ORIGINAL ARTICLE

Responsiveness, Minimal Detectable Change, and Minimal Clinically Important Difference of the Nottingham Extended Activities of Daily Living Scale in Patients With Improved Performance After Stroke Rehabilitation

Ching-yi Wu, ScD, OTR, Li-ling Chuang, PhD, PT,* Keh-chung Lin, ScD, OTR, Shin-da Lee, PhD, PT, Wei-hsien Hong, PhD*

ABSTRACT. Wu C-Y, Chuang L-L, Lin K-C, Lee S-D, Hong W-H. Responsiveness, minimal detectable change, and minimal clinically important difference of the Nottingham Extended Activities of Daily Living scale in patients with improved performance after stroke rehabilitation. Arch Phys Med Rehabil 2011;92:1281-7.

Objectives: To determine the responsiveness, minimal detectable change (MDC), and minimal clinically important differences (MCIDs) of the Nottingham Extended Activities of Daily Living (NEADL) scale and to assess percentages of patients' change scores exceeding the MDC and MCID after stroke rehabilitation.

Design: Secondary analyses of patients who received stroke rehabilitation therapy.

Setting: Medical centers.

Participants: Patients with stroke (N=78).

Interventions: Secondary analyses of patients who received 1 of 4 rehabilitation interventions.

Main Outcome Measures: Responsiveness (standardized response mean [SRM]), 90% confidence that a change score at this threshold or higher is true and reliable rather than measurement error (MDC_{90}), and MCID on the NEADL score and percentages of patients exceeding the MDC_{90} and MCID.

Results: The SRM of the total NEADL scale was 1.3. The MDC_{90} value for the total NEADL scale was 4.9, whereas minima and maxima of the MCID for total NEADL score were 2.4 and 6.1 points, respectively. Percentages of patients exceeding the MDC_{90} and MCID of the total NEADL score were 50.0%, 73.1%, and 32.1%, respectively.

Conclusions: The NEADL is a responsive instrument relevant for measuring change in instrumental activities of daily living after stroke rehabilitation. A patient's change score has to reach 4.9 points on the total to indicate a true change. The mean change score of a stroke group on the total NEADL scale should achieve 6.1 points to be regarded as clinically important. Our findings are based on patients with improved NEADL performance after they received specific interventions. Future research with larger sample sizes is warranted to validate these estimates.

Key Words: Clinimetrics; Nottingham Extended Activities of Daily Living; Rehabilitation; Stroke rehabilitation.

EACH YEAR, APPROXIMATELY 795,000 people in the United States experience a new or recurrent stroke and the estimated cost of stroke for 2010 was $73.7 billion.[1] Stroke is one of the leading causes of lifetime disabilities, with the consequences of functional difficulties and activity limitations.[1] One study reported that 62% of patients were still dependent in basic activities of daily living (BADLs) and 32% were inactive in instrumental activities of daily living (IADLs) at 3 years after stroke.[2] Performing IADLs requires increased interaction with the environment and appears to be a prerequisite for independent living in the household or community.[3]

A systematic review of ADL instruments identified 27 activities of daily living (ADLs) scales originally developed in the field of clinical neurology, and item content from the scales was categorized to BADL (basic mobility and self-care) or IADL (eg, household activities, community activities, social/recreational activities, cognitive activities) domains or non-ADL.[4] Ten neurologic ADL scales designed for stroke primarily focus on the BADL domain, whereas the Nottingham Extended Activities of Daily Living (NEADL) scale and Stroke

From the Department of Occupational Therapy and Graduate Institute of Clinical Behavioral Science, College of Medicine, Chang Gung University, Taoyuan (Wu); School of Occupational Therapy, College of Medicine, National Taiwan University Hospital, Taipei (Chuang, Lin) and Division of Occupational Therapy, Department of Physical Medicine and Rehabilitation, National Taiwan University Hospital, Taipei (Lin); and Department of Physical Therapy and Graduate Institute of Rehabilitation Science (Lee) and Department of Sports Medicine (Hong), China Medical University, Taichung, Taiwan.

*Wu and Chuang contributed equally to this article.

Supported by the China Medical University research fund (grant no. 99F008-310), National Science Council (grant nos. NSC-96-2628-B-002-033-MY2, NSC-97-2314-B-002-008-MY3), NSC-97-2314-B-182-004-MY3 and National Health Research Institutes (grant nos. NHRI-EX99-9742PI, NHRI-EX99-9920PI) in Taiwan.

No commercial party having a direct financial interest in the results of the research supporting this article has or will confer a benefit on the authors or on any organization with which the authors are associated.

Clinical trial registration numbers NCT00780533 and NCT00917605.

