Published in final edited form as:

Arch Phys Med Rehabil. 2019 January ; 100(1): 140–143.e1. doi:10.1016/j.apmr.2018.07.433.

Convergent Validity and Responsiveness of the Stroke Upper Limb Capacity Scale

Jayne S. Knutson, PhD,

MetroHealth Rehabilitation Institute, MetroHealth Medical Center, Cleveland, OH; Cleveland Functional Electrical Stimulation Center, Cleveland, OH; Department of Physical Medicine and Rehabilitation, Case Western Reserve University, Cleveland, OH

Amy S. Friedl, OTR/L,

MetroHealth Rehabilitation Institute, MetroHealth Medical Center, Cleveland, OH; Cleveland Functional Electrical Stimulation Center, Cleveland, OH

Kristine M. Hansen, PT,

MetroHealth Rehabilitation Institute, MetroHealth Medical Center, Cleveland, OH; Cleveland Functional Electrical Stimulation Center, Cleveland, OH

Terri Z. Hisel, OTR/L, and

MetroHealth Rehabilitation Institute, MetroHealth Medical Center, Cleveland, OH; Cleveland Functional Electrical Stimulation Center, Cleveland, OH

Mary Y. Harley, OT/L

MetroHealth Rehabilitation Institute, MetroHealth Medical Center, Cleveland, OH; Cleveland Functional Electrical Stimulation Center, Cleveland, OH

Abstract

Objective: To evaluate the convergent validity and responsiveness of the Stroke Upper Limb Capacity Scale (SULCS) in comparison to the Arm Motor Ability Test (AMAT), the Box and Blocks Test (BBT), and the upper limb Fugl-Meyer Assessment (FMA). The SULCS is a relatively new measure that was designed to be easier to score and less time consuming than some existing measures.

Design: Prospective repeated-measures design.

Setting: Clinical research laboratory of a large public hospital.

Participants: Patients (n=61) <2 years post-stroke with moderate to severe upper limb hemiparesis.

Correspondence: Jayne S. Knutson, PhD, MetroHealth Old Brooklyn Health Center, 4229 Pearl Road, Suite 5N-27, Cleveland, Ohio 44109, USA, Phone: 216-957-3557 Fax: 216-957-2885, jknutson@metrohealth.org.

Disclosures

Jayne Knutson, Ph.D. is an inventor on U.S. Patent 8,165,685: System and Method for Therapeutic Neuromuscular Electrical Stimulation, issued to Case Western Reserve University.

Clinical Trial Registration: URL: <http://www.clinicaltrials.gov>. Unique identifier: NCT01688856.

Intervention: Participants received 12 weeks of therapy that included neuromuscular electrical stimulation of the paretic finger and thumb extensors. The SULCS, AMAT, BBT, and FMA were administered at weeks 0, 6, 12 (end-of-therapy), 20, 28, and 36 (6 months post-therapy).

Main Outcome Measures: Convergent validity was evaluated with Spearman's correlation coefficients (ρ) between pairs of measures at each time point. Responsiveness from 0 to 12 weeks and 0 to 36 weeks was evaluated with the standardized response mean (SRM).

Results: The SULCS demonstrated strong correlation with the AMAT ($\rho=0.81-0.93$), BBT ($\rho=0.73-0.92$), and FMA ($\rho=0.78-0.92$), at all 6 time points. All four measures had moderate to large SRMs (SULCS, 0.71–0.77; AMAT, 0.83–0.97; BBT, 0.73–0.82; FMA, 0.75–0.76). There was no significant difference in responsiveness among the four measures.

Conclusions: The results support the use of the SULCS to measure upper limb capacity in patients who are less than 2 years post-stroke with moderate to severe hemiplegia.

Keywords

Stroke; rehabilitation; upper extremity; outcomes; psychometrics

The Stroke Upper Limb Capacity Scale (SULCS)¹ is a measure of *capacity*, the ability to execute a task in a standardized setting.² The SULCS was designed to address shortcomings of other measures,^{1,3} in particular, to be suitable for a wide range of patients, especially including those with poor hand-related upper limb capacity, to assess a *single* construct – capacity, and to be easily scored and administered without requiring special equipment or much time. The content validity, unidimensionality, internal consistency, and reliability of the SULCS have been established.^{1,3} Further evaluation of convergent validity, responsiveness (the ability to detect change over time), and floor and ceiling effects is needed before the SULCS can be widely recommended for clinical trials or practice.

