CrossMark

Validity, responsiveness, and minimal clinically important difference of EQ-5D-5L in stroke patients undergoing rehabilitation

Poyu Chen[1] · Keh-Chung Lin[2,3] · Rong-Jiuan Liing[1] · Ching-Yi Wu[1,4] ·
Chia-Ling Chen[5,6] · Ku-Chou Chang[7]

Abstract

Purpose To examine the criterion validity, responsiveness, and minimal clinically important difference (MCID) of the EuroQoL 5-Dimensions Questionnaire (EQ-5D-5L) and visual analog scale (EQ-VAS) in people receiving rehabilitation after stroke.

Methods The EQ-5D-5L, along with four criterion measures—the Medical Research Council scales for muscle strength, the Fugl–Meyer assessment, the functional independence measure, and the Stroke Impact Scale—was administered to 65 patients with stroke before and after 3- to 4-week therapy. Criterion validity was estimated using the Spearman correlation coefficient. Responsiveness was analyzed by the effect size, standardized response mean (SRM), and criterion responsiveness. The MCID was determined by anchor-based and distribution-based approaches. The percentage of patients exceeding the MCID was also reported.

Results Concurrent validity of the EQ-Index was better compared with the EQ-VAS. The EQ-Index has better power for predicting the rehabilitation outcome in the activities of daily living than other motor-related outcome measures. The EQ-Index was moderately responsive to change (SRM = 0.63), whereas the EQ-VAS was only mildly responsive to change. The MCID estimation of the EQ-Index (the percentage of patients exceeding the MCID) was 0.10 (33.8 %) and 0.10 (33.8 %) based on the anchor-based and distribution-based approaches, respectively, and the estimation of EQ-VAS was 8.61 (41.5 %) and 10.82 (32.3 %).

Conclusions The EQ-Index has shown reasonable concurrent validity, limited predictive validity, and acceptable responsiveness for detecting the health-related quality of life in stroke patients undergoing rehabilitation, but not for EQ-VAS. Future research considering different recovery stages after stroke is warranted to validate these estimations.

Keywords Stroke · Quality of life · Rehabilitation · Validity · MCID · EQ-5D-5L

Poyu Chen and Keh-Chung Lin have contributed equally to this article.

✉ Ching-Yi Wu
cywu@mail.cgu.edu.tw

[1] Department of Occupational Therapy and Graduate Institute of Behavioral Sciences, College of Medicine, Chang Gung University, 259 Wenhua 1st Road, Taoyuan, Taiwan

[2] School of Occupational Therapy, College of Medicine, National Taiwan University, Taipei, Taiwan

[3] Division of Occupational Therapy, Department of Physical Medicine and Rehabilitation, National Taiwan University Hospital, Taipei, Taiwan

[4] Healthy Aging Research Center, Chang Gung University, Taoyuan, Taiwan

[5] Department of Physical Medicine and Rehabilitation, Chang Gung Memorial Hospital, Taoyuan, Taiwan

[6] Graduate Institute of Early Intervention, College of Medicine, Chang Gung University, Taoyuan, Taiwan

[7] Department of Neurology, Chang Gung Memorial Hospital, Kaohsiung, Taiwan

Abbreviations

ES	Effect size
FIM	Functional independence measure
FMA	Fugl–Meyer assessment
HRQoL	Health-related quality of life
MCID	Minimal clinically important difference
MMSE	Mini-Mental State Examination

MRC	Medical Research Council scales for muscle strength
SIS	Stroke Impact Scale
SRM	Standardized response mean
UE	Upper extremity
VAS	Visual analog scale

Introduction

Stroke is one of the leading causes of serious lifetime disabilities. With possible chronic deficits of both functional difficulties and activity limitations [1], the patient's health-related quality of life (HRQoL) is deteriorated [2, 3]. Patient-centered HRQoL [4] has recently been recognized as one of the important outcome measures in stroke rehabilitation [5]. Obtaining information about patient-centered outcome HRQoL in stroke survivors is necessary for clinical practice and research areas to measure baseline data [3, 6] and set interventional goals for monitoring the success of interventions [7–9].

The EuroQoL 5-Dimensions Questionnaire (EQ-5D) is a generic HRQoL measurement with evidence of good reliability and validity in various disease populations [10–13], including stroke [14–18]. The EQ-5D contains the self-reported health state profile of five dimensions (mobility, self-care, usual activity, pain/discomfort, and depression/anxiety) and a visual analog scale (EQ-VAS) [19]. The health status measured by the five dimensions can be converted to a single utility value (EQ-Index score) to inform economic evaluations of health care interventions. The EQ-VAS derives information from a vertical 0- to 100-point VAS for rating overall subjective health status. Given the brevity and simplicity of EQ-5D, stroke patients were more able to complete the questionnaire without missing data compared with the Short-Form 36 (66 vs. 60 %, $p < 0.001$) [16].

The classic three-level version of the EQ-5D, the EQ-5D-3L, has been investigated in stroke trials and showed good psychometric properties, including stable test–retest reliability [14, 18], acceptable construct validity, and concurrent validity [15–18]. A five-level version of the EQ-5D (EQ-5D-5L) was recently developed by the EuroQoL Group to improve the sensitivity and psychometric properties [20]. To our knowledge, only three studies have adopted the EQ-5D-5L in stroke populations, demonstrating a reduced ceiling effect, increased discriminatory power [21], and better construct validity [22], but less responsive to change [23] compared with the EQ-5D-3L. Nevertheless, little is known about the criterion validity of the EQ-5D-5L. Moreover, the longitudinal validity of

changes (i.e., the responsiveness to changes) in stroke populations was inconsistent regardless of the EQ-5D-3L or EQ-5D-5L. Golicki et al. [23] and Pickard et al. [6] demonstrated a moderate to large responsiveness, but Hunger et al. [18] showed a rather limited responsiveness. The diverse findings might result from the stroke stage of patients recruited [23]. Patients included within 2 weeks of stroke onset [6, 23] were more responsive to the intervention because of their extreme health conditions than those included with onset beyond 2 weeks [17, 18]. Re-evaluating the responsiveness of the EQ-5D-5L with patients beyond 2 weeks might be necessary.

