J. Phys. Ther. Sci.
24: 893–899, 2012

Reliability, Concurrent Validity, and Responsiveness of the Fugl-Meyer Assessment (FMA) for Hemiplegic Patients

Heesoo Kim, PhD student, OT[1)], Jingang Her, PhD, PT[1)], Jooyeon Ko, PhD, PT[2)], Dong-sik Park, PhD, MD[3)], Ji-Hea Woo, PhD student, PT[1)], Youngyoul You, PhD student, PT[1)], Youngeun Choi, PhD student, PT[1)]

[1)] *Department of Rehabilitation Therapy, Graduate school of Hallym University : 1 Hallymdaehak-gil, Chuncheon,Gangwon-do, 200-702 Korea. TEL: +82 10-8712-5791, E-mail: jghur7@empas.com*
[2)] *Department of Rehabilitation Medicine, CHA Bundang Medical Center*
[3)] *Department of Rehabilitation Medicine, Hallym University College of Medicine*

Abstract. [Purpose] The purpose of this study was to determine the reliability, validity, and responsiveness of the Fugl-Meyer Assessment (FMA) for hemiplegic patients. [Subjects] For the reliability and validity study, 50 patients with stroke (26 males, 24 females) were recruited. For the responsiveness study, 16 hemiplegic patients (8 males, 8 females) participated. [Methods] Two physical therapists and one occupational therapist rated 50 video recordings of hemiplegic patients using the FMA to test the inter-rater reliability, and one physical therapist (rater A) rated each of the 50 video clips on two occasions, two weeks apart, to evaluate the test-retest reliability. Responsiveness was calculated three months after the baseline assessment. Reliability was calculated using the intraclass correlation coefficient (ICC), standard error of measurement (SEM), and smallest real difference (SRD). Concurrent validity was examined using Pearson's correlation coefficient and responsiveness was calculated using the effect size (ES) and standardized response mean (SRM). [Results] Assessment using the FMA showed high relative reliability, and the absolute reliability was satisfactory for the inter-rater and test-retest reliabilities. The correlations between motor function of the FMA and the Jebsen-Taylor hand function, grip power, motor assessment scale (MAS), and the Berg balance scale (BBS) were moderate to good, and were highly significant ($p<0.05$), while responsiveness was moderate to large. [Conclusion] The results indicate that the FMA is a reasonable assessment of the function of the upper and lower extremities of patient with stroke.
Key words: Reliability, Validity, Responsiveness

(This article was submitted Apr. 3, 2012, and was accepted May 8, 2012)

INTRODUCTION

Hemiplegic patients suffer from sensory-motor deficit, decreased passive joint motion, and joint pain[1)]. These impairments have a profound effect on patients' activities of daily living and on rehabilitation outcomes. Therefore, the function of patients with hemiplegia caused by stroke needs to be assessed in order to identify the degree of functional damages, and intervention plans need to be accompanied by evaluations of their outcomes[2)].

Outcome evaluation tools need to meet a diversity of psychometric criteria and should have high reliability. Reliability refers to the consistency of a test or measurement, and inter-rater, and test-retest reliabilities are necessary for clinical outcome measures. Inter-rater reliability measures the degree of consistency of the results when two or more raters test the same subject using the same assessment method at the same time[3)]. Test-rest reliability is the degree of consistency in the results when a single person measures the same item twice[4)]. Reliability may also be divided into relative and absolute reliabilities[5)]. Relative reliability is determined using the intraclass correlation coefficient (ICC)[6)], and absolute reliability is determined with the standard measurement error (SEM) and smallest real difference (SRD)[7)]. Some outcome evaluation tools have a high validity and the interrelationship between widely used valid evaluation tools and FMA, in other words, their concurrent validity, needs to be investigated[8)]. Responsiveness is the most important psychometric property of outcome evaluation tools, and it also needs to be estimated. Responsiveness refers to the ability of an instrument to detect a clinically relevant change, and the effect size (ES) and standard response mean (SRM) are usually used to quantify the change[9)].