Though not specified as a study aim a-priori, our ongoing interventional trial that used the SULCS to measure upper limb capacity provided this opportunity to assess its convergent validity and responsiveness. Convergent validity is established by comparing scores from one measure to scores on other measures of the same or similar constructs. Besides the SULCS, the ongoing trial used the Arm Motor Ability Test (AMAT), Box and Blocks Test (BBT), and upper limb Fugl-Meyer Assessment (FMA). Like the SULCS, the AMAT is a measure of upper limb capacity, while the BBT measures gross manual dexterity and FMA measures upper limb impairment. The purpose of this analysis was to determine if the SULCS is strongly correlated with any of these other commonly used measures and has similar or better responsiveness and acceptable floor and ceiling effects, which would support its use as an alternative to measures that have more complex scoring or are more time consuming.⁴⁻⁶

METHODS

Data from the first 61 participants (convenience sample) who were enrolled in a single-center randomized controlled trial of three neuromuscular electrical stimulation (NMES) therapies were used in this analysis. The primary selection criteria were: <2 years post-

stroke, finger extensor strength 4 out of 5 on the manual muscle test (Medical Research Council scale), 11 out of 14 points on the hand section of the FMA, sufficient active shoulder and elbow movement to volitionally position the paretic hand in the workspace for table-top tasks, and unable to fully extend the elbow and open the hand simultaneously. All participants received 12 weeks of therapy that included ten ~1-hr sessions per week of repetitive NMES of the paretic finger and thumb extensors at home and two 75-min sessions per week in the lab with an occupational therapist focusing on task practice with the paretic hand. The study protocol was approved by the local institutional review board, and all subjects provided written informed consent. The validation and responsiveness analysis of this article was not pre-specified in the study protocol.

A single therapist administered the FMA, BBT, SULCS, and AMAT, in that order, at weeks 0, 6, 12 (end-of-therapy), 20, 28, and 36 (6 months post-therapy). The SULCS consists of 10 tasks that relate to daily activities.^{1,3} Each receives a score of 0 (unable to perform) or 1 (able to perform), resulting in a total score of 0–10. Start-and-stop rules are based on a hierarchical ordering of items, which was reported to result in an average administration time of 6 minutes.¹ The English version of the SULCS, published in *Archives*, was used.¹ The AMAT had 9 compound tasks composed of 1 to 3 sub-tasks, each rated on a 0–5 ordinal scale. The scores for all sub-tasks were averaged to provide a total score of 0–5,⁴ and the test is reported to take approximately 20 minutes to administer.⁵ The BBT is a count of how many times a subject can pick up a 2.5-cm block from one compartment of a box and release it into another compartment of the box in 60 seconds;⁷ total administration time is approximately 5 minutes. The FMA grades volitional movements of the shoulder, elbow, forearm, wrist, and hand with a 3-point scale.⁶ Maximum score is 66, and the average administration time is reported to be 11 minutes.⁸

Convergent validity was evaluated using Spearman's rank correlation coefficients (ρ). We hypothesized that SULCS scores would be strongly and positively related to AMAT, BBT, and FMA scores at all time points. We also hypothesized that change in SULCS scores from baseline to 12 and 36 weeks would be positively related to changes in AMAT, BBT, and FMA scores. Correlations were considered strong if $\rho > 0.70$, inconclusive if $\rho = 0.40$ – 0.70 , and absent if $\rho < 0.40$.⁹ Change scores from 0 to 12 weeks and 0 to 36 weeks were assessed for significance with non-parametric Wilcoxon matched pairs signed rank tests. Responsiveness over those time intervals was assessed with standardized response means (SRM). The SRM is the mean change in score divided by the standard deviation of the change scores,¹⁰ and is considered large if > 0.80 , moderate if 0.50 – 0.80 , and small if 0.2 – 0.5 .¹¹ Tests for significant differences between SRMs of pairs of measures were also conducted (see Supplementary Materials). Floor and ceiling effects were considered present if 15% of participants scored the lowest or highest scores.¹²

RESULTS

The participants were 67% male, had a mean (SD) age of 58 (12.9), months post-stroke 4.3 (3.4), and baseline FMA 31 (9.3), indicating moderate to severe upper limb impairment (see table of Participant Characteristics in Supplementary Materials). The sample size analyzed

was 61 at week 0 and progressively decreased to 40 at week 36 (Table 1) due to 13 withdrawals and 8 losses to follow-up. Missing data were not replaced.

The correlations between the SULCS and the AMAT, BBT, and FMA were all strong ($p > 0.70$) (Table 1). The longitudinal correlations at week 12 were inconclusive ($p = 0.45$ to 0.52), but by week 36 were strong ($p = 0.68$ to 0.78) (Table 2).