Correspondence to Keh-chung Lin, ScD, OTR, 17, F4, Xu Zhou Rd, Taipei, Taiwan, e-mail: *kehchunglin@ntu.edu.tw*. Reprints are not available from the author.

0003-9993/11/9208-00970$36.00/0
doi:10.1016/j.apmr.2011.03.008

List of Abbreviations

ADLs	activities of daily living
BADL	basic activity of daily living
CI	confidence interval
IADL	instrumental activity of daily living
NEADL	Nottingham Extended Activities of Daily Living
MDC	minimal detectable change
MDC_{90}	90% confident that a change score at this threshold or higher is true and reliable rather than measurement error
MCID	minimal clinically important difference
SIS	Stroke Impact Scale
SRM	standardized response mean
SD_{pooled}	pooled standard deviation

Impact Scale (SIS) comprehensively assess both BADL and IADL domains.[4]

The NEADL scale is a measure of IADL outcomes developed specifically for the stroke population.[3,5] It also is 1 of the most frequently used IADL measures in randomized controlled trials of stroke rehabilitation.[6] According to the World Health Organization's International Classification of Functioning, Disability, and Health model,[7] the NEADL measures levels of activity and participation that may relate to quality of life after stroke.[8] The NEADL has well-established reliability, validity, and responsiveness in patients with stroke.[9-13] However, no estimates were found of measurement characteristics associated with responsiveness, minimal detectable change (MDC), and minimal clinically important difference (MCID) for the NEADL score in patients receiving stroke rehabilitation.

Responsiveness of a scale refers to the ability to detect change, which is an important quality for assessing treatment effectiveness.[14] Determination of the MDC and MCID is critical for judging whether treatments have resulted in real change and the magnitude of the benefit of interventions.[14,15] The MDC is the minimal amount of change score outside of measurement error that may reflect true change.[16] The MCID is the smallest meaningful change in score considered clinically important and constitutes a score difference related to a meaningful and beneficial change in health status perceived by the patient.[17,18] It has been advocated that MCID estimates should be based on multiple approaches,[19] such as a combination of the distribution- and anchor-based approaches, to triangulate a small interval of values for the MCID.[17,18,20] Distribution-based methods are based on the statistical distributions of change scores and proportions of the SD.[21] Anchor-based methods anchor or map the outcome measure of change onto an external criterion of change[22] that is considered important to patients or clinicians and directly reflects their point of view.[23]

Quantifying changes in NEADL scores after rehabilitation interventions should improve interpretation of the NEADL measure and extend our understanding of the meaning and clinical values of change scores in this measure. Reporting both the MDC and MCID values will provide information about 2 important benchmarks for whether the observed changes can be considered reliable and clinically important. The clinical significance of outcomes is critical to interpretation of results of a clinical trial,[24] and this information helps guide decision making about the discontinuation or alternation of an intervention that aims to improve IADL function. To date, the capacity of the NEADL to detect change after stroke interventions is unclear and no published estimates of MDC and MCID values for the NEADL are available. This study extended our recent study[13] by recruiting a larger sample to investigate both internal and external responsiveness for patients with improved NEADL performance after stroke rehabilitation. Internal responsiveness was defined as the ability of a measure to change over time, and external responsiveness reflects the extent to which changes in a measure over a specified time frame relate to corresponding changes in a reference measure of health status.[25] We used the standardized response mean (SRM) and MCID to estimate the internal and external responsiveness of the NEADL in patients with stroke, respectively. In addition, we estimated the MDC for judging whether treatments resulted in a real change beyond measurement error. Specifically, this study aimed to establish the responsiveness, MDC, and MCID of the NEADL and to assess percentages of patients' change scores exceeding the MDC and MCID in patients with stroke with improved NEADL performance after stroke rehabilitation.

METHODS

Participants

This study was a secondary analysis of data from randomized controlled trials of stroke rehabilitation therapies.[26-28] All participants in this study were recruited from 3 medical centers and enrolled in clinical investigations of the effects of distributed constraint-induced therapy, bilateral arm training, and robot-assisted therapy for stroke patients.[26-28] A cohort of 78 patients with no missing values for the 2 outcome measures was included in the present study. Inclusion criteria for participants were (1) improved performance on total NEADL scores after intervention, (2) poststroke onset of at least 6 months, (3) demonstration of Brunnstrom stage III or higher for the affected upper extremity,[29] (4) no excessive spasticity in the upper extremity (Modified Ashworth Scale score $\leq$2.5),[30] and (5) able to follow instructions to self-rate the measures included and perform therapeutic activities. Participants with cognitive impairments were excluded (score of $<$24 points on the Mini-Mental State Examination).[31] Institutional review board approval was obtained from the participating sites, and participants signed a document of informed consent.