Various tools are available to evaluate the functional recovery of hemiplegic patients after treatment or after a certain amount of time has lapsed. The Jebsen-Taylor hand function test, the grip power test[10, 11)], the Action Reach Arm Test (ARAT)[12)], the Berg Balance Scale (BBS)[13)], the Motor Assessment Scale (MAS)[14)], and the FMA[1)] which include domains such as motor function and balance, sensation qualities, passive range of motion, and joint pain are the

most widely used[15–17]. Research on the reliability and validity of the FMA has been conducted, but the majority of the research has only examined reliability and validity in one or two domains, not all the domains, or either of the upper or lower extremity[2, 18, 19]. Michaelsen et al. reported that the FMA has a reasonable reliability for total motor scores of the upper extremity (ICC=0.98), lower extremity (ICC = 0.90), movement sense (ICC = 0.98), upper and lower extremities' passive range of motion (ICC=0.84 and 0.90, respectively), and tactile sensitivity (ICC =0.75) however, they had only evaluated the inter-rater reliability[20].

As mentioned above, a reliable evaluation tool should have a high responsiveness as well as high reliability and validity. In particular, test-retest reliability is closely associated with responsiveness, therefore evaluating test-retest reliability and responsiveness together is desirable[4]. As far as we know, most research on the concurrent validity and responsiveness of the FMA has looked at the upper extremity functions. Motion function was highly correlated with the action research arm test (ARAT) and the box and block test (r= 0.925 and 0.921, respectively)[18], and van der Lee et al. reported that intensive treatment for 22 chronic stroke patients' upper extremity functions gave a responsiveness ratio of 0.41[19].

The aim of present study was to evaluate the inter-rater and test-retest reliabilities, concurrent validity, and responsiveness of the FMA in all the domains of motor function, sensation, passive range of motion, and joint pain of the upper and lower extremities. We also aimed to provide useful outcome evaluation tools for both clinical and research purposes.

SUBJECTS AND METHODS

The Fugl-Meyer Assessment was first presented in 1975 as 'a method for evaluation of physical performance following stroke[1]. The FMA version used in the present study, was a 212-point multi-item scale, which was divided into 4 domains: motor function, sensation qualities, passive range of motion, and joint pain. Each domain contains multiple items, each scored on a 3-point ordinal scale (0 = cannot perform, 1= perform partially, 2 = perform fully). The motor domain includes items that evaluate movement, coordination, and reflex activity of the shoulder, elbow, wrist, hand, hip, knee, and ankle. The motor score ranges from 0 to 100 points (66 points for the upper extremity and 34 points for the lower extremity). The sensation domain includes light touch and position, and the score range is 0 to 24 points (12 points each for the upper and lower extremities). The passive range of motion and joint pain scores range from 0 to 44 points (24 points for the upper extremity and 20 points for the lower extremity).

The subjects were a convenient sample of 50 inpatients (26 males and 24 females) with stroke from three rehabilitation hospitals in South Korea who we treated from May to October 2011. Their ages ranged from 47 to 72 years old. The inclusion criteria were: hemiplegia resulting from a cerebrovascular accident (CVA), not from trauma, brain tumor, surgery, or any other etiology; the date of onset of hemiplegia was at least six months before the date of assessment; the patient had been admitted for one-month of intensive rehabilitation therapy (two sessions of physical therapy and two sessions of occupational therapy per day, for 5 days a week); and the ability to understand instructions and to be oriented to name, time, and place as recommended by the original author[1]. All participants signed a consent form approved by the Hallym University Institutional Review Board (HIRB) which included consent to being videotaped.

Initially, a balance domain was added to motion function of the FMA to give a total of five domains with a total score of 226. However, in the present study, we used the original author's version of the FMA, which consists of four domains with a total score of 212.

For this reason, in the present study, we translated the English version of the FMA into Korean with the permission of the original author. The translation procedures followed the forward-backward-forward method by two of the authors, and we discussed any minor problems of the translation with the original author and corrected them.

Prior to the actual evaluation, three raters (physical therapists, A and B, and an occupational therapist C), and one assistant participated in a three hour workshop on two occasions. The workshop was run by a senior therapist with 20 or more years of clinical experience. After the workshop ended, a pre-test trial was performed twice with five patients as subjects. Then, all tests were carried out by the senior therapist and the procedures were video-recorded by the assistant. The three raters then started scoring the FMA performance of the patients using the video-recordings. The three raters' experience of the treatment of stroke patients averaged 7.2 years (from three to twelve years). Evaluation was made in a treatment room with which the subjects were familiar in order to minimize measurement errors, and efforts were made to equalize the time of each evaluation. The patients wore light clothes and had bare feet, and the senior therapist and the patients' guardians took part.

