. ICC는 반복 측정 시 검사 점수들의 일관성을 나타내지만 개개인의 진점수(관찰점수-오차)를 파악할 수 없고, 평가 사이 불일치와 측정 오차의 크기를 알 수 없다(Flansbjer 등, 2012; Lu 등, 2008). 평가에서 발생하는 임의 측정 오차는 절대적 신뢰도 지수인 SEM과 MDC에 의해 정량화 시킬 수 있다(Flansbjer 등, 2005; Flansbjer 등, 2012).

두 지수 모두 동일한 평가를 반복적으로 측정하였을 때 우연 변동(chance variation)에 의해 발생할 수 있는 임의 측정 오차(Random measurement error)의 범위를 결정하거나 실제 측정된 값이 95 % 신뢰수준에서 일정하게 유지되고 있는지 추정하기 위해 사용된다(Flansbjer 등, 2012; Lu 등, 2008). SEM은 측정 시 우연변동에 의해 발생하는 측정 오차의 크기를 의미하므로 피실험자들의 평가 점수에 관한 신뢰도에 관한 정보를 제공한다.

특히 MDC는 임상적인 관점에서 환자들의 치료 계획과 관리 측면에서 볼 때 치료 중재 후 효과를 규명하고 기능적인 회복을 예측할 수 있는 평가 방법으로 활용되고 있다(Flansbjer 등, 2012; Lu 등, 2008). 선행 연구에 의하면 물리치료 평가에서 사용하고 있는 버그 균형 척도(Berg Balance Scale, BBS), 뇌졸중 자세 평가 척도(Postural Assessment Scale for Stroke, PASS)(Hiengkaew 등, 2012; Liaw 등, 2008), 푸글 마이어 기능 평가(Fugl Meyer Assessment, FMA)(Hiengkaew 등, 2012; Hsueh 등, 2008), 뇌졸중 재활 움직임 평가 척도(Stroke Rehabilitation Assessment of Movement measure, STREAM), Rivermead 이동 지수(Rivermead Mobility Index, RMI)(Chen 등, 2007), 동적 보행 지수(Dynamic Gait Indexm, DGI), 기능적 보행 평가(Functional Gait Assessment, FGA)(Lin 등, 2010), 하지 운동 협응 검사(Lower Extremity Motor Coordination Test, LEMOCOT)(de Menezes 등, 2015), 일어나 걸어 가기 검사(Timed up & go test, TUG), 2분 보행 검사, 보행 속도(Hiengkaew 등, 2012) 등 많은 연구에서 절대적 신뢰도 지수는 보고되었고 이에 대한 연구는 활발히 진행되고 있다.

또한 MDC는 임상 연구에서 임상가와 연구가에게 있어 개개 환자들의 계측된 값과 추정된 값 사이의 실제 변화 점수가 향후 얼마나 좋아질 수(또는 악화 될 수) 있는지 치료 후 효과에 대한 크기를 결정할 수 있는 기준값으로 사용되므로 임상 의사 결정에 필요한 중요한 근거 자료로 활용이 가능하다(Flansbjer 등, 2012; Lu 등, 2008). 본 연구에서 ST 스텝 점수의 SEM은 마비측 0.22, 비마비측 0.46으로 ST의 평균 스텝 점수의 <10 %미만, MDC는 마비측 0.61, 비마비측 1.27로 획득 가능한 최고 스텝 점수의 <20 %로 신뢰 할만 수준 이었고 두 지수 모두 측정 오차는 수용할만한 수준이었다. 본 연구에 참여한 피실험자들의 현재 ST 마비·비마비측의 스텝 점수는 각각 0.61, 1.27이나 향후 2.2회, 2.4회 이상 개선 될 수 있음을 의미한다. 매우 신뢰할 수 있는 검사는 ICC값이 높고 MDC값은 낮아야 하는데(Chen 등, 2015) 이는 본 연구에서 증명되었다.

또한 Bland & Altman plot에서 마비측 Step 점수의 일치한계 범위는 -0.62~+0.63에서 폭은 1.25, 비마비측 Step 점

수의 일치 한계는 −1.29~+1.35에서 폭은 2.64로 두 검사의 평균값과 차이는 어떠한 관련성은 없었다. 마비측·비마비측 스텝 점수의 추정된 값은 실제 피 실험자들의 계측 가능한 값에 분포하고 있으므로 신뢰할 수 있다. 따라서 Bland −Altman plot의 0이 95 % 신뢰 구간 내에서 존재하고 있으므로 계측한 값에 체계적인 오차가 없음을 의미한다(Liaw 등, 2008).