Moreover, few studies have examined the clinimetric property—the minimal clinically important difference (MCID)—of the EQ-5D-5L. MCID refers to the smallest meaningful change in a score considered clinically important and constitutes a meaningful change in health status that patients perceive as beneficial important changes or as harmful [24, 25]. The importance of the MCID is that a statistical change is not synonymous with a clinically important change, which is important for consideration when patient-reported outcomes are interpreted.

The purpose of the present study was to assess the criterion validity, responsiveness, and MCID of the EQ-5D-5L in stroke patients undergoing rehabilitation. We also sought to re-evaluate the responsiveness of the EQ-5D-5L in stroke patients beyond 2 weeks.

Methods

Patients

The data used in this study were collected between July 2012 and December 2013 from the departments of rehabilitation in five medical centers in Taiwan. Inclusion criteria were (1) no serious cognitive function deficits as defined by a score of more than 21 on the Mini-Mental State Examination (MMSE) [26], (2) no excessive spasticity in the upper extremity (Modified Ashworth Scale score <3) [27], (3) able to follow instructions to complete the questionnaire and perform therapeutic activity, and (4) age 20–80. The institutional review board at each participating site approved the study, and all participants signed a consent form before entry into the study. Eligible participants were randomly assigned to receive an intensive 1.5- to 2-h therapy session, five times weekly, for 3–4 weeks.

Patients answered the self-reported questionnaires by themselves before and after the intervention, including the EQ-5D-5L, EQ-VAS, and Stroke Impact Scale (SIS) 3.0. Three objective assessments were conducted before and after the intervention by the raters, including the Medical Research Council scales for muscle strength (MRC), the

Fugl–Meyer assessment (FMA), and the functional independence measure (FIM). The raters were trained to administer the outcome measures properly.

Outcome measures

EQ-5D

In this study, patients chose five levels of severity (1, no problem; 2, slight problem; 3, moderate problem; 4, severe problem; and 5, unable to/extreme problem) in five dimensions (mobility, self-care, usual activity, pain/discomfort, and depression/anxiety) and rated their overall health status via the EQ-VAS. Thus, responses for the five dimensions could be combined as a five-digit health status profile, from "1–1–1–1–1," indicating no problem at all, to "5–5–5–5–5," indicating an extremely terrible health status in all five dimensions.

Although the three-level version of EQ-5D (EQ-5D-3L) can be converted to a utility index through some country-specific value sets, the EuroQoL Group has not released the official value set for the EQ-5D-5L. In the interim, we obtained the value set by using the EuroQoL Group's crosswalk methodology for converting the EQ-5D-3L to the EQ-5D-5L [28]. Specifically, we adopted the Japanese value set, which is the only available Asian value set in the crosswalk project [29]. Accordingly, a utility value "1" indicates the best health status, whereas "−0.111" suggests the worst health status.

Criterion measures

Medical Research Council scales

The MRC scale is a measure of muscle strength graded from 0 (no strength) to 5 (normal strength against full resistance) [30]. The MRC scale is a frequently used tool in routine clinical and scientific studies [31, 32], and has reasonable reliability and validity for assessing muscle strength [33].

Fugl–Meyer assessment

The FMA is one of the outcome measures for assessing UE motor function in stroke rehabilitation [34]. The FMA assesses the movement and reflexes of the shoulder/elbow/forearm, wrist, and coordination/speed with 33 three-level items (0, cannot perform; 1, performs partially; and 2, performs fully). Larger scores (maximum score is 66) on the FMA indicate the optimal recovery. The FMA has good psychometric properties [35, 36] and reasonable responsiveness [36].

Functional independence measure

The FIM is a widely used measure of functional independence in stroke rehabilitation [37, 38]. The FIM has 18 items for measuring basic activities of daily living (ADL) by the rater. Each item is rated from 1 (complete assistance) to 7 (complete independence) based on the required level of assistance to perform the tasks, with results combined into an overall score (maximum score, 126). The FIM has good reliability, validity, and responsiveness in stroke studies [39–41].

Stroke Impact Scale 3.0

The SIS 3.0, which is one of the most comprehensive stroke-specific HRQoL instruments, assesses various aspects of life function related to health. The SIS 3.0 includes eight domains (59 items): strength, hand function, ADL/instrumental ADL, mobility, communication, emotion, memory/thinking, and social participation. Each item is scored by the patients based on the difficulty they perceived during the last week (1, extremely difficult; 2, very difficult; 3, somewhat difficult; 4, a little difficult; and 5, not difficult at all). The subscales of strength, hand function, ADL/instrumental ADL, and mobility can be further summarized as the physical scores. The SIS 3.0 has good test–retest reliability [42], internal consistency [42], adequate construct validity [42], and appropriate responsiveness compared with other stroke-related HRQoL assessments [43].

In addition, the SIS includes a global rating scale to assess the patient's perceived recovery from stroke, with 0 indicating no recovery and 100 indicating full recovery. We chose the perceived recovery score from stroke as the anchor for the calculation of criterion responsiveness and MCID because the score directly reflects the patient's perspective, which studies suggested to use the global rating scale as the anchors [44].