Procedures

Participants were randomly assigned to receive 1 of 4 rehabilitation therapies: distributed constraint-induced therapy, bilateral arm training, robot-assisted therapy, or conventional rehabilitation for a total of 30 therapy hours in 3 to 4 weeks. The contents of the rehabilitation therapies are described in detail elsewhere.[26-28] Because the interventions involved functional training of movement, patients should directly obtain gains in physical performance and daily function. Positive effects of the interventions on IADL functions have been reported in previous studies.[13,26,27,32] Data for improved NEADL total scores after rehabilitation were used for analysis. The outcome measures were administered before and after treatment by 3 occupational therapists masked to participant group and trained to properly administer the outcome measures.

Outcome Measures

NEADL scale. The NEADL scale is a measure of independence in 4 areas of daily life: mobility, kitchen, domestic, and leisure activities.[12] Each item on the scale is rated on a 4-point scale (0 = unable to 3 = able). Higher scores represent greater independence. We used a modified version of the NEADL to address issues of cultural relevance. The validity of this modified NEADL scale was established in patients with stroke.[11]

SIS, version 3, ADL/IADL domain. The SIS, a stroke-specific quality-of-life measure that includes 59 items and assesses 8 domains related to activities and participation,[33] was administered before and after treatment. Only the ADL/IADL domain of the SIS was reported, including 10 questions about patients' perceived competency to perform their activities, such as dressing, bathing, control of bladder/bowels, light/heavy household tasks, and shopping. Each item is scored from "not difficult at all" to "could not do at all" on a 5-point scale. The maximum score for the SIS ADL/IADL is 50 points. The SIS has shown excellent clinimetric properties in responsiveness and concurrent and construct validity, as well as test-retest reliability (intraclass correlation coefficient, .57–.92).[34,35] When the therapists administered the SIS questionnaire, they explained the purpose of the SIS and asked patients questions about how stroke had affected their health and life from patients' points of view, without family or caregiver assistance or coaching.

We chose the ADL/IADL domain of the SIS as the anchor for calculation of the MCID estimates of the NEADL because scores on this domain directly reflect the point of view of patients regarding their perceived participation in activities after stroke. Returning to usual activities, social integration, and good quality of life are issues important to patients.[8] Competence in IADLs, especially leisure activities, showed a strong association with better overall well-being and quality of life after stroke.[8] The SIS is a widely accepted measure of activity and participation in stroke rehabilitation.[33-35] Compared with other quality-of-life measures, the SIS showed sound properties in patients with stroke.[36] In addition, previous research showed good correlations between scores of the ADL/IADL domain of the SIS and the Barthel Index, Lawton IADL, and NEADL.[35,36] Thus, patients' perceived competency to perform the activities in the ADL/IADL domain of the SIS appears to be a meaningful anchor for the MCID estimates of the NEADL in that it is clinically interpretable and conceptually associated with the NEADL.[19]

Although no defined range of the change score has been identified for determination of the MCID group, previous studies have set the MCID at 10% of the total range of the scales in a chronic hemiparetic population,[37] 11% on the Pediatric Evaluation of Disability Inventory,[38] and 15% on the visual analog scale of back pain.[39] These researchers arbitrarily set the MCID at a 10% to 15% change based on their clinical experience and the MCID of similar outcome measures because no established MCID estimates of the instruments used have been reported in the literature. On the basis of our previous work,[40] the anchor-based MCID estimate reported for the SIS ADL/IADL was established as 5.9 points, which also was within the range of 10% to 15% of the maximum score. If a patient's change score from pre- to posttreatment on the SIS ADL/IADL reached 5 to 7.5 points, which is 10% to 15% of the maximum score, the patient was classified in the MCID group.

Data Analysis

Responsiveness. The SRM provided measurements of responsiveness and was calculated by dividing the mean difference in scores of participants by the SD of the change scores.[41] The NEADL change score was calculated by subtracting the pretreatment score from the posttreatment score, with positive (negative) values representing improvement (deterioration) after intervention. It is arguably problematic to calculate responsiveness based on pooling improvement and deterioration together because a measure that has positive and negative change scores will have a small mean change score that results in a higher level of variability in change scores and a small SRM value. Therefore, it was recommended to calculate responsiveness for only the improved patients.[25] With a sufficient sample size, a comparison of change in the measure between patients who improved and those who did not should be analyzed separately.[25] In this study, we calculated the responsiveness of the NEADL score for patients who showed beneficial effects after stroke interventions in the light of previous research.[13,26,27,32] According to the criteria of Cohen,[42] SRM values of 0.2, 0.5, and 0.8 represent small, moderate, and large values for responsiveness, respectively. Bootstrap 1000 samples with replacement were used to estimate 95% confidence intervals (CIs) for the SRMs.[43] R software (Version 2.9.1)[a] was used for statistical computing.