본 연구 뇌졸중 환자의 ST 스텝 점수는 마비·비마비측은 각각 평균 6.75점, 7.94점 이었다. 이전 연구(Hill 등, 1996; Bernhardt 등, 1998; Mercer 등, 2009)에 따르면 뇌졸중 발병 후 유병일이 8주에서 6개월이 지난 환자들의 경우 마비측 6.4～7.3점, 비마비측은 6.5～7.5점 이었고, 뇌졸중 발병 후 평균 5.6년인 환자는 각각 8.1점, 11점으로 보고되었다(Hong 등, 2012). 급·아급성기 환자들은 만성 뇌졸중 환자들에 비해 기능적인 능력이 낮은 것으로 알려져 있다(Jorgensen 등, 1995). 본 연구의 피 실험자들의 유병기간은 8.96개월인 환자들로 Hong 등(2012)의 피 실험자들보다 ST수행은 낮은 수준이다. 뇌졸중 발병 후 ST 수행은 비사용 증후로 인한 보상적인 전략과 움직임에 영향을 받기 때문에 움직임의 질은 무시된다. 그러나 명확한 것은 건강한 노인(70세 이상)은 17점(Hill 등, 1996), 건강한 성인(57.3세)은 19점(Hong 등, 2012)에 비교하여 뇌졸중 환자들의 ST 수행은 현저히 낮다. 본 연구의 마비·비마비측 ST 스텝 점수는 10 m 보행 속도(r=0.77～0.79)와 유의한 상관관계가 있는 것으로 확인되었다. 이는 이전 연구와 일치하는 결과(Hsu 등, 2003; Lin 등, 2006)로 뇌졸중 환자의 마비측 ST 스텝 점수는 5 m 보행 속도(r=0.69)와 유의한 관련이 있고(Hong 등, 2012), 마비측의 발 디딤 동작의 하지 근력은 ST 스텝 수에 영향을 준다. 그러나 Hong 등(2012)의 연구에서 비마비측의 ST 스텝 점수(r=0.67)와는 관련이 없다고 하였는데 이 는 표본 수의 크기가 작기 때문이며 최소 20명은 되어야 유의한 관련성을 증명 할 수 있다(Curtin 등, 1988). Mercer 등(2009)에 의하면 마비·비마비측 ST 스텝 점수는 자가 선택 보행 속도(60~79 %)와 뇌졸중 영향 척도(Stroke Impact Scale)의 이동성 항목(34~50 %)에 영향을 주고, 보행 속도가 0.07 ㎧～0.09 ㎧로 증가할수록 유의한 관련이 있다고 하였다. 이러한 관계는 본 연구 피 실험자들의 보행 속도는 0.81 ㎧(ST 마비측 스텝 점수, 6.75점; 비마비측 스텝 점수, 7.94점), Mercer 등(2009)의 연구에서 보행 속도는 0.60～0.76 ㎧(마비측, 5.19～6.41점; 비마비측, 5.81～7.48점)이었고, Hong 등(2012)의 보행 속도는 1.27 ㎧(마비측, 8.1점; 비마비측, 11점)로 비교될 만하였다. 보행 속도가 증가할수록 ST 스텝 점수 또한 증가하는 것으로 나타나 ST는 뇌졸중 환자의 동적 균형과 이동 능력(보행 속도)을 반영할 수 유용한 평가 도구가 될 수 있다.

