Clinical Applicability and Psychometric Properties of Manual Function Test for Patients with Stroke

Hyun Jung Kim,¹ Kyunghoon Min,² Chan Ho Lee,¹ Yun Kyung Cho,¹ Jin Soo Yoon,² Dong Rae Cho,² Kyoung Suk Woo,² Je Young Song¹ and MinYoung Kim²

¹Department of Rehabilitation Medicine, Nowon Eulji Medical Center, Eulji University, Seoul, Korea ²Department of Rehabilitation Medicine, CHA Bundang Medical Center, CHA University, Seongnam, Gyeonggido, Korea

The Manual Function Test (MFT) is a tool to assess upper extremity motor impairment associated with stroke. This study investigated psychometric properties of the Korean version of the MFT and to establish normative data. Eighty-one patients were enrolled to evaluate MFT, Fugl-Meyer Assessment (FMA) and manual muscle test (MMT). MFT was completed by eight raters on two occasions separated by 6 weeks. Absolute and relative reliability and validity were examined. Additionally, MFT was assessed on 75 healthy controls of different ages. Intraclass correlation coefficient (ICC) (2,1) values for total and each dimension of Korean MFT ranged from 0.984 to 0.998 in the affected side of hemiplegic patients, indicating inter-rater reliability. Percentage values of standard error of measurement (SEM) and smallest real difference (SRD) ranged 3.10-10.57% and 8.58-29.29% respectively. Test-retest reliability ICCs for all raters were above 0.98. Effect size (ES) and standardized response mean (SRM) were larger in the acute-subacute group (onset to initial evaluation \leq 4 months) (ES = 0.12; SRM = 0.41) than those in the chronic group (onset to evaluation > 4 months) (ES = 0.01; SRM = 0.11). MFT score was significantly correlated with FMA score (p< 0.001) and MMT score (p < 0.001). In healthy controls, regression analysis indicated that age significantly predicts manual function scores on both dominant and non-dominant. The Korean MFT showed good reliability and validity. Modest responsiveness was observed in patients evaluated early after stroke onset. The Korean MFT is useful in evaluating upper extremity motor deficits for clinical and research purposes.

Keywords: outcome assessment; psychometrics; reproducibility of results; stroke; upper extremity Tohoku J. Exp. Med., 2017 October, **243** (2), 85-93. © 2017 Tohoku University Medical Press

Introduction

Upper extremity function is crucial for performing independent activities of daily livings. Patients with stroke often exhibit persistent impairment of an upper extremity, particularly in their dexterity (Kwakkel et al. 2003). Therein, valid and accurate assessment of manual function is indispensable to documenting disability and evaluating treatment outcomes. For upper extremity functional assessments to be useful, several criteria should be fulfilled: 1) use of standardized methods or equipment; 2) comprehensively measure upper extremity function, from proximal arm movements to distal dexterity; 3) have proven reliability and validity; 4) are simple and easy to use; and 5) have normative data available for comparison (Guyatt et al. 1987; Alt Murphy et al. 2015).

The Manual Function Test (MFT) is a measuring tool of upper extremity function developed by the Research Facility for Rehabilitation Medicine at Tohoku University (Miyamoto et al. 2009). The test was originally designed to assess motor impairment caused by stroke, the results from which could be used to develop individualized therapeutic programs. Comprising simple instructions and a scoring system that is easy learn, MFT can be readily used to assess overall upper extremity function in a relatively short time period (< 10 minutes) (Miyamoto et al. 2009). Another advantage is that MFT affords clinicians the ability to assess varying degrees of upper extremity impairment, from mild

Received July 4, 2017; revised and accepted September 13, 2017. Published online October 6, 2017; doi: 10.1620/tjem.243.85. Correspondence: MinYoung Kim M.D, Ph.D., Department of Rehabilitation Medicine, School of Medicine, CHA University, 59 Yatapro, Bundang-gu, Seongnam, Gyeonggi-do 13496, Republic of Korea.

[,] Bundang-gu, Seongnann, Gyeonggi-do

e-mail: kmin@cha.ac.kr

A part of the study was presented using an interim data at the 41st Annual meeting of the Korean Academy of Rehabilitation Medicine, October 26, 2013, Seoul, Korea; the 3rd Korea-Japan NeuroRehabilitation Conference, March 22, 2014, Busan, Korea and at the American Academy of Physical Medicine and Rehabilitation 2014 Annual assembly, November 15, 2014, San Diego, CA.

to severe (Miyamoto et al. 2009). Additionally, normative data for 333 healthy, Japanese subjects are available (Michimata et al. 2008).

While the relative reliability of MFT has been reported, its absolute reliability is uncertain. Moreover, the validity of the MFT has only been assessed in comparison to Brunnstrom stage of the upper limb and hand, the Stroke Impairment Assessment Set, and Barthel index (Miyamoto et al. 2009). The MFT has not been compared with Fugl-Meyer Assessment (FMA) (Fugl-Meyer et al. 1975), one of the stroke-specific measurement tools for evaluation of comprehensive upper extremity function. Also, responsiveness, which is defined as the capacity of an instrument to detect changes over time, has not yet been evaluated for MFT. Lastly, MFT has yet to be officially translated into many languages, including Korean. Therefore, the purposes of this study were to establish a Korean version of MFT in an attempt to facilitate wider and more accurate application of MFT in Korea and to evaluate its psychometric properties, which have not been reported. Also, we aimed to obtain normative data for Koreans.