Inter-rater reliability was measured by the three raters' watching the first video-recordings of the 50 subjects. Each scoring session was held over four days without discussion of the scores among the raters. To examine the test-retest reliability, all participants were scheduled to participate in two separate test sessions separated by an interval of two weeks. Rater A, who is a physical therapist with more than 7 years experience, scored 50 video recordings in the two scoring sessions, which were two-weeks apart. This assessment interval is generally believed to be a reasonable compromise between recollection bias and unwanted clinical change[21]. Fortunately, no patients dropped out before the second test.

To determine the concurrent validity, rater A performed the Jebsen-Taylor hand function test, the grip power test using a Jamar dynamometer, and scored the MAS and the BBS the day after making the first FMA evaluation of the video recordings. The Jebsen-Taylor hand function test is composed of seven standardized sub-tests of the most frequently conducted hand function activities such as writing short sentences and turning up cards. The time taken to perform the tasks was measured using a stopwatch. Grip

power was measured with the Jamar dynamometer using hydraulic pressure, with the subject seated, the shoulder joints in abduction and neutral rotation, the elbow joints flexed at 90 degrees, and the wrists in a neutral upward position. Measurements were taken on both hands three times and the average values were used in the analysis[10]. The MAS consists of nine items (supine to side lying, supine to sitting to standing, etc.). Eight items evaluate the functions of different body parts and one item assesses muscle tension on the affected side. A score from zero to six points is assigned to each item, which was measured three times. The best score was used[14]. The BBS is composed of 14 items assessing hemiplegic patients' balance abilities which are scored from zero to four points giving a maximum score of 56 points, with a higher score suggesting a better balance function[13].

To calculate the responsiveness for the same subjects, one additional follow-up assessment was conducted at 3 months after the baseline assessment. Among the fifty original participants, data were available for 16 who continued physical therapy after the baseline assessment.

Other information about general characteristics was obtained by reviewing medical records. The total score of each tool was analyzed[20] and descriptive statistics were calculated. For relative reliability, the $ICC_{3,1}$ with 95% confidence intervals (CI) was used to evaluate the inter-rater and test-retest reliabilities of each domain and total score of the FMA. According to Polit and Hungler, a reliability coefficient of at least 0.70 is sufficient for group comparisons[22]. For absolute reliability, the SEM and the SRD were calculated. The SEM represents the SD of measurement errors and is calculated as follows: SEM = SD $\sqrt{}$ 1-ICC[23]. A lower score means a smaller measurement error. The SRD value of less than $\sqrt{2} \times 1.96 \times SEM = 2.77 \times SEM$ is expected to admit the results of 95% of the ratings[24]. Pearson's correlation coefficient was determined to examine the inter-relationships among the FMA, the Jebsen-Taylor hand function test, the grip power test, the MAS, and the BBS. According to Meyer, a correlation coefficient $r \geq 0.8$ indicates a high correlation, $r = 0.6–0.8$ indicates a good correlation, $r = 0.4–0.6$ indicates a moderate correlation, and $r \geq 0.4$ indicates a poor correlation[25]. For responsiveness, the ES was calculated by dividing the mean change with the SD of the baseline. The SRM was calculated by dividing the mean change with SD of the changes. According to Cohen's criteria, an ES greater than 0.8 is large, 0.5 to 0.8 is moderate, and 0.2 to 0.5 is small[26].

RESULTS

Characteristics of the subjects and descriptive statistics are presented in Tables 1 and 2. For inter-rater reliability, the ICCs ranged from 0.930 to 1.000 and the SEMs were 0.00 to 2.15; the SRDs ranged from 0.39 to 5.96 for the motor function, sensation, passive joint motion, and joint pain of the FMA (Table 3). For test-retest reliability with an interval of two weeks, the ICCs were 0.834 to 0.972. The SEMs ranged from 0.85 to 3.25 and the SRDs from 2.35 to 9.00 (Table 3).

For concurrent validity, the relationship between the upper extremity motor function of the FMA and Jebsen-Taylor hand function was 0.757 ($p<0.05$), and the relationship between the lower extremity motor function of the FMA and the MAS was 0.725 ($p<0.05$) (Table 4).

The responsiveness of the FMA at three months after the baseline assessment is shown in Table 5. The ES and the SRM of the upper extremity motor function were 0.69 and 1.00, and 0.64, and 0.73 for the motor function of the lower extremity.