Hong 등(2012)은 마비·비마비측의 ST는 마비측 하지 운동 협응 검사 점수(lower extremity motor coordination test, LEMOCOT)(r=0.62～0.90)와 관련이 있고, Mercer 등(2009)에 따르면 ST는 균형 장애와 마비측 하지 운동 조절을 평가하는데 적합하다고 하였다. ST의 마비측 수행 시 비마비측은 단하지 입각기로 마비측 하지는 고관절과 슬관절의 신전에서 굴곡 패턴으로 부드럽게 전환해야 하는 협응력과 근력이 필요하다. 반대로 비마비측 수행 시 마비측 단하지 입각기 하지의 근력은 중력 중심점을 안정화시키고 동적인 균형을 유지하면서 비마비측 하지를 전방으로 빠르게 발 디딤 동작에 필요한 협력근으로 작용해야 한다(Wang 등, 2006). 임상에서 많이 사용되고 있는 FMA-L/E는 하지 운동 기능 조절과 반사, 굴곡과 신전의 조합된 협력 움직임 및 협응을 측정할 수 있는 포괄적인 평가이다(Fugl-Meyer 등, 1975). ST 스텝 점수와 FMA-L/E의 두 변수간의 관련성은 보고되지 않았다. 뇌졸중 환자의 FMA-L/E(평균)에 따른 ST 스텝 점수의 비교에서 본 연구 피실험자들의 FMA-L/E는 평균 20.52점(마비측, 6.75점; 비마비측, 7.94점), Mercer 등(2009)의 연구에 의하면 FMA-L/E는 17.82점(마비측, 6.41점; 비마비측 7.48점)이었다. 이는 본 연구에서 마비·비마비측의 ST 스텝 점수는 FMA-L/E(r=0.73～0.81)와 관련이 있음이 증명되었다. ST는 임상관점에서 볼 때 일상생활에서 빠르게 발걸음을 옮기기, 장애물을 탐색하여 피하기, 넘기, 회전하기 등 동적인 균형(무게 중심 조절) 소실 없이 발을 끌지 않고 충분히 들어 올릴 수 있는 근력(족관절 배측 굴곡근)과 협응 능력(입각기 와 유각기의 하지 굴곡·신전 패턴)이 필요하다(Said 등, 2005). 이는 FMA-L/E 평가 항목에 포함되어 있으므로 ST는 마비측 하지 운동 조절 능력과는 밀접한 관련이 있다. 본 연구의 마비·비 마비측 ST 스텝 점수는 BBS(r=0.72～0.76)와 관련성이 있었는데 이는 Hong 등(2012)의 연구와 일치하였다. ST 비마비측 스텝 점수는

BBS(r=0.73)와 유의한 관련이 있다고 하였는데 ST는 가능한 빨리 신속하게 효율적인 체 중심조절과 더불어 동적인 균형 조절 능력이 필요하다. 더욱이 BBS평가 14개 항목 중 2개 항목(교대로 양측발을 올리고 내리기, 외발 서기)은 ST 수행 과제를 포함하고 있으므로 ST와의 관련성은 명확하다. 본 연구 실험 과정에서 평가자는 ST 수행 과정에서 피 실험자가 입각기의 하지는 균형을 유지하고 반대측 하지는 교대로 발 디딤 동작에 필요한 하지의 협응 움직임을 반복적으로 수행 할 수 있는지 주의 깊은 관찰이 필요하였다. 본 연구의 제한점으로 ST는 15초 동안 가능한 얼마만큼 빨리 수행 할 수 있는지(스텝 수)에 관심을 둘 뿐 움직임의 질은 무시되었다. 또한 표준화된 7.5 ㎝ 높이의 계단(Hill 등, 1996)은 피 실험자의 키와 다리 길이의 변동성 때문에 모든 피 실험자에게 적합하지 않을 수 있다. 게다가 ST는 균형에 현저히 장애가 있거나 환측 하지로 수행하기 어려운 환자에겐 적용하기 어렵다. 더욱이 ST수행에 영향을 줄 수 있는 하지의 강직, 발의 고유수용성 감각 및 낙상에 대한 심리적 두려움은 평가하지 못하였다. 본 연구 피 실험자들은 기능이 우수한 환자들을 대상으로 하였기 때문에 위 기준에 부합한 환자들에게 본 연구 결과를 일반화할 수 있을 것 이다.

V. 결 론

ST는 측정자간·검사-재검사 일치율과 타당도는 매우 높고 우연변동에 의해 발생하는 측정 오차에서도 수용할 만한 평가인 것으로 확인되었다. 따라서 ST는 임상가와 연구가들이 만성뇌졸중 환자들의 기능적인 변화를 관찰하거나 동적 균형과 이동성을 평가하는데 사용될 수 있을 것이다. 본 연구는 ST의 스텝 수에 관한 신뢰도와 타당도 검증이 이루어졌으므로 추후 연구에서는 ST 수행과 관련이 있는 기능적인 변수들과의 인과관계 분석이 필요할 것이다.

참고문헌

Altman DG, Bland JM(1983). Measurement in medicine: the analysis of method comparison studies. Statistician, 32(3), 307-317.

Atkinson G, Nevill AM(1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med, 26(4), 217-238.

Berg KO, Maki BE, Williams JI, et al(1992). Clinical and laboratory measures of postural balance in an elderly population. Arch Phys Med Rehabil, 73(11), 1073-1080.