MDC calculation. The MDC is calculated by multiplying the standard error of measurement by the z score associated with the desired confidence level and the square root of 2, adjusting for sampling from 2 different measures.[44] The standard error of measurement is estimated as the pooled standard deviation (SD_{pooled}) of pre- and posttreatment assessments multiplied by the square root of $(1-r)$, where r is the intraclass correlation coefficient.[45] The standard error of measurement quantifies within-subject variability and reflects the amount of measurement error.[44,46] The MDC is estimated based on the 90% confidence interval (CI) ($z=1.65$). The MDC_{90} is the most common standard used in the literature[47] and means that one can be 90% confident that a change score at this threshold or higher is true and reliable rather than measurement error.[14] Test-retest reliability was determined by using a set of independent data from 19 patients with chronic stroke without intervention in 2 baselines with a 2-week interval. The single measure of the intraclass correlation coefficient was calculated by using a 2-way mixed-effect model with an agreement coefficient. SPSS software, Version 16.0[b] was used for statistical analyses.

The MDC_{90} was calculated by using the following formula: $MDC_{90}=1.65\times\sqrt{2}\times$standard error of measurement$=1.65\times\sqrt{2}\times SD_{pooled}\times\sqrt{(1-r)}$, where 1.65 is the 2-tailed tabled z value for the 90% CI, $\sqrt{2}$ represents the variance of 2 measures, and r is the coefficient of test-retest reliability. The test-retest reliability coefficient for the NEADL total score was .97, which is the same as .97 for patients undergoing total hip replacement[48] and healthy volunteers.[49]

Estimates of MCID. Distribution- and anchor-based approaches were both used to determine the MCID of the NEADL. The distribution-based MCID estimate was determined by using the Cohen effect-size benchmark. Cohen suggested that score differences of 0.2SD units correspond to small but important changes in treatment-effectiveness research.[42] Although the cutoff value used to describe the magnitude of the effect size was set arbitrarily, most research has widely accepted Cohen's benchmark.[50-52] An effect size of 0.2 was advocated as a reasonable method to estimate the MCID.[51,52] Thus, 0.2SD of baseline was used to estimate the distribution-based MCID in this study.

The anchor-based approach applying a relevant external criterion provides meaningful estimates of the measure's MCID[18]; that is, relating change scores in an instrument to an external standard of clinical change (eg, patients' global ratings of change in health) to establish the MCID.[25] The MCID reflects the magnitude of change in the measure associated with an arbitrary definition of the smallest important change in the external criterion, and it may be estimated by the average change score in patients rating some improvement on NEADL scores.[25] Selecting an anchor should be based on criteria of relevance to the study context for which the anchor has a relationship with the patient-reported outcome.[18,50] Therefore, the anchor-based MCID estimate was calculated as mean change score on the NEADL, corresponding to patients defined as having MCID; that is, those with a perceived overall change score of 5 to 7.5 points (10%–15% of the total scale score range) on the ADL/IADL domain of the SIS.

MDC and MCID percentages. Percentages of patients with change scores exceeding the values of the MDC_{90} and MCID estimates were examined to evaluate the extent of patients' changes after interventions detected by using the NEADL. Patients who did not benefit from interventions were excluded from the estimation of responsiveness because of limited sample sizes. The greater the percentages of patients who exceed the values, the more responsive the measure.

RESULTS

Demographic and clinical characteristics of the 78 participants (51 men, 27 women; mean age, 54y) with an im-

Table 1: Demographics and Clinical Characteristics of Participants

Characteristics	Value
Age (y)	54.3±11.9
Sex	
Men	51
Women	27
Side of stroke lesion	
Left	35
Right	43
Time since stroke (mo)	19.7±17.0
Median Brunnstrom stage of the proximal/distal UE	4/4
Mini-Mental State Examination score	27.4±2.3
Pretreatment evaluations	
NEADL total score	28.1±12.1
SIS ADL/IADL score	71.8±17.0
Posttreatment evaluations	
NEADL total score	33.6±12.2
SIS ADL/IADL score	77.2±16.8

NOTE. N=78. Values expressed as mean ± SD unless noted otherwise.
Abbreviation: UE, upper extremity.

proved NEADL total score at posttreatment are listed in table 1. Although the mean change in NEADL total score in these 78 patients after treatment is positive, some showed negative values in 1 or 2 subscales, but the sum of the mean change score in the 4 subscales was higher than 0. Therefore, the change score of the overall NEADL rather than subscale scores was reported to avoid confusion. NEADL total score and the SRM for patients who improved after interventions are listed in table 2. The SRM was 1.3 for total score. The NEADL was highly responsive to change by using the Cohen criteria.