DISCUSSION

Our results indicate a satisfactory level of responsiveness as well as relative and absolute reliabilities of the FMA, based on calculations of the ICC, SEM, SRD, ES, and SRM. The upper and lower motor functions of the FMA demonstrated good concurrent validity with other outcome measures.

For relative reliability, the ICC of the inter-rater reliability was more than 0.930 for all domains, whereas in previous research, the ICC of sensation and pain domains were 0.85 and 0.61, respectively[27]. Diverse factors such as study design, subject groups, and treatment intervention methods affect the reliability evaluation. The factors related to study design include measurement time interval, evaluation method, and the number of people being measured. Factors associated with study subject groups are the diagnosed disease, age, gender, emotional conditions, and cognitive levels[28]. All of these factors can greatly influence not only error dispersion but also deviation among subjects, and accordingly the present study considered measurement intervals. In addition, all of the raters received training in the use of the assessment tools, discussed their implementation, and conducted pretests.

The ICC of the test-retest reliability was higher than 0.9 for the motor function total, upper extremity motor function, and lower extremity sensation, 0.962, 0.972, and 0.921 respectively. For the other domains, the ICC ranged from 0.828 to 0.883. As far as we know, there has been almost no previous research on the test-retest reliability of the FMA. The establishment of trustworthy test-retest reliability is very important for the clinical application of outcome evaluation tools in a rehabilitation environment. This is because patients' conditions and the effects of treatment interventions with lapse of time need to be quantified through comprehensive evaluation[29]. However, research on reliability mostly employs ICC, which does not provide information about measurement errors. Therefore, the SEM should also be calculated, as a high ICC does not necessarily mean a small measurement error[30]. For instance, the ICCs for motor function and joint pain of the upper extremity were 0.972 and 0.830, respectively, and the SEMs were 3.25 and 2.23. The ICC indicates the motor function of upper extremity is more reliable, but the SEM indicates that joint pain of the upper extremity is more reliable. The SEM and the SRD represent the reliability better than the ICC[31].

For absolute reliability, we calculated the SRD for each of the FMA domain and the FMA total score. This provides

Table 1. Summary of sample characteristics

Variables	Reliability & validity (N=50)	Responsiveness (n=16)
Gender		
Male	26 (52.0)	8 (50.0)
Female	24 (48.0)	8 (50.0)
Type of stroke		
Hemorrhage	29 (58.0)	10 (62.5)
Infarction	21 (42.0)	6 (37.5)
Affected side		
Right	22 (44.0)	3 (18.8)
Left	28 (56.0)	13 (81.3)
Age (years)	59.5 ± 12.4	55.1 ± 12.7
Stroke onset time (years)	2.4 ± 1.8	1.6 ± 0.5
MMSE	24.3 ± 3.4	24.5 ± 3.6
Jebsen-Taylor Hand function	20.1 ± 2.2	21.2 ± 23.3
Grip power	7.3 ± 9.8	6.0 ± 7.3
MAS total	36.2 ± 12.5	37.5 ± 12.3
BBS total	50.5 ± 13.8	56.3 ± 12.4

Values are n (%) or Mean (SD), MMSE: Mini mental state examination, MAS: Motor assessment scale, BBS: Berg balance scale

Table 2. Descriptive statistics of the first and second Fugl-Meyer Assessment (FMA) of hemiplegic patients