Materials and Methods

Participants

Patients undergoing inpatient and outpatient rehabilitation from two university affiliated hospitals were recruited between November 2012 and December 2013. Inclusion criteria were 1) diagnosis of cerebral hemorrhage or infarction; 2) neurological motor sequelae affecting unilateral upper extremity; and 3) willingness to be videotaped. Exclusion criteria were 1) an inability to understand and follow task instructions and/or unable to perform tasks due to poor postural control; 2) any auditory and/or visual dysfunction; 3) neurological conditions other than stroke; and 4) previous injury or any conditions associated with upper extremity functions.

For normative study, we enrolled healthy volunteers between 30 and 80 years of age who had no history of a neurological or psychiatric problem, previous injury, or condition that might affect upper extremity function.

This study was approved by the institutional review boards of the CHA Bundang Medical Center and Nowon Eulji Medical Center, Eulji University. All participants were informed of the purposes of this study and its procedures, and agreed to participate therein in writing.

Measures

Manual Function Test: The MFT is designed to determine a subject's ability to perform eight tasks related with arm motion and manipulation (Table 1). Thirty-two standardized criteria are evaluated during the assessment. According to the protocol, subjects are to attempt each task three times, and the highest score is recorded (Nakamura and Moriyama 2000). Completion of each criterion is scored as 1, and total MFT scores range from 0 to 32 points. Manual Function Score (MFS) can then be obtained by multiplying the total MFT score by a scaling factor of 3.125 for a maximum total score of 100 (Nakamura and Moriyama 2000). Further description of the MFT is provided in the MFT manual (Nakamura and Moriyama 2000).

We obtained permission from the original developer and copy-

Table 1.	Eight tasks	of the Manual	Function	Test

	Score
Arm motions	
Forward elevation of the upper limb	0-4
Lateral elevation of the upper limb	0-4
Touch the occiput with the palm	0-4
Touch the dorsum with the palm	0-4
Manipulative activities	
Grasp and pinch	
Grip	0-3
Pinch	0-3
Arm and hand activities	
Carry cubes	0-4
Peg board	0-6
Total score	32

Observation items of the Manual Function Test.

right holding company (SAKAI Medical Co., Ltd.) to translate MFT into Korean. The translation process was conducted according to general guidelines for cross-cultural adaptation (Beaton et al. 2000) as follows: 1) forward translation, for which direct translation was preferred; 2) reconciliation and synthesis by the authors for clearer understanding; 3) back translation at a notary translation office; and 4) expert review and confirmation by the MFT copyright holders.

Fugl-Meyer Assessment and muscle strength of the upper extremities: The FMA assesses motor function, sensory, balance, range of motion, and joint pain. The maximum motor score for the upper extremities is 66 points, with individual items rated on a threepoint ordinal scale. According to studies, FMA shows excellent intrarater and inter-rater reliability, as well as good construct validity (Gladstone et al. 2002; Sullivan et al. 2011).

Muscle strength of the upper extremity on the hemiplegic side was measured using the manual muscle test (MMT) (Medical Research Council (Great Britain) and University of Edinburgh Department of Surgery 1976). We assessed flexors, extensors, abductors, adductors, internal rotators, and external rotators of the shoulder; flexors and extensors of the elbow and wrist; and flexors, extensors, abductors, and adductors of the fingers. The grading system, ranging from "zero" to "normal," was converted to numerical scoring as follows: "Zero" to 0, "trace" to 10, "Poor-" to 15, "Poor" to 25, "Poor+" to 35, "Fair-" to 40, "Fair" to 50, "Fair+" to 60, "Good-" to 70, "Good" to 80, and "Normal" to 100. Thus, summed MMT scores ranged from 0 to 1,400. We utilized grades of MMT by converting into numerical values to evaluate the recovery from paralysis by adopting previous study methods (Min et al. 2013).

Study protocol

After obtaining informed consent, demographic and clinical data for each participant were collected through interviews and medical chart review. The MFT was administered to all participants while sitting in a chair. The MFT and FMA were conducted by eight proficient occupational therapists, four from each hospital. All therapists received detailed training on the measurements, and were given sufficient practice prior to administer the tests. MMT was assessed by two well-trained physicians. Each subject's performance on MFT and FMA was videotaped for initial evaluations. We divided patients according to the time at which the tests were conducted after stroke onset: acute-subacute group (onset to initial evaluation ≤ 4 months, n = 43) and chronic group (onset to evaluation > 4 months, n = 38) (Kreisel et al. 2007). For responsive analysis, we conducted followup assessment of MFT at six weeks after initial evaluation for both patient groups. Six weeks of follow-up interval was determined according to previous study results (Hsieh et al. 2007; Sale et al. 2014; da Silva et al. 2015; Wolf et al. 2015) and data availability in hospitalized patients. To determine inter-rater reliability, all assessors independently scored the videotaped performances of the initial MFT conducted by other therapists. Intra-rater reliability analysis was conducted by rescoring the video recordings two weeks after the first scoring.