Results of the MDC_{90} and MCID estimates of the total NEADL scores are listed in table 3. The MDC_{90} of the total NEADL score was 4.9 (8.6% of scale width). Distribution-based MCID estimates (ie, 0.2 times baseline SD) for total NEADL score equated to a change of 2.4 (4.2% of scale width). As calculated from the 15 patients with SIS ADL/IADL change scores of 5 to 7.5 points, the corresponding mean change score on the NEADL was 6.1 (10.6% of scale width).

As listed in table 4, 50.0% of patients had a positive change that exceeded the MDC_{90} of the total NEADL. Approximately three-quarters of the change scores exceeded the 0.2SD of the total NEADL (73.1%). Furthermore, changes in approximately one-third of the patients achieved the anchor-based MCID estimates of the total NEADL (32.1%).

DISCUSSION

The present study was conducted to identify the responsiveness, MDC, and MCID for the total NEADL scores after stroke rehabilitation. Overall, total NEADL score was highly responsive after stroke rehabilitation. Analyses indicated that when change score points of an individual patient with stroke reached 4.9 on the total NEADL, clinicians may be 90% confident that the changes are true and reliable. Combining 2 approaches of the MCID, participants in a stroke group who achieve a mean score of 2.4 to 6.1 points on the total NEADL are likely to have had a clinically important change. Our results showed that mean change in the total NEADL score was 5.5 points, which reached the threshold of clinical importance.

The validity of using an assessment to detect clinically important change rests in part on the assumption that the measurement tool is responsive to change.[53] Showing responsiveness is a key component to establish construct validity and is important to determine the MCID.[18] An instrument with good responsiveness is sensitive to clinically meaningful change over time within a patient, and the change is responsive to effects of therapeutic interventions.[54] Therefore, lack of a differential effect of an instrument may be caused by inadequate responsiveness of the instrument rather than ineffectiveness of the treatments.[37] In the present study, the magnitude of the SRM provides evidence that the NEADL is a responsive instrument for outcome study in stroke rehabilitation, consistent with our previous study using the SRM[13] and another stroke study that used effect size to assess responsiveness and showed that NEADL score was sensitive to changes after stroke.[10] The responsiveness index of the NEADL (SRM=1.3) was higher in the present study than in the previous study (SRM=0.9)[13] because only patients with improved NEADL performance after intervention were included in this study. In addition, estimates of responsiveness of the NEADL in our research appear to be different from previous studies[48,55] in which the NEADL was poorly responsive for the differences before and after hip replacement by calculating the total score effect size of 0.3 to 0.5. Possible contributing factors that may influence study results include patient populations, the intervention involved, and the responsiveness index.

In this study, MDC and MCID estimates were both used to interpret change scores on the NEADL after stroke rehabilitation. To decide whether an improvement is clinically important or a measurement error, MDC scores would be useful to help distinguish actual change from measurement error.[56] Furthermore, identifying the MCID is helpful to determine meaning benefit, provide a basis for sample-size estimation in clinical trials, and provide a more precise measure of patient-reported treatment effect.[50] Our results showed that the MDC and MCID differed notably, with the former at 4.9 points, indicating a "minimal" score of change that was free of error, whereas the MCID of 2.4 to 6.1 points was more indicative of patients who showed a clinically important difference. The MDC estimate of the total NEADL was within the range of the corresponding MCID. It was suggested that both the MDC and MCID be considered in clinical decision making when the MDC exceeds the MCID.[57]

In addition to directly comparing the magnitudes of the MDC and MCID estimates, we computed the percentage of scale width, dividing the MDC or MCID estimates by the maximum scale score. The MDC_{90} percentage of scale width was 8.6% for total NEADL score. Intervals of the 0.2SD- and anchor-based approaches of the MCID percentages of scale width were 4.2% to 10.6% for total NEADL score. Revicki et al[18] recommended that the MCID is best estimated primarily from anchor-based approaches. Distribution-based methods can be used to support MCID estimates

Table 2: Scores and Responsiveness Index of the NEADL Scale in Participants Who Improved After Interventions

Scale	Score (points)	SRM (95% CI)
NEADL total score		
Pretreatment	28.1±12.1	NA
Posttreatment	33.6±12.2	NA
Posttreatment–pretreatment	5.5±4.2	1.3 (1.13–1.53)

NOTE. N=78. Values expressed as mean ± SD unless otherwise noted.
Abbreviation: NA, not applicable.