	Rater A		Rater B		Rater C	
	Mean ± SD	Min-Max	Mean ± SD	Min-Max	Mean ± SD	Min-Max
1st Motor function total	60.2 ± 25.3	4–99	60.1 ± 25.6	4–99	59.8 ± 25.4	4–99
1st U/E Motor function	38.9 ± 19.2	4–66	38.8 ± 19.2	4–66	38.8 ± 19.3	4–66
1st L/E Motor function	21.3 ± 8.0	0–34	21.3 ± 8.4	0–34	20.9 ± 8.2	0–34
1st sensation total	18.1 ± 6.4	0–24	18.1 ± 6.5	0–24	18.1 ± 6.3	0–24
1st U/E Sensation	8.4 ± 3.6	0–12	8.5 ± 3.5	0–12	8.4 ± 3.6	0–12
1st L/E Sensation	9.7 ± 3.4	0–12	9.6 ± 3.5	0–12	9.7 ± 3.4	0–12
1st Passive joint motion total function	40.6 ± 4.9	24–47	40.6 ± 4.7	26–44	40.6 ± 4.6	26–44
1st U/E Passive joint motion	21.8 ± 3.3	10–24	21.8 ± 3.2	10–24	21.8 ± 3.3	10–24
1st L/E Passive joint motion	18.8 ± 2.4	11–24	18.7 ± 2.2	11–20	18.8 ± 2.1	12–20
1st Joint pain total	39.0 ± 8.4	0–44	39.1 ± 8.4	0–44	39.1 ± 7.9	0–44
1st U/E Joint pain	20.7 ± 5.1	0–24	20.8 ± 5.1	0–24	20.6 ± 5.0	0–24
1st L/E Joint pain	18.3 ± 4.3	0–20	18.3 ± 4.3	0–20	18.5 ± 3.7	0–20
2nd Motor function total	64.4 ± 26.4	4–100	63.8 ± 26.0	4–100	63.4 ± 27.0	4–100
2nd U/E Motor function	41.4 ±19.8	4–66	41.5 ± 19.8	4–66	40.4 ± 20.8	1–66
2nd L/E Motor function	23.0 ± 7.9	0–34	22.3 ± 8.4	0–34	23.0 ±7.8	0–34
2nd Sensation total	19.0 ± 6.7	0–24	19.4 ± 7.5	0–40	18.8 ± 6.9	0–24
2nd U/E Sensation	9.0 ± 3.8	0–12	9.1 ± 3.8	0–12	9.0 ± 3.8	0–12
2nd L/E Sensation	9.9 ± 3.3	0–12	10.3 ± 4.7	0–34	9.9 ± 3.3	0–12
2nd Passive joint motion total	40.9 ± 4.6	24–44	40.8 ± 4.6	24–44	40.8 ± 4.6	24–44
2nd U/E Passive joint motion	22.2 ± 2.7	14–24	22.1 ± 2.7	14–24	22.1 ± 2.7	14–24
2nd L/E Passive joint motion	18.7 ± 2.4	10–20	18.7 ± 2.4	10–20	18.7 ± 2.4	10–20
2nd Joint pain total	38.2 ± 10.4	0–44	38.2 ± 10.4	0–44	38.1 ± 10.7	0–44
2nd U/E Joint pain	20.8 ± 5.8	0–24	20.8 ± 5.8	0–24	20.7 ± 6.0	0–24
2nd L/E Joint pain	17.5 ± 5.1	0–20	17.5 ± 5.1	0–20	17.5 ± 5.1	0–20

(N=50) U/E: Upper extremity, L/E: Lower extremity

Table 3. Inter-rater and test-retest reliabilities of the Fugl-Meyer Assessment (FMA) for hemiplegic patients

	Inter-rater reliability				
	Mean ± SD	ICC	95% CI	SEM	SRD
Motor function total	60.0 ± 25.3	0.992	0.988 – 0.995	2.30	6.26
U/E Motor function	38.8 ± 19.1	1.000	1.000 – 1.000	0.00	–
L/E Motor function	21.2 ± 8.1	0.930	0.892 – 0.957	2.15	5.96
Sensation total	18.1 ± 6.4	0.994	0.991 – 0.996	0.49	1.36
U/E Sensation	8.4 ± 3.5	0.992	0.987 – 0.995	0.31	0.86
L/E Sensation	9.7 ± 3.4	0.988	0.981 – 0.993	0.37	1.02
Passive joint motion total	40.6 ± 4.7	0.992	0.987 – 0.995	0.42	1.16
U/E Passive joint motion	21.8 ± 3.2	0.998	0.997 – 0.999	0.14	0.39
L/E Passive joint motion	18.8 ± 2.2	0.966	0.947 – 0.980	0.41	1.14
Joint pain total	39.1 ± 8.2	0.991	0.995 – 0.998	0.78	2.16
U/E Joint pain	20.7 ± 5.0	0.992	0.988 – 0.995	0.45	1.25
L/E Joint pain	18.4 ± 4.1	0.974	0.959 – 0.984	0.65	1.80
	Test-retest reliability (Rater A)				
	Mean ± SD	ICC	95% CI	SEM	SRD
Motor function total	62.3 ± 21.8	0.961	0.917 – 0.980	4.31	11.94
U/E Motor function	40.2 ± 19.4	0.972	0.943 – 0.985	3.25	9.00
L/E Motor function	22.2 ± 7.9	0.868	0.762 – 0.926	2.88	7.98
Sensation total	18.5 ± 6.5	0.883	0.798 – 0.932	2.24	6.20
U/E Sensation	8.7 ± 3.7	0.834	0.836 – 0.944	1.50	4.16
L/E Sensation	9.8 ± 3.3	0.921	0.861 – 0.955	0.93	2.58
Passive joint motion total	40.7 ± 4.7	0.828	0.715 – 0.898	1.95	5.40
U/E Passive joint motion	22.0 ± 3.0	0.864	0.761 – 0.922	1.11	3.07
L/E Passive joint motion	18.8 ± 2.3	0.869	0.770 – 0.926	0.85	2.35
Joint pain total	38.6 ± 8.3	0.848	0.732 – 0.913	3.24	8.97
U/E Joint pain	20.7 ± 5.4	0.830	0.700 – 0.904	2.23	6.18
L/E Joint pain	17.9 ± 4.7	0.859	0.750 – 0.920	1.76	4.88