Statistical analyses

Descriptive analyses were conducted for all measures in each group. For normative data, the following analyses were conducted: simple regression to assess the relationship between age and MFS score; Mann-Whitney's U-test to compare total MFT scores for subjects in their 30th decade with those for older subjects in each age decade and to compare performances between male and female subjects; and Wilcoxon's signed rank test to compare performance scores for dominant and non-dominant hands in each age group. Spearman correlation coefficients between age and scores for each dexterity task were calculated.

The internal consistency of the MFT was explored using Cronbach's alpha for total score for all subjects as well as for patients with stroke only. Values greater than 0.8 were deemed indicative of acceptable internal consistency (Nunnally 1978).

Intraclass correlation coefficients (ICCs) were calculated to obtain inter- and intra-rater reliabilities; ICC (2,1) was for inter-rater reliability, and ICC (1,1) was for intra-rater reliability. ICCs greater than 0.9 are required for clinical decision making (Nunnally 1978; Ko and Kim 2013).

Absolute reliability was measured by calculating standard error of measurement (SEM) and the smallest real difference (SRD). SEM reflects the standard deviation of measurement errors, providing information on the magnitude of score error variabilities in repeated measurements (Atkinson and Nevill 1998, Beckerman et al. 2001). The equation for calculating SEM is as follows: SEM = $SD\sqrt{(1-ICC)}$ (Atkinson and Nevill 1998), where SD is the standard deviation of the assessed scores. SEM% was calculated to determine the error magnitude independent of the units of measurement as follows: SEM% = (SEM/mean) × 100 (Flansbjer et al. 2005b). Expressing SEM as a percentage, error magnitude could be compared across samples and experimental conditions. Thus, SEM% represents the smallest change at which improvement in a group of subjects following any intervention can be detected (Flansbjer et al. 2005b). Tests with a SEM% of less than 10% are considered sensitive enough to detect clinically relevant changes (Flansbjer et al. 2005b; Liaw et al. 2008). SRD represents the smallest true change exceeding measurement noise on repeated testing (Beckerman et al. 2001). A score exceeding this range indicates a true clinical change. The SRD was calculated using SEM as follows: SRD = $1.96\sqrt{2}$ (SEM) (Beckerman et al. 2001), where 1.96 constitutes the z score associated with a 95% level of confidence and $\sqrt{2}$ signifies the variance of two measurements. SRD was also expressed as a percentage value as SRD% (Flansbjer et al. 2005b). The equation for calculating SRD% was as follows: SRD% = (SRD/mean) × 100 (Flansbjer et al. 2005b).

Responsiveness was determined by effect size (ES) (Kazis et al. 1989) and standardized response mean (SRM) (Liang et al. 1990), and paired t test (Husted et al. 2000) was used to examine the significance of the change in scores evaluated six weeks apart for acute-subacute and chronic groups respectively. ES was calculated by dividing the mean change in scores after six weeks by the standard deviation of the baseline scores (Kazis et al. 1989). SRM was calculated as the mean change in scores after six weeks divided by the standard deviation of the change in scores (Liang et al. 1990). According to Cohen's criteria, an effect size greater than 0.8 is to be considered large, 0.5 to 0.8 as moderate, and 0.2 to 0.5 as small (Cohen 1988).

To determine the construct validity of the MFT, Spearman correlation coefficients between MFT scores and FMA and MMT scores were calculated, respectively. For further analysis, we extracted the sum scores of shoulder/elbow/forearm section of FMA as an arm subscale, and sum scores of hand section as a hand subscale (Fugl-Meyer et al. 1975). We adapted the following criteria to interpret the magnitude of the correlation coefficients: > 0.6 indicated excellent validity, 0.3 to 0.6 adequate validity, and < 0.3 poor validity (Sivan et al. 2011). All data were analyzed using SPSS software for Windows, version 23.0.

Results

Characteristics of the participants

A total of 81 patients with stroke and 75 healthy controls were recruited in this study. The characteristics of the participants are reported in Table 2.

Analysis of normative data

Mean total MFT score for all healthy participants was 31.6 ± 0.8 . We observed significantly poorer performance on the MFT in individuals older than 50 compared to those in their 30s (Table 3). Regression analysis indicated age as significant predictor of lower MFS scores on the non-dominant side ($\beta = -0.60$, t [86] = -6.34, p < 0.001), and age explained 35.5% of the variance in MFS scores for the nondominant hand ($R^2 = 0.36$, F (1,73) = 40.15, p < 0.001). For the dominant side, age significantly predicted lower MFS scores ($\beta = -0.45$, t [92] = -4.27, p < 0.001) and explained 20% of the variance therein ($R^2 = 0.20$, F (1,73) = 18.20, p < 0.001) (Fig. 1). MFT scores for dominant hands were significantly better than those for non-dominant hands among subjects in their 60s and 70s (60s, z = -2.53, p =0.011 r = -0.21; 70s, z = -2.0, p = 0.046, r = -0.16); this difference was not observed in those under the age of 60 years.