The change in scores from week 0 to 12 and week 0 to 36 were significant for all four measures (Table 2). All four measures had moderate to large SRMs (Table 2), and there were no statistically significant differences between the SRMs of any pair of measures at either 12 or 36 weeks (see Supplementary Materials). The floor and ceiling effects for the SULCS across the 6 time points were acceptable ($< 15\%$); however, the floor effect for the BBT exceeded 15% at three time points (Table 1). The AMAT and FMA showed no floor or ceiling effects.

DISCUSSION

The SULCS scores were strongly correlated with the AMAT, BBT, and FMA, and these scales were all moderately to highly responsive to change over the 12- and 36-week time periods of the study. Because both the AMAT and SULCS are designed to measure the same construct – upper limb capacity – the strong correlation found between them was expected. The slightly lower correlations with BBT the FMA were also expected as those measures do not as strongly reflect upper limb capacity as the AMAT, the BBT focusing mainly on dexterity and the FMA measuring impairment. Positive and significant correlations between the respective change scores, especially at 36 weeks (6 months post-therapy) further support the SULCS' convergent validity.

Comparing responsiveness metrics (e.g., SRMs) generated from different studies may reflect differences in interventions or subject populations rather than a property of the outcome measure itself. However, the relative responsiveness of different measures within a study may be useful in helping decide which measures to use. In terms of SRM, there were no statistically significant differences between SRMs of the different measures, suggesting that all four measures have similar responsiveness. The findings of this study further suggest that the SULCS may have no significant ceiling or floor effects in patients who have moderate to severely impaired upper limb capacity and undergo up to 12 weeks of therapy and are followed for 6 more months.

Study Limitations

While this study is in line with COSMIN guidelines,¹³ it was not powered a-priori for the analyses reported here, and the modest sample size may limit the conclusions that can be made. Only 20% of the participants were > 15 months post-stroke by the end of the study; therefore, the findings here may not be generalizable to patients who are beyond that timeframe. Future work should study the properties of the SULCS in patients who are beyond 2yrs post-stroke. Also, only patients with moderate to severe impairment were included; therefore, the results may not apply to those with mild impairment.

CONCLUSIONS

The SULCS strongly correlates with and is similarly responsive as the AMAT, BBT, and FMA in patients who are less than 2 years post-stroke with moderate to severe hemiplegia. These findings support the use of the SULCS, a recently developed measure of post-stroke upper limb capacity that may be easier to score and less time consuming than other upper limb capacity measures.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Financial Support

This work was supported by the NIH National Institute of Child Health and Human Development's National Center for Medical Rehabilitation Research grant number R01HD068588.

Abbreviations

AMAT	Arm Motor Ability Test
BBT	Box and Blocks Test
FMA	Fugl-Meyer Assessment
NMES	Neuromuscular Electrical Stimulation
SRM	Standardized Response Mean
SULCS	Stroke Upper Limb Capacity Scale

REFERENCES

1. Roorda LD, Houwink A, Smits W, Molenaar IW, Geurts AC. Measuring upper limb capacity in poststroke patients: development, fit of the monotone homogeneity model, unidimensionality, fit of the double monotonicity model, differential item functioning, internal consistency, and feasibility of the stroke upper limb capacity scale, SULCS. *Arch Phys Med Rehabil* 2011;92:214–227. [PubMed: 21272717]
2. World Health Organization. How to use the ICF: a practical manual for using the International Classification of Functioning, Disability and Health (ICF). Geneva: World Health Organization; 10, 2013.
3. Houwink A, Roorda LD, Smits W, Molenaar IW, Geurts AC. Measuring upper limb capacity in patients after stroke: reliability and validity of the stroke upper limb capacity scale. *Arch Phys Med Rehabil* 2011;92:1418–1422. [PubMed: 21802656]
4. Kopp B, Kunkel A, Flor H, Platz T, Rose U, Mauritz KH, et al. The Arm Motor Ability Test: reliability, validity, and sensitivity to change of an instrument for assessing disabilities in activities of daily living. *Arch Phys Med Rehabil* 1997;78:615–620. [PubMed: 9196469]
5. O'Dell MW, Kim G, Finnen LR, Polistena C. Clinical implications of using the arm motor ability test in stroke rehabilitation. *Arch Phys Med Rehabil* 2011;92:830–836. [PubMed: 21435635]
6. Fugl-Meyer AR, Jaasko L, Leyman I, Olsson S, Steglind S. The post-stroke hemiplegic patient. 1. a method for evaluation of physical performance. *Scand J Rehabil Med* 1975;7:13–31. [PubMed: 1135616]