(N=50) ICC: Intraclass correlation coefficient, CI: Confidence interval, SEM: Standard error of measurement, SRD: Smallest real difference.

Table 4. Concurrent validities of the Fugl-Meyer Assessment (FMA) for hemiplegic patients

	FMA U/E motor	FMA L/E motor	BBS	MAS	Grip power	Jebsen-Taylor hand function
Jebsen-Taylor hand function	0.757*	0.466*	2.08	0.560*	0.734*	1
Grip power	0.719*	0.549*	0.241	0.543*	1	
MAS	0.692*	0.725*	0.661*	1		
BBS	0.470*	0.661*	1			
FMA L/E motor	0.723*	1				
FMA U/E motor	1					

(N=50) *p <0.05. L/E: Lower extremity, U/E: Upper extremity, MAS: Motor assessment scale, BBS: Berg balance scale.

a threshold for interpreting the scores of the FMA over time. For instance, the SRD threshold for a clinically significant change after a one-month intensive treatment is 9.00 in the upper extremities and 7.98 in the lower extremities. In other words, that a significant improvement has occurred with 95% reliability when the scores change in the upper and lower extremity functions are over 9.00 and 7.98, respectively.

As for concurrent validity, lower extremity motor functions of the FMA were closely associated with the MAS and BBS, and upper extremity motion functions of the FMA

Table 5. Responsiveness of the Fugl-Meyer Assessment (FMA) between baseline assessment and 3 months after baseline assessment for hemiplegic patients

	Mean change ± SD	ES	SRM
Motor function total	16.3 ± 14.4	0.80	1.13
U/E Motor function	11.2 ± 11.2	0.69	1.00
L/E Motor function	5.1 ± 7.0	0.64	0.73
Sensation total	3.1 ± 3.9	0.54	0.81
U/E Sensation	1.9 ± 2.7	0.58	0.72
L/E Sensation	1.2 ± 1.7	0.40	0.71
Passive joint motion total	2.3 ± 3.2	0.51	0.71
U/E Passive joint motion	1.8 ± 2.4	0.53	0.74
L/E Passive joint motion	0.5 ± 1.1	0.26	0.45
Joint pain total	3.7 ± 4.8	0.60	0.77
U/E Joint pain	2.7 ± 4.1	0.60	0.66
L/E Joint pain	1.0 ± 1.8	0.43	0.56

(n=16) L/E: Lower extremity, U/E: Upper extremity, ES: Effect size, SRM: Standardized response mean.

were closely associated with the Jebsen-Taylor hand function and grip power. Previous studies have reported that balance and motor functions are closely related to each other[13, 32]. Moreover, Boonsinsukh et al.[33] proposed that balance and the lower extremity motor function of the FMA can be used together as assessment of gait training for patients with hemiplegia. As demonstrated by the above studies, the interrelation between motor domains of the BBS and the FMA, two representative balance measurement tools for stroke patients, needs to be examined in addition to assessment with motor function tools like the Motor Assessment Scale.