The scores of two manual dexterity tasks (carrying cubes and moving pegs) decreased with increasing age. Relationship analyses revealed strong and moderate negative correlations between age and peg-board scores for

H.J. Kim et al.

Table 2. Subject characteristics.

	Patient $(n = 81)$	Healthy control $(n = 75)$
Age (years)	60.2 ± 12.4	53.9 ± 13.6
Sex (male/female)	42/39	34/41
Stroke type (n)		
Infarction	45	N.A.
Hemorrhage	36	
Time from onset to initial evaluation (months)	23.7 ± 41.3	N.A.
Acute-subacute/Chronic*	43/38	N.A.
Hemiplegic side Right/ Left	47/34	N.A.
Hemiplegic dominant hand Right/ Left	47/0	N.A.

Values are mean \pm SD for age and time from onset to initial evaluation. N.A., not applicable.

*Groups were divided according to the initial evaluation time from the stroke onset by 4 months.

Table 3. Score of total Manual Function Test in each age groups (n = 75). Dominant side

Age group	All	All (n = 75)		Male (n = 34)		Female $(n = 41)$		<i>p</i> value	
(yrs)	n	score	n	score	n	score	vs. 30s	Male vs. Female	
30-39	15	32.0 ± 0.0	6	32.0 <u>+</u> 0.0	9	32.0 <u>+</u> 0.0	_	n.s.	
40-49	15	31.9 <u>+</u> 0.4	7	31.9 <u>+</u> 0.4	8	31.9 <u>+</u> 0.4	n.s.	n.s.	
50-59	15	31.7 <u>+</u> 0.5	7	31.6 <u>+</u> 0.5	8	31.9 <u>+</u> 0.4	0.035	n.s.	
60-69	15	31.6 <u>+</u> 0.7	7	31.4 <u>+</u> 0.8	8	31.8 <u>+</u> 0.7	0.035	n.s.	
70-79	15	30.7 <u>+</u> 1.3	7	30.3 <u>+</u> 1.4	8	31.0 <u>+</u> 1.2	0.001	n.s.	

Non-dominant side

Age group	All	All (n = 75)		Male $(n = 34)$		Female $(n = 41)$		<i>p</i> value	
	n	score	n	score	n	score	vs. 30s	Male vs. Female	
30-39	15	32.0 <u>+</u> 0.0	6	32.0 <u>+</u> 0.0	9	32.0 <u>+</u> 0.0	_	n.s.	
40-49	15	31.9 <u>+</u> 0.3	7	32.0 <u>+</u> 0.0	8	31.9 <u>+</u> 0.4	n.s.	n.s.	
50-59	15	31.5 <u>+</u> 0.6	7	31.3 <u>+</u> 0.8	8	31.6 <u>+</u> 0.5	0.003	n.s.	
60-69	15	31.1 <u>+</u> 1.0	7	31.0 <u>+</u> 1.4	8	31.1 <u>+</u> 0.6	0.001	n.s.	
70-79	15	30.1 <u>+</u> 1.3	7	29.3 <u>+</u> 1.0	8	30.9 <u>+</u> 1.0	< 0.001	0.014	

Values are mean \pm SD for total MFT score.

MFT, Manual function test; n.s., not significant.

p values were calculated using the Mann-Whitney U-test.



Fig. 1. Relationship between age and Manual Function Score in healthy controls. Panel A shows normative data of dominant hands. Panel B shows normative data of non-dominant hands.

dominant and non-dominant hands, respectively ($r_s = -0.41$, n = 75, p < 0.0001: $r_s = -0.64$, n = 75, p < 0.0001). Weak correlations were identified between age and task scores on the cube carrying task for both dominant and non-dominant hands ($r_s = -0.35$, n = 75, p = 0.002; $r_s = -0.32$, n = 75, p = 0.005). Scores for other six tasks evaluating arm motion, grasping, and pinching were excellent in all age groups.

Internal consistency

Cronbach's coefficient alpha for the eight MFT tasks performed by all participants including healthy subjects was 0.966 (95% confidence interval (CI) = 0.959-0.972). For patients only, it was 0.971 (95% CI = 0.960-0.979), indicating very good internal consistency for all tasks of the MFT.

Reliability

As for inter-rater reliability, ICC (2,1) values for total scores and scores for each dimension ranged from 0.984 to 0.998, showing excellent inter-rater reliability. The SEM values for each dimension ranged from 0.08-0.18, and those for total MFT and MFS scores were 0.47 and 1.48, respectively. SEM% values ranged from 3.17% for the forward elevation task to 10.57% for the carry a cube task. SRD values for each dimension were all less than 1; values of total MFT and MFS scores were 1.31 and 4.10, respectively. SRD% values ranged from 8.78% for the forward elevation task to 29.29% for the carry a cube task (Table 4).

The relative and absolute intra-rater reliabilities of the MFT for each therapist are shown in Table 5, from the smallest to the largest values for each dimension. All ICC (1,1) values for total scores and scores for each dimension were above 0.98. SEM and SRD values for each dimension were all less than one.