7. Desrosiers J, Bravo G, Hebert R, Dutil E, Mercier L. Validation of the Box and Block Test as a measure of dexterity of elderly people: reliability, validity, and norms studies. *Arch Phys Med Rehabil* 1994;75:751–755. [PubMed: 8024419]
8. De Weerdt WJG. Measuring recovery of arm-hand function in stroke patients: a comparison of the Brunnstrom-Fugl-Meyer test and the Action Research Arm test. *Physiother Can* 1985;37:65–70.
9. Post MW. What to Do With “Moderate” Reliability and Validity Coefficients? *Arch Phys Med Rehabil* 2016;97:1051–1052. [PubMed: 27095143]
10. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. *Journal of clinical epidemiology* 2000;53:459–468. [PubMed: 10812317]
11. Cohen J Statistical Power Analysis for the Behavioral Sciences. New York: Academic Press; 1977.
12. van de Ven-Stevens LA, Munneke M, Terwee CB, Spauwen PH, van der Linde H. Clinimetric properties of instruments to assess activities in patients with hand injury: a systematic review of the literature. *Arch Phys Med Rehabil* 2009;90:151–169. [PubMed: 19154842]
13. Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018;27:1147–1157. [PubMed: 29435801]

Table 1.Scores, Spearman's ρ Between Measures, Floor and Ceiling Effects

	Therapy Period			Follow-Up Period		
	Week 0 (N=61)	Week 6 (N=53)	Week 12 (N=47)	Week 20 (N=47)	Week 28 (N=43)	Week 36 (N=40)
SULCS						
Score $\bar{x}$, mean (SD)	3.6 (2.3)	4.4 (2.9)	4.8 (2.9)	5.0 (2.9)	5.5 (3.2)	5.3 (3.2)
vs AMAT, ρ	0.81 *	0.89 *	0.87 *	0.84 *	0.90 *	0.93 *
vs BBT, ρ	0.73 *	0.87 *	0.89 *	0.88 *	0.89 *	0.92 *
vs FMA, ρ	0.78 *	0.85 *	0.86 *	0.86 *	0.92 *	0.89 *
Floor effect, %	11	13	2	6	2	5
Ceiling effect, %	0	2	2	2	7	10
AMAT						
Score $\bar{x}$, mean (SD)	2.05 (0.82)	2.4 (1.08)	2.55 (1.16)	2.67 (1.23)	2.81 (1.32)	2.86 (1.31)
vs BBT, ρ	0.87 *	0.92 *	0.94 *	0.93 *	0.96 *	0.94 *
vs FMA, ρ	0.87 *	0.88 *	0.92 *	0.86 *	0.92 *	0.93 *
Floor effect, %	0	0	0	0	0	0
Ceiling effect, %	0	0	0	2	0	0
BBT						
Score $\bar{x}$, mean (SD)	12 (11.7)	18 (15.7)	20 (17.6)	21 (17.8)	23 (19.0)	24 (20.0)
vs FMA, ρ	0.74 *	0.83 *	0.87 *	0.86 *	0.90 *	0.89 *
Floor effect, %	20	15	11	11	16	18
Ceiling effect, %	n/a	n/a	n/a	n/a	n/a	n/a
FMA						
Score $\bar{x}$, mean (SD)	30.3 (9.3)	34.3 (10.5)	34.6 (10.9)	36.5 (11.5)	38.4 (12.2)	37.3 (12.6)
Floor effect, %	0	0	0	0	0	0
Ceiling effect, %	0	0	0	2	0	0

$\bar{x}$ higher score indicates better outcome

* statistically significant, $p < .001$

Table 2.

Change, Responsiveness, and Longitudinal Correlations of the SULCS with the AMAT, BBT, and FMA

	Week 0 to Week 12 (End-of-Therapy)					Week 0 to Week 36 (6 Months Post-Therapy)						
	Change	Responsiveness		Longitudinal Correlation (ρ)		Change	Responsiveness		Longitudinal Correlation (ρ)			
	Mean (SD)	Wilcoxon Test Z-Value	SRM	AMAT	BBT	FMA	Mean (SD)	Wilcoxon Test Z-Value	SRM	AMAT	BBT	FMA
SULCS	1.5 (2.1)	4.3 *	0.71	0.52 *	0.52 *	0.45	2.0 (2.6)	4.2 *	0.77	0.78 *	0.68 *	0.71 *
AMAT	0.57 (0.69)	5.1 *	0.83		0.80 *	0.54 *	0.84 (0.86)	4.9 *	0.97		0.85 *	0.76 *
BBT	8.8 (12.0)	4.9 *	0.73			0.46	11.8 (14.3)	4.8 *	0.82			0.64 *
FMA	5.1 (6.9)	4.4 *	0.75				7.1 (9.4)	4.0 *	0.76			

* p<0.001