With regard to responsiveness, the ES and SRM of motor functions of the FMA were more than 0.8 (Table 5), which suggests a high responsiveness. Hsueh et al.[34] measured the motor function responsiveness of 50 stroke patients using the FMA, prior to hospitalization and after discharge, ES at 0.38 and a SRM of 1.16. The reason why their ES result was small, unlike that of the present study, is probably because the standard deviation at the baseline was larger than that of the change after discharge. Nonetheless, the SRM represents responsiveness better than ES and the larger the SRM, the higher responsiveness[35]. This study had some limitations. Since it was conducted in three regional rehabilitation hospitals, the results cannot be generalized. In addition, a long-term follow-up study may provide better information on the FMA responsiveness. Also, we didn't investigate possible differences of the test-retest reliability among the three raters according to their clinical experience.

In conclusion, each dimension and the total score of the FMA showed acceptable levels of relative and absolute reliabilities. Concurrent validity and responsiveness were moderate to good, and moderate to large, respectively.

REFERENCES

1) Fugl-Meyer AR, Jääskö L, Leyman I, et al.: The post-stroke hemiplegic patient 1. A method for evaluation of physical performance. Scand J Rehabil Med, 1975, 7: 13–31. [Medline]
2) Sullivan KJ, Tilson JK, Cen SY, et al.: Fugl-meyer assessment of sensorimotor function after stroke: standardized training procedure for clinical practice and clinical trial. Stroke, 2011, 42: 427–432. [Medline] [CrossRef]
3) Stemler SE: A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. Pract Assess, Res Eval, 2004, 9: 80–89.
4) Diehr P, Chen L, Patrick D, et al.: Reliability, effect size, and responsiveness of health status measures in the design of randomized and cluster-randomized trials. Contemp Clin Trials, 2005, 26: 45–58. [Medline] [CrossRef]
5) Lu YM, Lin JH, Hsiao SF, et al.: The relative and absolute reliability of leg muscle strength testing by a handheld dynamometer. J Strength Cond Res, 2011, 25: 1065–1071. [Medline]
6) Harris KD, Heer DM, Roy TC, et al.: Reliability of a measurement of neck flexor muscle endurance. Phys Ther, 2005, 85: 1349–1355. [Medline]
7) Flansbjer UB, Holmbäck AM, Downham D, et al.: What change in isokinetic knee muscle strength can be detected in men and women with hemiparesis after stroke? Clin Rehabil, 2005, 19: 514–522. [Medline] [CrossRef]
8) Jette AM, Tao W, Norweg A: Interpreting rehabilitation outcome measurements. J Rehabil Med, 2007, 39: 585–590. [Medline] [CrossRef]
9) Kurtaiş Y, Küçükdeveci A, Elhan A, et al.: Psychometric properties of the Rivermead motor assessment: its utility in stroke. J Rehabil Med, 2009, 41: 1055–1061. [Medline] [CrossRef]
10) Bolognini N, Vallar G, Casati C, et al.: Neurophysiological and behavioral effects of tDCS combined with constraint-induced movement therapy in poststroke patients. Neurorehabil Neural Repair, 2011, 25: 819–829. [Medline] [CrossRef]
11) Wu CW, Seo HJ, Cohen LG: Influence of electric somatosensory stimulation on paretic-hand function in chronic stroke. Arch Phys Med Rehabil, 2006, 87: 351–357. [Medline] [CrossRef]
12) Lang CE, Wagner JM, Dromerick AW, et al.: Measurement of upper-extremity function early after stroke: properties of the action research arm test. Arch Phys Med Rehabil, 2006, 87: 1605–1610. [Medline] [CrossRef]
13) Sullivan JE, Andrews AW, Lanzino D, et al.: Outcome measures in neurological physical therapy practice: part II. A patient-centered process. J Neurol Phys Ther, 2011, 35: 65–74. [Medline]
14) Langhammer B, Stanghelle JK: Bobath or motor relearning programme? A comparison of two different approaches of physiotherapy in stroke rehabilitation: a randomized controlled study. Clin Rehabil, 2000, 14: 361–369. [Medline] [CrossRef]
15) Sabut SK, Sikdar C, Kumar R, et al.: Functional electrical stimulation of dorsiflexor muscle: effects on dorsiflexor strength, plantarflexor spasticity, and motor recovery in stroke patients. NeuroRehabilitation, 2011, 29: 393–400. [Medline]
16) Hijmans JM, Hale LA, Satherley JA, et al.: Bilateral upper-limb rehabilitation after stroke using a movement-based game controller. J Rehabil Res Dev, 2011, 48: 1005–1013. [Medline] [CrossRef]
17) Lin KC, Chang YF, Wu CY, et al.: Effects of constraint-induced therapy versus bilateral arm training on motor performance, daily functions, and quality of life in stroke survivors. Neurorehabil Neural Repair, 2009, 23: 441–448. [Medline] [CrossRef]
18) Platz T, Pinkowski C, van Wijck F, et al.: Reliability and validity of arm function assessment with standardized guidelines for the Fugl-meyer test, action research arm test and box and block test: a multicenter study. Clin Rehabil, 2005, 19: 404–411. [Medline] [CrossRef]
19) van der Lee JH, Beckerman H, Lankhorst GJ, et al.: The responsiveness of the action research arm test and the fugl-meyer assessment scale in chronic stroke patients. J Rehabil Med, 2001, 33: 110–113. [Medline] [CrossRef]
20) Michaelsen SM, Rocha AS, Knabben RJ, et al.: Translation, adaptation and inter-rater reliability of the administration manual for the Fugl-meyer assessment. Rev Bras Fisioter, 2011, 15: 80–88. [Medline] [CrossRef]
21) Andresen EM, Bowley N, Rothenberg BM, et al.: Test-retest performance of a mailed version of the Medical Outcomes Study 36-Item Short-Form Health Survey among older adults. Med Care, 1996, 34: 1165–1170. [Medline] [CrossRef]
22) Meyer CR: Measurement in physical education. New York: NY Ronald Press.
23) Damstra J, Huddleston Slater JJ, Fourie Z, et al.: Reliability and the smallest detectable differences of lateral cephalometric measurements. Am J Orthod Dentofacial Orthop, 2010, 138: 546.e1-546e8.
24) Bland JM, Altman DG: Measurement error and correlation coefficients. BMJ, 1996, 313: 41–42. [Medline] [CrossRef]
25) Polit DF, Hungler BP: Essential of nursing research: methods, appraisal, and utilization. Philadelphia: Lippincott Williams & Wilkins, 1989.
26) Cohen C: Statistical Power Analysis for the behavioral sciences. 2nd ed.