Data analysis

Criterion validity

Criterion validity of the EQ-5D-5L, including concurrent and predictive validity, was examined using the Spearman rank correlation coefficient (ρ). To assess the concurrent validity of the EQ-Index, EQ-VAS, and the response level of each dimension, pre-intervention and post-intervention scores were correlated with their respective pre-intervention and post-intervention scores on the criterion measures. To assess the predictive validity of EQ-5D-5L, these scores at pre-intervention were correlated with criterion measures at post-intervention. The strength of the relationship was

considered low ($\rho \leq 0.25$), fair ($\rho = 0.25$–0.5), good ($\rho = 0.5$–0.75), and excellent ($\rho > 0.75$).

Because of the multidimensional concept of the EQ-5D, we used different gold standards as criterion measures to cover the domains of physical function and functional independence contained in the EQ-5D. The MRC scales and FMA for assessing muscle strength and UE function, respectively, were adopted as criterion measures because previous studies demonstrated that physical function was negatively related to the HRQoL [45, 46]. Functional independence is also one of the major components of HRQoL in stroke survivors [47], and so we adopted the FIM as another criterion measure. Furthermore, because validating whether the brief EQ-5D-5L assesses similar concepts as those evaluated by the comprehensive HRQoL is essential, we included the SIS 3.0 as a criterion measure. If the association strength between outcome measures and criterion measures that measure similar constructs is good, accordingly, there will be good concurrent validity between the EQ-Index, EQ-VAS, and the criterion measures. In contrast, there would be absent to low concurrent validity between the EQ-Index, EQ-VAS, and cognitive and communication domains of the SIS because there is no such descriptive concept in the EQ-5D-5L.

Responsiveness

There is no consensus among researchers regarding which methods are better for calculating responsiveness [48, 49]. In the present study, we used three approaches to examine the responsiveness: effect size (ES), standardized response mean (SRM), and criterion responsiveness. ES is defined as the observed change in scores between pre-intervention and post-intervention divided by the standard deviation (SD) of the baseline (pre-intervention) score. SRM is defined as the change scores between pre-intervention and post-intervention measures divided by the SD of the change scores [48]. For the criterion-based responsiveness, we used the "perceived recovery scores" of the SIS 3.0 as a criterion by calculating the Spearman correlation between the change in the EQ-5D and the change in the perceived recovery scores of the SIS 3.0. According to Cohen criteria [50], the values are categorized as large (>0.8), moderate (0.5–0.8), and small (<0.5).

MCID

It has been suggested that the estimation of MCID should be based on multiple approaches, such as the distribution-based and anchor-based approaches, to triangulate the MCID [24, 25]. The distribution-based method was based on the Cohen ES benchmark, for which Norman et al. [51] suggest 0.5 SD of baseline as a small but important change for the intervention effectiveness. The anchor-based method estimates

the MCID by comparing change scores with an external anchor, which is often based on a patient's global rating of change. In the present study, the perceived recovery score of the SIS 3.0 was taken as the external anchor, corresponding to patients who were defined as having MCID. We used a previous study's suggestion for considering a 10–15 % change as a clinically important change [52].

Results

Of the recruited 70 stroke patients, three patients refused to participate in the post-intervention assessment, and two were excluded from the final analysis because of missing data on the SIS subscales. Table 1 summarizes the demographic and clinical features of the 65 participants. The median time between stroke onset and intervention was 17 (range 0.4–94) months. The response distribution of each dimension of the EQ-5D-5L at pre-intervention and post-intervention is shown in Fig. 1. The most frequently reported problems were self-care, usual activity, and mobility, followed by pain/discomfort and anxiety/depression. At the pre-intervention assessment, 13 respondents (20 %) reported the best health status (1–1–1–1–1), and 20 respondents (30.7 %) reported the best possible health status after intervention. The median EQ-Index was 0.719 (range −0.018 to 1) at the pre-intervention session and 0.813 (range 0.364–1) at the post-intervention session. The median score for EQ-VAS was 65 (range 0–100) at the pre-intervention session and 70 (range 5–100) at the post-intervention session.

Pre-intervention and post-intervention concurrent validity

Table 2 reports the Spearman correlation coefficients for the EQ-Index, EQ-VAS, the response levels of each dimension, and the criterion measures related to physical function (i.e.,

Table 1 Demographic and clinical characteristics of the participants ($n = 65$)

Gender	
Male/female	48/17
Age (years)	52.8 ± 11.6
Side of stroke	
Left/right	32/33
Months (median) since stroke onset (range)	19.7 (0.4–94)
Median Brunnstrom stage of proximal/distal UE (range)	3 (1–6)/3 (1–5)
Mini-Mental State Examination Scores	27.3 ± 2.5

Continuous variables are expressed as mean ± SD or median (range) and categoric variables as number

UE upper extremity

MRC and FMA), ADL (i.e., FIM), and comprehensive HRQoL (i.e., SIS 3.0) at pre-intervention and post-intervention sessions. The correlations were generally higher at the post-intervention session than at the pre-intervention session. The EQ-Index demonstrated fair to good concurrent validity with the FIM, SIS-ADL, SIS mobility, and SIS physical scores ($\rho = 0.255$–0.703, $P < 0.05$), whereas the EQ-VAS only showed low to fair concurrent validity with the FIM, SIS mobility, and SIS physical scores ($\rho = 0.249$–0.345, $P < 0.05$). As predicted, associations were absent between the outcome measures and the criterion

measures without concept overlapping (i.e., SIS memory/thinking and SIS communication).