In intra-rater study, two experienced therapists in administering MFT showed perfect test-retest reliabilities, such that their SEM and SRD values were calculated as zero. Meanwhile, four other therapists, who were novice MFT administrators, showed lower reliabilities with less than 10% of SEM% values for total MFT scores, which were quite good.

Validity

Responsiveness: To determine the responsiveness of the MFT, we calculated ES and SRM for two different groups divided according to the time elapsed since stroke onset (Table 6). Even though all values were less than 0.5, ES and SRM values for the acute-subacute group were larger than those for the chronic group. Paired *t* test revealed significant improvement in MFT scores in the acute-subacute group (t [42] = 1.33, p = 0.011), but not in the chronic group (t [37] = 3.25, p = 0.52). Further analysis was conducted for a subgroup of patients from the acutesubacute group: We extracted data from patients (n = 13) who underwent baseline assessment earlier than one month after stroke onset. Therein, ES was 0.34 and SRM was 0.62, indicating a greater mean change between assessments in acute-onset patients than in chronic patients.

Construct validity

Correlations between MFT scores and comparable measures are shown in Table 7. Overall, strong correlations were found between MFT and total scores for FMA ($|\mathbf{r}_s| \ge 0.861$, p < 0.001) and MMT ($|\mathbf{r}_s| \ge 0.790$, p < 0.001). Correlation coefficients between arm motion items in MFT and the FMA arm subscale were higher than those for other MFT items. Grasping and pinching items in MFT showed strong correlations with the FMA hand subscale ($\mathbf{r}_s = 0.90$, p < 0.001).

Discussion

Normal study

Analysis of our normative data revealed that age sig-

Table 4. Scores on the Manual Function Test and values of inter-rater reliability test assessed by eight raters (n = 81).

Dimension	Mean score	ICC _(2,1)	95%CI	SEM	SEM%	SRD	SRD%
Forward elevation	2.6 ± 1.5	.997	.995998	0.08	3.17	0.23	8.78
Lateral elevation	2.4 ± 1.4	.987	.983991	0.16	6.72	0.46	18.62
Occiput with the palm	2.6 ± 1.4	.991	.987993	0.14	5.30	0.38	14.68
Dorsum with the palm	2.2 ± 1.4	.984	.979989	0.18	8.02	0.50	22.24
Grasp	1.7 ± 1.5	.996	.994997	0.09	5.37	0.26	14.89
Pinch	1.2 ± 1.3	.992	.989994	0.12	9.55	0.33	26.47
Carry a cube	1.2 ± 1.3	.991	.987993	0.13	10.57	0.35	29.29
Peg-board	1.3 ± 1.6	.995	.993996	0.12	8.90	0.32	24.66
Total -MFT	15.3 ± 10.6	.998	.998999	0.47	3.10	1.31	8.58

Values are mean \pm SD for mean score.

ICC_(2,1), intraclass correlation coefficient for inter-rater reliability; 95% CI, 95% confidence interval; SEM, standard error of measurement; SRD, smallest real difference.

Dimension	ICC _(1,1)	SEM [*]	SEM%	SRD^*	SRD%
Forward elevation	.994-1.0	011	0-4.44	032	0-12.29
Lateral elevation	.982-1.0	019	0-7.97	054	0-22.08
Occiput with the palm	.992-1.0	013	0-5.07	037	0-14.04
Dorsum with the palm	.982-1.0	019	0-8.57	053	0-23.76
Grasp	.986-1.0	017	0-9.94	048	0-27.55
Pinch	.983-1.0	017	0-13.88	048	0-38.48
Carry a cube	.986-1.0	016	0-13.47	044	0-37.35
Peg-board	.984-1.0	021	0-15.92	058	0-44.13
Total MFT	.998-1.0	047	0-3.08	0-1.31	0-8.53

Table 5. Ranges of values of intra-rater reliability test assessed by eight raters (n = 81).

Values are range of minimum and maximum of each statistic values.

ICC_(1,1), intraclass correlation coefficient for intra-rater reliability; SEM, standard error of measurement; SRD, smallest real difference.

*Lower scores correspond to better statistics.

stroke (n = 81	.).				
Subgroup	Mean change	SD	ES	SRM	<i>p</i> -value
Acute-subacute $(n = 43)$	1.33	3.25	0.12	0.41	0.011
Chronic $(n = 38)$	0.08	0.75	0.01	0.11	0.520
Total group $(n = 81)$	0.74	2.49	0.07	0.30	0.009

Table 6. Responsiveness of Manual Function Test assessed by six weeks interval according to chronicity of stroke (n = 81).

SD, standard deviation of changes; ES, effect size; SRM, standardized response mean.

p values were calculated using the Paired t-test.

Table 7. Associations between performance on Manual Function Test and other tests of upper extremity.

	FMA	FMA-arm	FMA-hand	MMT
Forward elevation	.890*	.906*	.837*	.814*
Lateral elevation	.878*	.887*	.782*	.798*
Occiput with the palm	.909*	.909*	.875*	.826*
Back with the palm	.887*	.868*	.861*	.790*
Grasp	.861*	.827*	.900*	.825*
Pinch	.888*	.866*	.900*	.830*
Carry a cube	.886*	.853*	.846*	.843*
Peg-board	.873*	.855*	.850*	.838*
MFS	.960*	.947*	.889*	.877*

The values are Spearman rho scores (*p < 0.001).