Table 3: MDC and MCID Estimates of the Total NEADL Scale Score

Scale	SD*	ICC	Standard Error of Measurement	MDC_{90}	% Scale Width of the MDC_{90}	MCID Estimates			
						0.2 SD	% Scale Width of 0.2SD	Anchor-Based MCID	% Scale Width of the Anchor-Based MCID
NEADL total score	12.1	.97	2.1	4.9	8.6	2.4	4.2	6.1	10.6

Abbreviation: ICC, intraclass correlation coefficient.
*Square root of $[(SD_{pretreatment})^2+(SD_{posttreatment})^2/2]$.

from anchor-based approaches.[18] The MCID for total NEADL scale score in this analysis was about 10% to 15% of the total scale score, similar to that of other instruments used in patients with chronic stroke.[37,58] It should be noted that using 10% to 15% of the total scale score range as an external criterion of the anchor-based MCID of the NEADL is based on arbitrary standards of the previous MCID research.

Moreover, reporting the percentages of patients who met the MDC and MCID requirements provided more insightful interpretations than considering only overall mean change scores.[17,44] To determine whether the treatment really is effective for an individual patient, it is important to look at the data to determine how many subjects reach or exceed the values of the MDC and MCID. Therefore, the percentage of individuals in a group that achieve the MDC and MCID can be considered a significant benchmark or an index of responsiveness to evaluate an intervention's effectiveness.[14] We calculated percentages of study participants exceeding the MDC and MCID for a comparison. Our results showed that 50.0% of patients achieved a degree of improvement beyond measurement error for the total NEADL. Nearly three-quarters of patients exceeded the distribution-based MCID, and approximately one-third achieved the anchor-based MCID estimates of the total NEADL scale. These values show that some patients in this sample considered themselves to have become "better," and these data are more informative for evaluating the effects of treatment than overall mean changes.

Study Limitations

Our findings have some limitations that warrant consideration. Responsiveness, MDC, and MCID may vary across populations and treatments and must be shown and documented for the particular study group involving the NEADL scale. Part of the present data overlapped with data from patients with improved NEADL performance in our previous studies[13,26-28] because this study is a secondary analysis for estimating treatment-related threshold values of clinically important change in patients with an improved NEADL total score after any of the specific interventions involved in the study. Our findings show that some patients receiving rehabilitation treatment perceived more or less benefits from the interventions. Future study with a larger sample size may focus on analyzing changes after specific interventions. Additionally, we examined an independent group of 19 patients with chronic stroke for calculating the test-retest reliability of the NEADL. This small sample may not be representative of a broad enough range of patients with stroke, which may limit generalization. Further research based on a larger sample is needed to verify the finding. Moreover, we analyzed only patients with no cognitive disturbances with improved NEADL scores after rehabilitation therapy. Therefore, results may not be generalized to patients with stroke with more cognitive impairments and those who deteriorated in NEADL performance after intervention. The results of responsiveness were specific for patients who rated their NEADL performance as improved, and the major weakness of this approach is that it does not involve a comparison of change in the measure between patients who report improvement in the NEADL and those who do not, which certainly would be overestimated. Future research may use larger samples to allow comparison of improved versus deteriorated patients.

Finally, the treatment period was limited to 3 to 4 weeks; however, rehabilitation studies may require an extended period to achieve optimal change in ADL function. Follow-up assessments of a longer intervention period are warranted in future research.

Table 4: Patients Who Met Criteria for the MDC and MCID on the NEADL Scale

Scale	Those Exceeding the MDC_{90}	Those Exceeding the 0.2SD	Those Exceeding the Anchor-Based MCID
NEADL total score	39 (50)	57 (73.1)	25 (32.1)

NOTE. Values expressed as N (%).

CONCLUSIONS

Findings of this study indicate that total NEADL score is a responsive measure based on the responsiveness index and percentages of patients who met both the MDC and MCID criteria. A clinician can be 90% confident that a change of 4.9 points on total NEADL score in individual patients with stroke indicates a true change. If mean change scores within a stroke group are in a range of 2.4 to 6.1 points on the total NEADL, a clinically important change may have occurred after treatment. Our findings are based on patients with improved NEADL performance after they received specific interventions. Further research with larger samples is needed to validate the findings.

References

1. Lloyd-Jones D, Adams RJ, Brown TM, et al. Heart disease and stroke statistics—2010 update: a report from the American Heart Association. Circulation 2010;121:e46-215.
2. Pettersen R, Dahl T, Wyller TB. Prediction of long-term functional outcome after stroke rehabilitation. Clin Rehabil 2002;16:149-59.
3. Chong DK. Measurement of instrumental activities of daily living in stroke. Stroke 1995;26:1119-22.
4. Lindeboom R, Vermeulen M, Holman R, De Haan RJ. Activities of daily living instruments: optimizing scales for neurologic assessments. Neurology 2003;60:738-42.
5. Ward G, Jagger C, Harper W. A review of instrumental ADL assessments for use with elderly people. Rev Clin Gerontol 1998;8:65-71.