The individual dimensions of "mobility" and "self-care" showed fair to good concurrent validity with the physical function criterion measures (i.e., SIS strength, SIS mobility, and SIS physical scores) and the ADL criterion measures (i.e., FIM and SIS-ADL; $\rho = -0.249$ to -0.771, $P < 0.05$). The "usual activity" dimension showed fair concurrent validity with the FIM, SIS-ADL, SIS mobility, and SIS physical scores, and "pain/discomfort" and "anxiety/depression" showed fair concurrent validity with
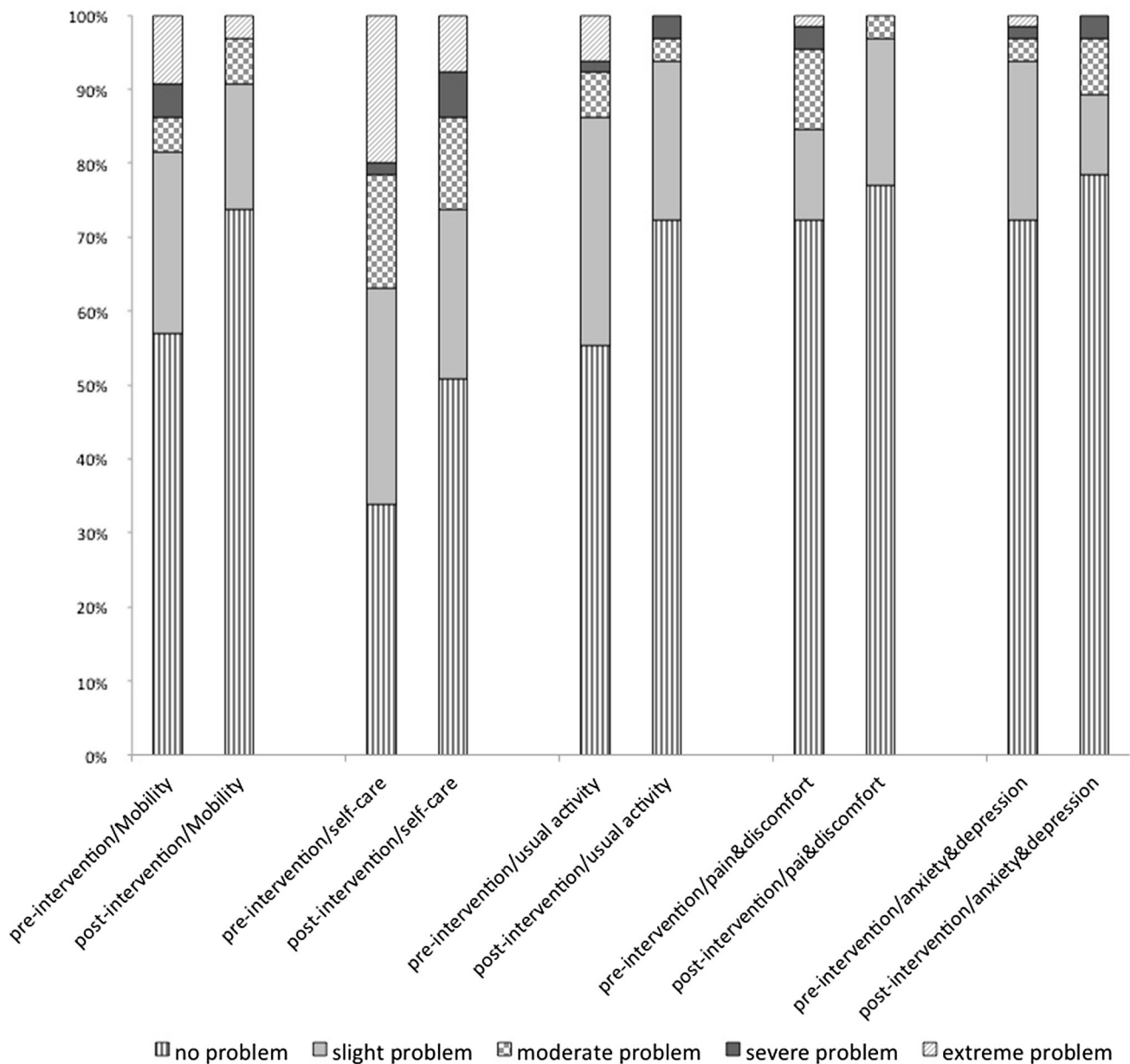

Fig. 1 Responses to each dimension of the EQ-5D-5L at pre-intervention and post-intervention assessments

Table 2 Concurrent validity of EQ-5D-5L at pre-intervention and post-intervention stages

	Muscle strength MRC Pre	MRC Post	UE function FMA Pre	FMA Post	ADL FIM Pre	FIM Post	HRQoL Strength Pre	Strength Post	Memory Pre	Memory Post	Emotion Pre	Emotion Post	Communication Pre	Communication Post
EQ-Index	0.09	−0.04	0.00	0.10	0.255*	0.703**	0.22	0.431**	0.14	0.19	0.05	0.269*	0.07	0.07
EQ-VAS	−0.11	−0.10	0.01	0.03	0.08	0.249*	0.07	0.330**	0.18	0.01	0.14	0.14	−0.09	−0.08
EQ-5D dimension														
Mobility	0.04	0.13	0.04	0.14	−0.249*	−0.559**	−0.14	−0.349**	−0.01	−0.09	−0.08	−0.18	−0.06	−0.10
Self-care	0.13	−0.10	0.11	−0.17	−0.15	−0.771**	−0.02	−0.486**	−0.11	−0.13	−0.03	−0.02	−0.14	−0.02
Usual activity	−0.17	0.03	−0.19	−0.09	−0.14	−0.246*	−0.16	−0.20	−0.22	−0.07	0.02	−0.15	0.00	−0.12
Pain/discomfort	−0.23	0.07	−0.14	−0.04	−0.19	−0.22	−0.18	−0.08	−0.13	0.03	−0.04	−0.412**	−0.02	0.12
Anxiety/depression	−0.06	−0.19	−0.04	−0.355**	0.00	−0.09	−0.14	−0.258*	−0.13	−0.22	0.10	−0.298*	0.08	−0.07