FMA, the total score of Fugl-Meyer Assessment-Upper Extremity; FMA-arm, the total score of arm subscale of Fugyl-Meyer Assessment-Upper Extremity; FMA-hand, the total score of hand subscale of Fugyl-Meyer Assessment-Upper Extremity; MMT, total score of upper extremity part of the Manual muscle test.

nificantly predicts performance on the MFT. In regression analysis, age explained a greater amount of variance in scores for the subject's non-dominant side than scores for the dominant side. Also, similar to a previous study (Michimata et al. 2008), lower scores on the cube-carrying and peg-board tasks were associated with increasing age. Accordingly, we cautiously suggest that poor performance on MFT tasks evaluating motor functions might suggest problems in upper extremity function.

As in Rotterdam study, age is regarded as the most influential factor which has impact on dexterity (Hoogendam et al. 2014). However, there are conflicting results on factors influencing the determination of manual asymmetries (Weller and Latimer-Sayer 1985; Przybyla et al. 2011). It is considered that manual asymmetries or gender differences in measures of dexterity are mostly taskspecific (Francis and Spirduso 2000; Francis et al. 2015; Sivagnanasunderam et al. 2015).

Reliability

In this study, we found the intra-rater and inter-rater relative reliabilities, in terms of ICC, to be excellent (ICC > .90) for total scores on the MFT as well as each of its eight individual tasks. Our reliability coefficients were comparable to those reported in a previous study on patients with stroke in Japan, where MFT was first developed (Miyamoto et al. 2009). However, high ICCs do not necessarily imply that the test is suitable for clinical use. For more meaning-

ful information, assessment tools should display small measurement errors that are sensitive enough to detect real changes in both individual subjects and groups thereof. Thus, we assessed absolute reliability by calculating SEM, SEM%, SRD, and SRD%. To the best of our knowledge, this study was the first to investigate the absolute reliability of the MFT.

In our study, the inter-rater and intra-rater SRD% values for total MFT scores were satisfactory low. However, scores for individual subtests are reported higher. Notwithstanding, as percentage values of SEM and SRD are influenced by the mean score of each item in the test, since mean values serve as the denominator for percentage calculations, measurement noise should be taken into consideration when interpreting reliability results. The benchmark score of a functional hand is 22 or above for total MFT score (Sone et al. 2015), and in the present study, more than 50% of the patients had a non-functional hemiplegic hand (total MFT score less than 22). Thus, mean scores for tasks with manipulative activities were low, and as a consequence thereof, percentage values of SEM and SRD increased. Accordingly, we would not suggest that our results invalidate MFT assessment.

Herein, SRD values for scores in each dimension were all below 1 point, and that for total MFT score was below 2 points. These findings indicate that a change in sub-dimension scores by 1 point and a change in total MFT scores by 2 points, as assessed by different raters for an individual patient after intervention, imply a real change with 95% certainty. As the results of our inter-rater reliability study reflect computation data for MTF assessments conducted by eight therapists, we deemed the SRDs small enough to be significant.

Validity

Validity in terms of responsiveness: While paired t test indicated that there was a significant improvement in scores evaluated six weeks apart for the acute-subacute group, the overall results only showed modest improvement. The low values of ES and SRM may have arisen from the fact that the state of our patients was somewhat severe and that the 6-week interval between the baseline and follow-up measurements was not long enough to detect significant changes. Indeed, Chen et al. (2014) reported that ES and SRM values for FMA increased with longer time intervals between assessments.

Construct validity: In this study, the construct validity of the MFT was estimated by calculating Spearman correlation coefficients for associations between MFT and FMA and MMT scores. The validity of the MFT showed good correlations with both measures. Notably, closer associations between MFT and FMA were identified for subdomains of similar movements (e.g., arm motion on MFT and the arm subscale on FMA; grasp and pinch items on MFT and the hand subscale on FMA). As FMA is a widely used functional scale of the upper extremities in patients with stroke and as its validity and reliability are well-established, these findings confirm the validity of using the MFT to assess upper extremity function in patients with stroke. Importantly, this study is the first to report validation results for the MFT in relation to FMA.

Study Limitations

In this study, we primarily assessed the effect of the examiner on reliability. The reliability of an assessment tool can also be influenced by day-to-day variability, the test procedure itself, and the environmental conditions in which the assessment was performed (Sole et al. 2007). Hence, further study of the reliability of MFT in relation to these factors may be required. In addition, previous studies have indicated that SEM and SRD values can differ depending on whether the tasks are conducted by either the paretic or non-paretic side of patients with hemiplegia (Patten et al. 2003; Flansbjer et al. 2005a). Thus, motor impairment would likely affect the absolute reliability values of tests such as the MFT. Therefore, a larger sample of patients with stroke differing in levels of motor impairment is required to study whether the absolute reliability of the MFT differs according to impairment severity. Also, as seen in this study, a rater's familiarity with and proficiency in a specific tool affects performance. In this respect, further reliability study including raters of similar proficiency in MFT may be warranted. Lastly, although fine motor skill is mostly affected by age in non-demented elders (Hoogendam et al. 2014), cognitive deficit may influence dexterity (de Paula et al. 2016). Since we did not evaluate cognitive function in our healthy participants, any influence by cognitive deficits may not be excluded especially in our older participants.