6. Salter KL, Teasell RW, Foley NC, Jutai JW. Outcome assessment in randomized controlled trials of stroke rehabilitation. Am J Phys Med Rehabil 2007;86:1007-12.
7. Wade DT, Halligan P. New wine in old bottles: the WHO ICF as an explanatory model of human behaviour. Clin Rehabil 2003;17: 349-54.
8. Sveen U, Thommessen B, Bautz-Holter E, Wyller TB, Laake K. Well-being and instrumental activities of daily living after stroke. Clin Rehabil 2004;18:267-74.
9. Gladman JR, Lincoln NB, Adams SA. Use of the extended ADL scale with stroke patients. Age Ageing 1993;22:419-24.
10. Gompertz P, Pound P, Ebrahim S. Validity of the extended activities of daily living scale. Clin Rehabil 1994;8:275-80.
11. Hsueh IP, Huang SL, Chen MH, Jush SD, Hsieh CL. Evaluation of stroke patients with the extended activities of daily living scale in Taiwan. Disabil Rehabil 2000;22:495-500.
12. Nouri FM, Lincoln NB. An extended activities of daily living scale for stroke patients. Clin Rehabil 1987;1:301-5.
13. Wu CY, Chuang LL, Lin KC, Horng YS. Responsiveness and validity of two outcome measures of instrumental activities of daily living in stroke survivors receiving rehabilitative therapies. Clin Rehabil 2011;25:175-83.
14. Portney LGW. MP foundations of clinical research: applications to practice. 3rd ed. Upper Saddle River: Pearson/Prentice Hall; 2009.
15. Wright JG. The minimal important difference: who's to say what is important? J Clin Epidemiol 1996;49:1221-2.
16. Liang MH, Lew RA, Stucki G, Fortin PR, Daltroy L. Measuring clinically important changes with patient-oriented questionnaires. Med Care 2002;40(Suppl 4):II45-51.
17. Hays RD, Woolley JM. The concept of clinically meaningful difference in health-related quality-of-life research. How meaningful is it? Pharmacoeconomics 2000;18:419-23.
18. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. J Clin Epidemiol 2008;61: 102-9.
19. Guyatt GH, Osoba D, Wu AW, Wyrwich KW, Norman GR. Methods to explain the clinical significance of health status measures. Mayo Clin Proc 2002;77:371-83.
20. Revicki DA, Cella D, Hays RD, Sloan JA, Lenderking WR, Aaronson NK. Responsiveness and minimal important differences for patient reported outcomes. Health Qual Life Outcomes 2006; 4:70.
21. Wyrwich KW, Tierney WM, Wolinsky FD. Further evidence supporting an SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. J Clin Epidemiol 1999;52:861-73.
22. Stratford PW, Binkley FM, Riddle DL. Health status measures: strategies and analytic methods for assessing change scores. Phys Ther 1996;76:1109-23.
23. Lang CE, Edwards DF, Birkenmeier RL, Dromerick AW. Estimating minimal clinically important differences of upper-extremity measures early after stroke. Arch Phys Med Rehabil 2008;89:1693-700.
24. Dobkin BH. Progressive staging of pilot studies to improve phase III trials for motor interventions. Neurorehabil Neural Repair 2009;23:197-206.
25. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. J Clin Epidemiol 2000;53:459-68.
26. Lin KC, Chang YF, Wu CY, Chen YA. 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;23:441-8.
27. Lin KC, Wu CY, Liu JS, Chen YT, Hsu CJ. Constraint-induced therapy versus dose-matched control intervention to improve motor ability, basic/extended daily functions, and quality of life in stroke. Neurorehabil Neural Repair 2009;23:160-5.
28. Wu CY, Chuang LL, Lin KC, Chen HC, Tsay PK. Randomized trial of distributed constraint-induced therapy versus bilateral arm training for the rehabilitation of upper-limb motor control and function after stroke. Neurorehabil Neural Repair 2011;25: 130-9.
29. Brunnstrom S. Movement therapy in hemiplegia. New York: Harper & Row; 1970.
30. Bohannon RW, Smith MB. Interrater reliability of a Modified Ashworth Scale of muscle spasticity. Phys Ther 1987;67:206-7.
31. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12:189-98.
32. Wu CY, Chen CL, Tsai WC, Lin KC, Chou SH. A randomized controlled trial of modified constraint-induced movement therapy for elderly stroke survivors: changes in motor impairment, daily functioning, and quality of life. Arch Phys Med Rehabil 2007;88: 273-8.
33. Duncan PW, Bode RK, Min Lai S, Perera S. Rasch analysis of a new stroke-specific outcome scale: the Stroke Impact Scale. Arch Phys Med Rehabil 2003;84:950-63.
34. Duncan PW, Wallace D, Lai SM, Johnson D, Embretson S, Laster LJ. The Stroke Impact Scale version 2.0. Evaluation of reliability, validity, and sensitivity to change. Stroke 1999;30:2131-40.
35. Duncan PW, Lai SM, Tyler D, Perera S, Reker DM, Studenski S. Evaluation of proxy responses to the Stroke Impact Scale. Stroke 2002;33:2593-9.
36. Lin KC, Fu T, Wu CY, Hsieh YW, Chen CL, Lee PC. Psychometric comparisons of the Stroke Impact Scale 3.0 and Stroke-Specific Quality of Life Scale. Qual Life Res 2010;19:435-43.
37. van der Lee JH, Wagenaar RC, Lankhorst GJ, Vogelaar TW, Deville WL, Bouter LM. Forced use of the upper extremity in chronic stroke patients: results from a single-blind randomized clinical trial. Stroke 1999;30:2369-75.
38. Iyer LV, Haley SM, Watkins MP, Dumas HM. Establishing minimal clinically important differences for scores on the pediatric evaluation of disability inventory for inpatient rehabilitation. Phys Ther 2003;83:888-98.
39. Hagg O, Fritzell P, Nordwall A. The clinical importance of changes in outcome scores after treatment for chronic low back pain. Eur Spine J 2003;12:12-20.
40. Lin KC, Fu T, Wu CY, et al. Minimal detectable change and clinically important difference of the Stroke Impact Scale in stroke patients. Neurorehabil Neural Repair 2010;24:486-92.
41. Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. Med Care 1990;28: 632-42.
42. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale: Lawrence Erlbaum Assoc; 1988.
43. Efron B, Tibshirani RJ. An introduction to the bootstrap. New York: Chapman & Hall; 1993.
44. Schmitt JS, Di Fabio RP. Reliable change and minimum important difference (MID) proportions facilitated group responsiveness comparisons using individual threshold criteria. J Clin Epidemiol 2004;57:1008-18.
45. de Vet HC, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. J Clin Epidemiol 2006;59: 1033-9.
46. Wagner JM, Rhodes JA, Patten C. Reproducibility and minimal detectable change of three-dimensional kinematic analysis of reaching tasks in people with hemiparesis after stroke. Phys Ther 2008;88:652-63.