	HRQoL ADL Pre	ADL Post	Mobility Pre	Mobility Post	Hand function Pre	Hand function Post	Social Pre	Social Post	Physical Pre	Physical Post
EQ-VAS	0.24	0.427**	0.24	0.525**	−0.24	−0.19	0.16	0.327**	0.18	0.311*
EQ-5D dimension										
Mobility	−0.18	−0.426**	−0.281*	−0.587**	0.28	0.24	−0.19	−0.14	−0.13	−0.251*
Self-care	−0.14	−0.414**	−0.17	−0.509**	0.28	0.23	−0.14	−0.348**	−0.09	−0.284*
Usual activity	−0.09	−0.276*	0.00	−0.297*	0.18	−0.07	−0.04	−0.10	−0.01	−0.256*
Pain/discomfort	−0.23	−0.07	−0.19	−0.17	−0.17	0.01	−0.15	−0.09	−0.252*	−0.06
Anxiety/depression	0.00	0.04	−0.07	−0.14	0.04	−0.14	−0.13	−0.314*	−0.05	−0.24

For the EQ-Index and EQ-VAS, higher score indicates better health status. On the other hand, the response level of EQ-5D dimension is otherwise: Higher score indicates more severe problem

Spearman ρ: ** $P < 0.01$ level; * $P < 0.05$

the psychological domain of the SIS (i.e., SIS emotion; $\rho = -0.298$ to -0.412, $P < 0.05$).

Predictive validity

The EQ-Index at the pre-intervention session only showed a fair predictive validity ($\rho = 0.25$, $P < 0.05$) with the SIS-ADL at the post-intervention session (Table 3). With regard to the individual dimension, the mobility dimension showed fair predictive validity ($\rho = -0.27$, $P < 0.05$) with the FIM; the pain/discomfort dimension had fair predictive validity with the four SIS subscales, including strength, emotion, mobility, and physical score ($\rho = -0.27$ to -0.34, $P < 0.05$); the anxiety/depression dimension was able to predict the SIS hand function ($\rho = -0.26$, $P < 0.05$).

Responsiveness

Table 4 reports the responsiveness results of the EQ-Index and EQ-VAS. Although the ES and criterion-based approaches revealed that the EQ-Index indicated small responsiveness (0.40 and 0.46, respectively), the SRM of the EQ-Index indicated moderate responsiveness (0.63). As for the EQ-VAS, limited responsiveness was demonstrated for ES (0.3), SRM (0.34), and criterion-based responsiveness (0.29).

MCID

The MCID estimations, derived from the anchor-based and distribution-based methods, are presented in Table 5. The number of participants whose changes exceeded the MCID of the EQ-Index and EQ-VAS are presented as well. The anchor-based MCID, which was based on 26 participants whose self-rated recovery score of the SIS reached 10–15 %, equated to a change of 0.10 for the EQ-Index and 8.61 for the EQ-VAS. Accordingly, changes in 22 participants (33.8 %) and 27 participants (41.5 %) reached the anchor-based MCID estimation of the EQ-Index and the EQ-VAS, respectively. However, the distribution-based MCID estimation (i.e., 0.5 SD of baseline) was 0.10 for the EQ-Index and 10.82 for the EQ-VAS. The estimation indicated that 22 (33.8 %) and 21 participants (32.3 %) had positive changes that exceeded the distribution-based MCID of the EQ-Index and EQ-VAS, respectively.

Discussion

To our knowledge, this is the first study to investigate the criterion validity and MCID of the EQ-5D-5L in stroke patients receiving rehabilitation. Inconsistent with the

Table 3 Predictive validity of EQ-5D-5L

	Muscle strength MRC	UE function FMA	ADL FIM	HRQoL SIS								
				Strength	Memory	Emotion	Communication	ADL	Mobility	Hand function	Social participation	Physical summary score
EQ-Index	0.02	0.08	0.24	0.04	0.06	0.17	0.01	0.25*	0.02	0.01	0.20	-0.06
EQ-VAS	-0.02	0.11	0.02	0.12	0.17	0.06	-0.03	0.06	0.01	0.09	0.14	0.20
EQ-5D dimension												
Mobility	0.03	-0.03	-0.27*	-0.19	0.02	-0.10	-0.10	-0.10	-0.10	0.06	-0.22	-0.05
Self-care	0.17	0.05	-0.20	0.05	-0.08	-0.06	-0.14	-0.05	0.01	0.16	-0.05	0.15
Usual activity	-0.10	-0.22	-0.14	-0.10	-0.21	0.05	-0.05	0.01	0.11	-0.11	-0.13	-0.08
Pain/discomfort	-0.19	-0.19	-0.14	-0.28*	-0.20	-0.27*	-0.05	-0.11	-0.33*	-0.20	-0.23	-0.34*
Anxiety/depression	-0.04	-0.11	0.04	-0.10	-0.15	-0.20	0.01	-0.02	-0.08	-0.26*	0.09	-0.10

For the EQ-Index and EQ-VAS, higher score indicates better health status. On the other hand, the response level of EQ-5D dimension is otherwise: Higher score indicates more severe problem