Conclusion

Our results confirmed that the Korean version of MFT is a highly reliable and valid instrument for the assessment of upper extremity function in patients with hemiplegic stroke.

Acknowledgments

Manual function test kit was kindly loaned by Yerang Medical Co. to Nowon Eulji Medical Center.

Conflict of Interest

The authors declare no conflict of interest.

References

- Alt Murphy, M., Resteghini, C., Feys, P. & Lamers, I. (2015) An overview of systematic reviews on upper extremity outcome measures after stroke. *BMC Neurol.*, 15, 29.
- Atkinson, G. & Nevill, A.M. (1998) Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med.*, 26, 217-238.
- Beaton, D.E., Bombardier, C., Guillemin, F. & Ferraz, M.B. (2000) Guidelines for the process of cross-cultural adaptation of selfreport measures. *Spine (Phila Pa 1976)*, 25, 3186-3191.
- Beckerman, H., Roebroeck, M.E., Lankhorst, G.J., Becher, J.G., Bezemer, P.D. & Verbeek, A.L. (2001) Smallest real differ-

ence, a link between reproducibility and responsiveness. *Qual. Life Res.*, **10**, 571-578.

- Chen, K.L., Chen, C.T., Chou, Y.T., Shih, C.L., Koh, C.L. & Hsieh, C.L. (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**, 941-949.
- Cohen, J. (1988) Statistical power analysis for the behavior sciences, 2nd ed., Lawrence Erlbaum Associates, Hillsdalem, NJ.
- da Silva, P.B., Antunes, F.N., Graef, P., Cechetti, F. & Pagnussat Ade, S. (2015) Strength training associated with task-oriented training to enhance upper-limb motor function in elderly patients with mild impairment after stroke: a randomized controlled trial. Am. J. Phys. Med. Rehabil., 94, 11-19.
- de Paula, J.J., Albuquerque, M.R., Lage, G.M., Bicalho, M.A., Romano-Silva, M.A. & Malloy-Diniz, L.F. (2016) Impairment of fine motor dexterity in mild cognitive impairment and Alzheimer's disease dementia: association with activities of daily living. *Rev. Bras. Psiquiatr.*, 38, 235-238.
- Flansbjer, U.B., Holmback, A.M., Downham, D. & Lexell, J. (2005a) What change in isokinetic knee muscle strength can be detected in men and women with hemiparesis after stroke? *Clin. Rehabil.*, **19**, 514-522.
- Flansbjer, U.B., Holmback, A.M., Downham, D., Patten, C. & Lexell, J. (2005b) Reliability of gait performance tests in men and women with hemiparesis after stroke. *J. Rehabil. Med.*, 37, 75-82.
- Francis, K.L. & Spirduso, W.W. (2000) Age differences in the expression of manual asymmetry. *Exp. Aging Res.*, 26, 169-180.
- Francis, K.L., MacRae, P.G., Spirduso, W.W. & Eakin, T. (2015) Age and practice effects on inter-manual performance asymmetry. *Front. Psychol.*, 5, 1585.
- Fugl-Meyer, A.R., Jaasko, L., Leyman, I., Olsson, S. & Steglind, S. (1975) The post-stroke hemiplegic patient. 1. a method for evaluation of physical performance. *Scand. J. Rehabil. Med.*, 7, 13-31.
- Gladstone, D.J., Danells, C.J. & Black, S.E. (2002) The Fugl-Meyer assessment of motor recovery after stroke: a critical review of its measurement properties. *Neurorehabil. Neural Repair*, 16, 232-240.
- Guyatt, G., Walter, S. & Norman, G. (1987) Measuring change over time: assessing the usefulness of evaluative instruments. *J. Chronic Dis.*, 40, 171-178.
- Hoogendam, Y.Y., Hofman, A., van der Geest, J.N., van der Lugt, A. & Ikram, M.A. (2014) Patterns of cognitive function in aging: the Rotterdam Study. *Eur. J. Epidemiol.*, **29**, 133-140.
- Hsieh, Y.W., Hsueh, I.P., Chou, Y.T., Sheu, C.F., Hsieh, C.L. & Kwakkel, G. (2007) Development and validation of a short form of the Fugl-Meyer motor scale in patients with stroke. *Stroke*, **38**, 3052-3054.
- Husted, J.A., Cook, R.J., Farewell, V.T. & Gladman, D.D. (2000) Methods for assessing responsiveness: a critical review and recommendations. J. Clin. Epidemiol., 53, 459-468.
- Kazis, L.E., Anderson, J.J. & Meenan, R.F. (1989) Effect sizes for interpreting changes in health status. *Med. Care*, 27, S178-189.
- Ko, J. & Kim, M. (2013) Reliability and responsiveness of the gross motor function measure-88 in children with cerebral palsy. *Phys. Therapy*, **93**, 393-400.
- Kreisel, S.H., Hennerici, M.G. & Bazner, H. (2007) Pathophysiology of stroke rehabilitation: the natural course of clinical recovery, use-dependent plasticity and rehabilitative outcome. *Cerebrovasc. Dis.*, 23, 243-255.
- Kwakkel, G., Kollen, B.J., van der Grond, J. & Prevo, A.J. (2003) Probability of regaining dexterity in the flaccid upper limb: impact of severity of paresis and time since onset in acute stroke. *Stroke*, **34**, 2181-2186.