Berg KO, Wood-Dauphinee SL, Williams JI, et al(1992). Measuring balance in the elderly: validation of an instrument. Can J Public Health, 83(suppl 2), S7-S11.

Bernhardt J, Ellis P, Denisenko S, et al(1998). Changes in balance and locomotion measures during rehabilitation following stroke. Physiother Res Int, 3(2), 109-122.

Bland JM, Altman DG(1986). Statistical methods for assessing agreement between two methods of clinical measurement. Lancet, 1(8476), 307-310.

Blennerhassett JM, Dite W, Ramage ER, et al(2012). Changes in balance and walking from stroke rehabilitation to the community: a follow-up observational study. Arch Phys Med Rehabil, 93(10), 1782-1787.

Chen HM, Hsieh CL, Sing Kai Lo, et al(2007). The test-retest reliability of 2 mobility performance tests in patients with chronic stroke. Neurorehabil Neural Repair, 21(4), 347-352.

Chen KL, Chen CT, Chou YT, et al(2014). Is the long form of the Fugl-Meyer motor scale more responsive than the short form in patients with stroke? Arch Phys Med Rehabil, 95(5), 941-949.

Chen KL, Chou YT, Yu WH, et al(2015). A prospective study of the responsiveness of the original and the short form Berg Balance Scale in people with stroke. Clin Rehabil, 29(5), 468-476.

Chou CY, Chien CW, Hsueh IP, et al(2006). Developing a short form of the Berg Balance Scale for people with stroke. Phys Ther, 86(2), 195-204.

Curtin F, Schulz P(1998). Multiple correlations and Bonferroni's correction. Biol Psychiatry, 44(8), 775-777.

Desrosiers J, Malouin F, Bourbonnais D, et al(2003). Arm

and leg impairments and disabilities after stroke rehabilitation: relation to handicap. Clin Rehabil, 17(6), 666-673.

de Menezes KK, Scianni AA, Faria-Fortini I, et al(2015). Measurement properties of the lower extremity motor coordination test in individuals with stroke. J Rehabil Med, 47(6), 502-507.

Duncan PW, Propst M, Nelson SG(1983). Reliability of the Fugl-Meyer assessment of sensorimotor recovery following cerebrovascular accident. Phys Ther, 63(10), 1606-1610.

Flansbjer UB, Blom J, Brogardh C(2012). The reproducibility of berg balance scale and the single-leg stance in chronic stroke and the relationship between the two tests. PM R, 4(3), 165-170.

Flansbjer UB, Holmback AM, Downham D, et al(2005). Reliability of gait performance tests in men and women with hemiparesis after stroke. J Rehabil Med, 37(2), 75-82.

Fugl-Meyer AR, Jaasko L, Leyman I, et al(1975). The post-stroke hemiplegic patient, 1: a method for evaluation of physical performance. Scand J Rehabil Med, 7(1), 13-31.

Geurts ACH, de Haart M, van Nes IJW, et al(2005). A review of standing balance recovery from stroke. Gait Posture, 22(3), 267-281.

Goh EY, Chua SY, Hong SJ, et al(2013). Reliability and concurrent validity of four square step test scores in subjects with chronic stroke: A pilot study. Arch Phys Med Rehabil, 94(7), 1306-1311.

Hiengkaew V, Jitaree K, Chaiyawat P(2012). Minimal detectable changes of the Berg Balance Scale, Fugl-Meyer Assessment Scale, Timed "Up & Go" Test, gait speeds, and 2-minute walk test in individuals with chronic stroke with different degrees of ankle plantarflexor tone. Arch Phys Med Rehabil, 93(7), 1201-1208.

Hill KD, Berhardt J, McGann AM, et al(1996). A new test of dynamic standing balance for stroke patients: reliability, validity and comparison with healthy elderly. Physiother Can, 48(4), 257-262.

Hong SJ, Goh EY, Chua SY, et al(2012). Reliability and validity of step test scores in subjects with chronic stroke. Arch Phys Med Rehabil, 93(6), 1065-1071.

Hsieh YW, Hsueh IP, Chou YT, et al(2007). Development and validation of a short form of the Fugl-Meyer motor scale in patients with stroke. Stroke, 38(11), 3052-3054.

Hsu AL, Tang PF, Jan MH(2003). Analysis of impairments influencing gait velocity and asymmetry of hemiplegic patients after mild to moderate stroke. Arch Phys Med Rehabil, 84(8), 1185-1193.