Spearman ρ: ** $P < 0.01$ level; * $P < 0.05$

Table 4 Responsiveness statistics

	EQ-Index	EQ-VAS
Responsiveness		
ES	0.40	0.30
SRM	0.63	0.34
Criterion approach	0.46	0.29

Table 5 MCID estimation with anchor-based and distribution-based approaches

	EQ-Index	EQ-VAS
MCID		
Anchor-based: SIS recovery score (10–15 %)	0.10	8.61
Distribution-based: 0.5 SD	0.10	10.82
Those exceeding MCID [n (%)]		
Anchor-based: SIS recovery score (10–15 %)	22 (33.8)	27 (41.5)
Distribution-based: 0.5 SD	22 (33.8)	21 (32.3)

SD standard deviation

results reported by Golicki et al. [23] that suggested moderate responsiveness for both the EQ-Index and EQ-VAS of the EQ-5D-5L, our results showed that only the EQ-Index was moderately responsive to changes based on the SRM value. Also, the MCID values identified in the present study might help clinicians and researchers determine whether change scores indicate clinical improvement after an intervention.

Criterion validity

By examining the correlation between the EQ-5D-5L and criterion measures, the present study showed limited concurrent validity at the pre-intervention session and acceptable concurrent validity at post-intervention (Table 2). Moreover, our preliminary findings indicated the inferior predicting power of the EQ-5D-5L (Table 3). These issues are further discussed in the following section.

Concurrent validity

At the pre-intervention session, the EQ-Index and the EQ-mobility dimension showed fair correlation with FIM. However, no other correlations between the EQ-5D and objective criterion measures at the pre-intervention session were identified. This finding suggested that a discrepancy exists between the patient's self-perception of QoL and the objective evaluation of physical function and functional independence. Furthermore, the lack of correlation between the EQ-5D and subjective SIS at the pre-

intervention session might be related to the questionnaire discrepancy—the subjective SIS tends to evaluate different aspects of patients' impairments for performing designated activities, whereas EQ-5D emphasizes the general performance during usual activities. For example, the SIS-ADL asks whether the patient could cut the food, dress the top part of the body, etc., and EQ-5D asks whether patients have problems in household, work, and leisure activities without focusing on how difficult it is for them to perform subtasks comprising one household, work, or leisure task.

After the 3- to 4-week intervention, correlations were increased between the EQ-5D and criterion measures, including objective and subjective measures. The intervention might facilitate a realistic perception by patients about QoL, and such perception might better agree with objective assessment outcomes in patients' performance of functional independence. In addition, the intervention improved stroke-related impairments that were reflected by the SIS subscales such as SIS-ADL, increasing the consistency of self-rating scores between the general scale of EQ-5D and the more specific and comprehensive scale of SIS.

Specifically, EQ-Index, EQ-mobility, and EQ-self-care had good to excellent concurrent validity with FIM and SIS mobility at the post-intervention session, supporting the important role of functional independence in HRQoL for stroke survivors [47]. Also, the EQ-Index had fair correlation with SIS-ADL, strength, social participation, and emotion subscales at the post-intervention session. The lower correlation between the EQ-Index and SIS emotion might have been caused by the poor measurement properties of the SIS emotion subscale [18, 53]. As hypothesized, there were absent correlations with SIS memory/thinking and SIS communication given the non-overlapping concept in the EQ-5D self-classifier. Compared with the EQ-Index, EQ-VAS only showed low to fair correlation with FIM, SIS strength, and SIS mobility at the post-intervention session, suggesting the poor validity of EQ-VAS.

Although EQ-5D showed fair correlation with SIS strength, neither the EQ-Index nor individual dimensions showed correlation with objective MRC scores. Given that EQ-5D emphasizes the general HRQoL, SIS strength, addressing muscle strength based on the daily activity (i.e., grip of your hand that was mostly affected by your stroke), would show a more similar concept with EQ-5D compared with the fundamental strength assessment by MRC based on individual joint movement (i.e., wrist extension, with 0 as no movement to 5 as a lot of strength). As for the invalid concurrent validity with FMA, we speculated that the FMA assessed the impairment of the motor component so that the patient's subjective HRQoL was not directly related to this impairment assessment. Future study with different criterion measures should be done to clarify these speculations.

Predictive validity

Our preliminary findings of predictive validity showed the EQ-Index at the pre-intervention session could fairly predict the SIS-ADL at the post-intervention session, indicating stroke patients with better perceived health status before the intervention were associated with a more favorable rehabilitation outcome of the ADL function.

This finding implicates that the predictive validity of EQ-5D-5L is better to apply in the ADL rehabilitation outcome measures than the objective motor-related outcome measures. However, EQ-VAS had no predictive power for predicting the rehabilitation outcomes. The general concept of health status in stroke survivors does not seem to be a good index for predicting rehabilitation outcome. Future studies should address this issue.

As for the individual dimension, EQ-mobility could fairly predict the FIM, suggesting a better functional independence outcome in patients who had better mobility at the pre-intervention session. Interestingly, the EQ-pain/discomfort dimension had fair predictive power for predicting post-intervention rehabilitation outcomes of the SIS subscales (i.e., strength, emotion, mobility, and physical scores), suggesting a better rehabilitation outcome in patients who experienced less pain at the pre-intervention session. Pain that occurs after stroke is one of the most common medical complications to adversely affect the course of rehabilitation. Appropriate and timely pain management leads to maximum function in ADL and adequate QoL [54]. Also, the EQ-anxiety/depression domain had fair power for predicting the subjective SIS hand function but not for the objective UE function measure (i.e., FMA). This result was in line with the argument that both psychological and physical impairments are important for HRQoL in stroke survivors [55].