Hillsdale, NJ: Lawrence Erlbaum Associates, 1988.
27) Gladstone DJ, Danells G, Black SF: The Fugl-meyer assessment of motor recovery after stroke: a critical review of its measurement properties. Neurorehabil Neural Repair, 2002, 16: 232–240. [Medline] [CrossRef]
28) Deakin A, Hill H, Pomeroy VM: Rough guide to the Fugl-meyer assessment. Physiotherapy, 2003, 89: 751–763. [CrossRef]
29) Marx RG, Menezes A, Horovitz L, et al.: A comparison of two time intervals for test-retest reliability of health status instruments. J Clin Epidemiol, 2003, 56: 730–735. [Medline] [CrossRef]
30) Stratford PW, Goldsmith CH: Use of the standard error as a reliability index of interest: an applied example using elbow flexor strength data. Phys Ther, 1997, 77: 745–750. [Medline]
31) Beckerman H, Roebroeck ME, Lankhorst GJ, et al.: Smallest real difference, a link between reproducibility and responsiveness. Qual Life Res, 2001, 10: 571–578. [Medline] [CrossRef]
32) Furanzo AC, Guerreiro CT, Galetti FC, et al.: Modified constant-induced movement therapy and modified forced-use therapy for stroke patients are both effective to promote balance and gait improvements. Res Bras Fisioter. 2012.
33) Boonsinsukh R, Panichareon L, Saengsirisuwan V, et al.: Clinical identification for the use of light touch cues with a cane in gait rehabilitation poststroke. Top Stroke Rehabil, 2011, 18: 633–642. [Medline] [CrossRef]
34) Hsueh IP, Hsu MJ, Sheu CF, et al.: Psychometric comparisons of 2 versions of the fugl-meyer motor scale and 2 versions of the stroke rehabilitation assessment of movement. Neurorehabil Neural Repair, 2008, 22: 737–744. [Medline] [CrossRef]
35) Wang HH, Liao HF, Hsieh CL: Reliability, sensitivity to change, and responsiveness of the Peabody developmental motor scales-second edition for children with cerebral palsy. Phys Ther, 2006, 86: 1351–1359. [Medline] [CrossRef]