Liang, M.H., Fossel, A.H. & Larson, M.G. (1990) Comparisons of

five health status instruments for orthopedic evaluation. *Med. Care*, **28**, 632-642.

- Liaw, L.J., Hsieh, C.L., Lo, S.K., Chen, H.M., Lee, S. & Lin, J.H. (2008) The relative and absolute reliability of two balance performance measures in chronic stroke patients. *Disabil. Rehabil.*, **30**, 656-661.
- Medical Research Council (Great Britain) & University of Edinburgh Department of Surgery. (1976) *Aids to the examination of the peripheral nervous system*, edited by H.M. Stationery Off., London.
- Michimata, A., Kondo, T., Suzukamo, Y., Chiba, M. & Izumi, S. (2008) The manual function test: norms for 20- to 90-yearolds and effects of age, gender, and hand dominance on dexterity. *Tohoku J. Exp. Med.*, 214, 257-267.
- Min, K., Song, J., Kang, J.Y., Ko, J., Ryu, J.S., Kang, M.S., Jang, S.J., Kim, S.H., Oh, D., Kim, M.K., Kim, S.S. & Kim, M. (2013) Umbilical cord blood therapy potentiated with erythropoietin for children with cerebral palsy: a double-blind, randomized, placebo-controlled trial. *Stem Cells*, **31**, 581-591.
- Miyamoto, S., Kondo, T., Suzukamo, Y., Michimata, A. & Izumi, S. (2009) Reliability and validity of the Manual Function Test in patients with stroke. *Am. J. Phys. Med. Rehabil.*, 88, 247-255.
- Nakamura, R. & Moriyama, S. (2000) Rehabilitation manual 8: Manual function test(MFT) and functional occupational therapy for stroke patients, National Rehabilitation Center for the Disabled, Tokorozawa.
- Nunnally, J.C. (1978) *Psychometric theory*, 2nd ed., McGraw-Hill, New York, NY.
- Patten, C., Kothari, D., Whitney, J., Lexell, J. & Lum, P.S. (2003) Reliability and responsiveness of elbow trajectory tracking in chronic poststroke hemiparesis. *J. Rehabil. Res. Dev.*, 40, 487-500.
- Przybyla, A., Haaland, K.Y., Bagesteiro, L.B. & Sainburg, R.L. (2011) Motor asymmetry reduction in older adults. *Neurosci. Lett.*, 489, 99-104.
- Sale, P., Franceschini, M., Mazzoleni, S., Palma, E., Agosti, M. & Posteraro, F. (2014) Effects of upper limb robot-assisted therapy on motor recovery in subacute stroke patients. J. Neuroeng. Rehabil., 11, 104.
- Sivagnanasunderam, M., Gonzalez, D.A., Bryden, P.J., Young, G., Forsyth, A. & Roy, E.A. (2015) Handedness throughout the lifespan: cross-sectional view on sex differences as asymmetries change. *Front. Psychol.*, 5, 1556.
- Sivan, M., O'Connor, R.J., Makower, S., Levesley, M. & Bhakta, B. (2011) Systematic review of outcome measures used in the evaluation of robot-assisted upper limb exercise in stroke. J. Rehabil. Med., 43, 181-189.
- Sole, G., Hamren, J., Milosavljevic, S., Nicholson, H. & Sullivan, S.J. (2007) Test-retest reliability of isokinetic knee extension and flexion. *Arch. Phys. Med. Rehabil.*, 88, 626-631.
- Sone, T., Nakaya, N., Iokawa, K., Hasegawa, K., Tsukada, T., Kaneda, M., Hamaguchi, T. & Suzuki, K. (2015) Prediction of upper limb recovery in the acute phase of cerebrovascular disease: evaluation of "functional hand" using the manual function test. J. Stroke Cerebrovasc. Dis., 24, 815-822.
- Sullivan, K.J., Tilson, J.K., Cen, S.Y., Rose, D.K., Hershberg, J., Correa, A., Gallichio, J., McLeod, M., Moore, C., Wu, S.S. & Duncan, P.W. (2011) Fugl-Meyer assessment of sensorimotor function after stroke: standardized training procedure for clinical practice and clinical trials. *Stroke*, **42**, 427-432.
- Weller, M.P. & Latimer-Sayer, D.T. (1985) Increasing right hand dominance with age on a motor skill task. *Psychol. Med.*, 15, 867-872.
- Wolf, S.L., Sahu, K., Bay, R.C., Buchanan, S., Reiss, A., Linder, S., Rosenfeldt, A. & Alberts, J. (2015) The HAAPI (Home Arm Assistance Progression Initiative) Trial: a novel robotics delivery approach in stroke rehabilitation. *Neurorehabil. Neural Repair*, 29, 958-968.