Furthermore, a close look at the raw data of predictive validity suggested there might be differential associations in patients at subacute (beyond 2 weeks and <3 months after stroke onset) and chronic (beyond 3 months after stroke onset) stages. An additional analysis based on subacute and chronic stages showed the predictive validity was better in subacute patients compared with chronic patients. Interestingly, EQ-pain/discomfort and EQ-anxiety/depression in subacute patients might be more important predictive factors of UE function, functional independence level, and SIS subscales ($\rho = -0.49$ to -0.69, $P < 0.05$). The additional analyses demonstrate that the predictive power of the various dimensions of the EQ-5D depends on the patient's post-stroke stage. Further research should address the role of the different dimensions of the EQ-5D and its predictive validity at different post-stroke stages.

Responsiveness

Compared with the moderate to large responsiveness results for the EQ-Index and EQ-VAS in the Golicki et al. [23] study, our results were rather limited: The EQ-Index was moderately responsive to changes based on the SRM index, whereas the EQ-VAS was only mildly responsive to changes. As indicated by Pickard et al. [6], the EQ-5D is highly responsive in patients with extreme health conditions before intervention. The difference in the results obtained by Golicki et al. [23] and the present results might be related to the patients' stage of stroke. Golicki et al. [23] recruited acute patients, whereas the present study recruited subacute and chronic patients. We further performed an additional responsiveness analysis based on subacute and chronic patients. As expected, subacute patients with worse functional status before the intervention were more responsive to the rehabilitation than the chronic patients. However, a conclusion that the subacute stroke patients are more eligible for the EQ-5D-5L as a measure of HRQoL than chronic stroke patients should be exercised cautiously because the findings are based on a subgroup analysis with a small sample size. Future study should further address the responsiveness issue by directly incorporating a larger sample size for the different stroke stages.

MCID

To our knowledge, empirical work has not been performed to assess the MCID of the EQ-5D-5L in stroke populations, although it has been tested in other diseases [56, 57]. On the basis of the anchor-based approach, we found the MCID values were 0.10 for the EQ-Index and 8.61 for the EQ-VAS. The MCID values calculated by the distribution-based method were 0.10 for the EQ-Index and 10.82 for the EQ-VAS. Combining the two approaches for the MCID, participants who achieved values of 0.10 on the EQ-Index and 8.61–10.82 on the EQ-VAS are likely to have a clinically important change.

Moreover, to determine whether the rehabilitation is effective for patients, an examination of how many participants (i.e., the percentages of a group) reach or exceed the values of the MCID instead of focusing on the MCID values alone is necessary [25]. Thus, the percentages of patients who achieve MCID can be considered as another benchmark for evaluating the effectiveness of an intervention. The present results showed that 33.8 and 41.5 % of patients achieved the anchor-based MCID of the EQ-Index and EQ-VAS estimations, respectively, whereas 33.8 and 32.3 % of patients achieved the distribution-based MCID of the EQ-Index and EQ-VAS. Finally, the MCID estimation also depends on the stages of stroke; the MCID estimations of the EQ-Index and EQ-VAS in subacute patients were larger than those in chronic patients.

Study limitations

Two limitations may influence the interpretation of our findings. First, the predictive validity, responsiveness, and MCID may vary across demographic characteristics and interventions. Thus, the results of the present study might not generalize to other populations with different characteristics. Second, only patients with an MMSE of 21 or more were included in the present study; thus, the results may not be generalized to patients with cognitive impairment.

Conclusions

This is the first study to explore the criterion validity and MCID of the EQ-5D-5L in stroke patients receiving rehabilitation. The present study generally showed the reasonable criterion validity and responsiveness of the EQ-Index compared with the EQ-VAS. The results also suggested that the EQ-5D-5L has better power for predicting the rehabilitation outcome in ADL compared with other motor-related outcome measures. The MCID estimation provides clinicians with the EQ-5D-5L benchmark for evaluating the therapeutic effect in stroke patients undergoing rehabilitation.

Finally, the psychometric and clinimetric properties of the EQ-5D-5L seemed to further depend on the patient's stroke stage. Further research with larger samples is required to explore the psychometric and clinimetric properties of the EQ-5D-5L in patients at different stages of stroke recovery.

Compliance with ethical standards

Conflict of interest 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.

Informed consent The institutional review board at each participating site approved the study, and all participants signed a consent form before entry into the study.

References

1. Go, A. S., Mozaffarian, D., Roger, V. L., Benjamin, E. J., Berry, J. D., Borden, W. B., et al. (2013). Heart disease and stroke statistics—2013 update: A report from the American Heart Association. *Circulation, 127*(1), e6–e245. doi:10.1161/CIR.0b013e31828124ad.

2. Lai, S. M., Perera, S., Duncan, P. W., & Bode, R. (2003). Physical and social functioning after stroke: Comparison of the Stroke Impact Scale and Short Form-36. *Stroke, 34*(2), 488–493.

3. Golomb, B. A., Vickrey, B. G., & Hays, R. D. (2001). A review of health-related quality-of-life measures in stroke. *Pharmacoeconomics, 19*(2), 155–185.

4. Schipper, H., Clinch, J. J., & Olweny, C. L. M. (1996). Quality of life studies: Definitions and conceptual issues. In B. Spilker (Ed.), *Quality of life and pharmacoeconomics in clinical trials* (pp. 11–23). Philadelphia, PA: Lippincott-Raven Publishers.