47. Haley SM, Fragala-Pinkham MA. Interpreting change scores of tests and measures used in physical therapy. Phys Ther 2006;86: 735-43.
48. Harwood RH, Ebrahim S. The validity, reliability and responsiveness of the Nottingham Extended Activities of Daily Living scale in patients undergoing total hip replacement. Disabil Rehabil 2002;24:371-7.
49. Sahin F, Yilmaz F, Ozmaden A, Kotevoglu N, Sahin T, Kuran B. Reliability and validity of the Turkish version of the Nottingham Extended Activities of Daily Living scale. Aging Clin Exp Res 2008;20:400-5.
50. Eton DT, Cella D, Yost KJ, et al. A combination of distribution- and anchor-based approaches determined minimally important differences (MIDs) for four endpoints in a breast cancer scale. J Clin Epidemiol 2004;57:898-910.
51. Hawkes WG, Williams GR, Zimmerman S, et al. A clinically meaningful difference was generated for a performance measure of recovery from hip fracture. J Clin Epidemiol 2004;57:1019-24.
52. Samsa G, Edelman D, Rothman ML, Williams GR, Lipscomb J, Matchar D. Determining clinically important differences in health status measures: a general approach with illustration to the Health Utilities Index Mark II. Pharmacoeconomics 1999; 15:141-55.
53. Stratford PW, Binkley JM, Riddle DL, Guyatt GH. Sensitivity to change of the Roland-Morris Back Pain Questionnaire: part 1. Phys Ther 1998;78:1186-96.
54. Guyatt G, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. J Chronic Dis 1987;40:171-8.
55. Harwood RH, Ebrahim S. A comparison of the responsiveness of the Nottingham Extended Activities of Daily Living scale, London handicap scale and SF-36. Disabil Rehabil 2000;22: 786-93.
56. Steffen T, Seney M. Test-retest reliability and minimal detectable change on balance and ambulation tests, the 36-Item Short-Form Health Survey, and the Unified Parkinson Disease Rating Scale in people with parkinsonism. Phys Ther 2008;88: 733-46.
57. Resnik L, Dobrykowski E. Outcomes measurement for patients with low back pain. Orthop Nurs 2005;24:14-24.
58. Kwakkel G, Meskers CG, van Wegen EE, et al. Impact of early applied upper limb stimulation: the EXPLICIT-stroke programme design. BMC Neurol 2008;8:49.

Suppliers

a. R software, Available at: http://www.r-project.org/.
b. SPSS Inc, 233 S Wacker Dr, 11th Fl, Chicago, IL 60606.