Hsueh IP, Hsu MJ, Sheu CF, et al(2008). Psychometric comparisons of 2 versions of the Fugl-Meyer motor scale and 2 versions of the stroke rehabilitation assessment of movement. Neurorehabil Neural Repair, 22(6), 737-744.

Jorgensen HS, Nakayama H, Raaschou HO, et al(1995). Outcome and time course of recovery in stroke. Part II: time course of recovery. Arch Phys Med Rehabil, 76(5), 406-412.

Levin MF, Hui-Chan CW(1992). Relief of hemiparetic spasticity by TENS is associated with improvement in reflex and voluntary motor functions. Electroencephalogr Clin Neurophysiol, 85(2), 131-142.

Liaw LJ, Hsieh CL, Lo SK, et al(2008). The relative and absolute reliability of two balance performance measures in chronic stroke patients. Disabil Rehabil, 30(9), 656-666.

Lin JH, Hsu MJ, Hsu HW, et al(2010). Psychometric comparisons of 3 functional ambulation measures for patients with stroke. Stroke, 41(9), 2021-2025.

Lin PY, Yang YR, Cheng SJ, et al(2006). The relation between ankle impairments and gait velocity and symmetry in people with stroke. Arch Phys Med Rehabil, 87(4), 562-568.

Lu WS, Wang CH, Lin JH, et al(2008). The minimal detectable change of the simplified stroke rehabilitation assessment of movement measure. J Rehabil Med, 40(8), 615-619.

Mackintosh SF, Hill KD, Dodd KJ, et al(2006). Balance

score and a history of falls in hospital predict recurrent falls in the 6 months following stroke rehabilitation. Arch Phys Med Rehabil, 87(12), 1583-1589.

Malouin F, Pichard L, Bonneau C, et al(1994). Evaluating motor recovery early after stroke: comparison of the Fugl-Meyer assessment and the motor assessment scale. Arch Phys Med Rehabil, 75(11), 1206-1212.

Mayo N, Wood-Dauphinee S, Coyte R, et al(2002). Activity, participation, and quality of life 6 months poststroke. Arch Phys Med Rehabil, 83(8), 1035-1042.

Mercer VS, Freburger JK, Chang SH, et al(2009). Step test scores are related to measures of activity and participation in the first 6 months after stroke. Phys Ther, 89(10), 1061-1071.

Patten C, Kothari D, Whitney J, et al(2003). Reliability and responsiveness of elbow trajectory tracking in chronic poststroke hemiparesis. J Rehabil Res Dev, 40(6), 487-500.

Patterson SL, Forrester LW, Rodgers MM, et al(2007). Determinants of walking function after stroke: differences by deficit severity. Arch Phys Med Rehabil, 88(1), 115-119.

Said CM, Goldie PA, Culham E, et al(2005). Control of lead and trail limbs during obstacle crossing following stroke. Phys Ther, 85(5), 413-427.

Salbach NM, Mayo NF, Higgins J, et al(2001). Responsiveness and predictability of gait speed and other disability measures in acute stroke. Arch Phys Med Rehabil, 82(9), 1204-1212.

Sanford J, Moreland J, Swanson LR, et al(1993). Reliability of the Fugl-Meyer assessment for testing motor performance in patients following stroke. Phys Ther, 73(7), 447-454.

Salter K, Jutai JW, Teasell R, et al(2005). Issues for selection of outcome measures in stroke rehabilitation: ICF body functions. Disabil Rehabil, 27(4), 191-207.

Stevenson TJ(2001). Detecting change in patients with stroke using the berg balance scale. Aust J Physiother, 47(1), 29-38.

Tyson SF, DeSouza LH(2004). Reliability and validity of functional balance tests post stroke. Clin Rehabil, 18(8), 916-923.

Tyson SF, Hanley M, Chillala J, et al(2006). Balance disability after stroke. Phys Ther, 86(1), 30-38.

Wang Y, Zatsiorsky VM, Latash ML(2006). Muscle synergies involved in preparation to step made under the self-paced and reaction time instructions. Clin Neurophysiol, 117(1), 41-56.

Weir JP(2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. J Strength Cond Res, 19(1), 231-240.

Wong SS, Yam MS, Ng SS(2013). The figure of eight walk test: reliability and associations with stroke-specific impairments. Disabil Rehabil, 35(22), 1896-1902.