5. Carod-Artal, F. J., & Egido, J. A. (2009). Quality of life after stroke: The importance of a good recovery. *Cerebrovascular Diseases, 27*(Suppl 1), 204–214. doi:10.1159/000200461.

6. Pickard, A. S., Johnson, J. A., & Feeny, D. H. (2005). Responsiveness of generic health-related quality of life measures in stroke. *Quality of Life Research, 14*(1), 207–219.

7. Gray, L. J., Sprigg, N., Bath, P. M., Boysen, G., De Deyn, P. P., Leys, D., et al. (2007). Sex differences in quality of life in stroke survivors: Data from the Tinzaparin in Acute Ischaemic Stroke Trial (TAIST). *Stroke, 38*(11), 2960–2964. doi:10.1161/STROKEAHA.107.488304.

8. Alguren, B., Fridlund, B., Cieza, A., Sunnerhagen, K. S., & Christensson, L. (2012). Factors associated with health-related quality of life after stroke: A 1-year prospective cohort study. *Neurorehabilitation and Neural Repair, 26*(3), 266–274. doi:10.1177/1545968311414204.

9. Lin, K. C., Fu, T., Wu, C. Y., & Hsieh, C. J. (2011). Assessing the stroke-specific quality of life for outcome measurement in stroke rehabilitation: Minimal detectable change and clinically important difference. *Health and Quality of Life Outcomes, 9*, 5. doi:10.1186/1477-7525-9-5.

10. Brazier, J., Jones, N., & Kind, P. (1993). Testing the validity of the Euroqol and comparing it with the SF-36 health survey questionnaire. *Quality of Life Research, 2*, 169–180.

11. Harrison, M. J., Lunt, M., Verstappen, S. M., Watson, K. D., Bansback, N. J., & Symmons, D. P. (2010). Exploring the validity of estimating EQ-5D and SF-6D utility values from the health assessment questionnaire in patients with inflammatory arthritis. *Health and Quality of Life Outcomes, 8*, 21. doi:10.1186/1477-7525-8-21.

12. Linde, L., Sorensen, J., Ostergaard, M., Horslev-Petersen, K., & Hetland, M. L. (2008). Health-related quality of life: Validity, reliability, and responsiveness of SF-36, 15D, EQ-5D RAQoL, and HAQ in patients with rheumatoid arthritis. *Journal of Rheumatology, 35*(8), 1528–1537.

13. Obradovic, M., Lal, A., & Liedgens, H. (2013). Validity and responsiveness of EuroQol-5 dimension (EQ-5D) versus Short Form-6 dimension (SF-6D) questionnaire in chronic pain. *Health and Quality of Life Outcomes, 11*, 110. doi:10.1186/1477-7525-11-110.

14. Dorman, P., Slattery, J., Farrell, B., Dennis, M., & Sandercock, P. (1998). Qualitative comparison of the reliability of health status assessments with the EuroQol and SF-36 questionnaires after stroke. United Kingdom Collaborators in the International Stroke Trial. *Stroke, 29*(1), 63–68.

15. Dorman, P. J., Dennis, M., & Sandercock, P. (1999). How do scores on the EuroQol relate to scores on the SF-36 after stroke? *Stroke, 30*(10), 2146–2151.

16. Dorman, P. J., Slattery, J., Farrell, B., Dennis, M. S., & Sandercock, P. A. (1997). A randomised comparison of the EuroQol and Short Form-36 after stroke. United Kingdom collaborators in the International Stroke Trial. *BMJ, 315*(7106), 461.

17. Dorman, P. J., Waddell, F., Slattery, J., Dennis, M., & Sandercock, P. (1997). Is the EuroQol a valid measure of health-related quality of life after stroke? *Stroke, 28*(10), 1876–1882.

18. Hunger, M., Sabariego, C., Stollenwerk, B., Cieza, A., & Leidl, R. (2012). Validity, reliability and responsiveness of the EQ-5D in German stroke patients undergoing rehabilitation. *Quality of Life Research, 21*(7), 1205–1216. doi:10.1007/s11136-011-0024-3.

19. EuroQol, G. (1990). EuroQol—A new facility for the measurement of health-related quality of life. *Health Policy, 16*(3), 199–208.

20. Herdman, M., Gudex, C., Lloyd, A., Janssen, M., Kind, P., Parkin, D., et al. (2011). Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of Life Research, 20*(10), 1727–1736. doi:10.1007/s11136-011-9903-x.

21. Janssen, M. F., Pickard, A. S., Golicki, D., Gudex, C., Niewada, M., Scalone, L., et al. (2013). Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: A multi-country study. *Quality of Life Research, 22*(7), 1717–1727. doi:10.1007/s11136-012-0322-4.

22. Golicki, D., Niewada, M., Buczek, J., Karlinska, A., Kobayashi, A., Janssen, M. F., et al. (2015). Validity of EQ-5D-5L in stroke. *Quality of Life Research, 24*(4), 845–850. doi:10.1007/s11136-014-0834-1.

23. Golicki, D., Niewada, M., Karlinska, A., Buczek, J., Kobayashi, A., Janssen, M. F., et al. (2014). Comparing responsiveness of the EQ-5D-5L, EQ-5D-3L and EQ VAS in stroke patients. *Quality of Life Research, 24*(6), 1555–1563. doi:10.1007/s11136-014-0873-7.

24. Revicki, D., Hays, R. D., Cella, D., & Sloan, J. (2008). Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *Journal of Clinical Epidemiology, 61*(2), 102–109. doi:10.1016/j.jclinepi.2007.03.012.

25. Hays, R. D., & Woolley, J. M. (2000). The concept of clinically meaningful difference in health-related quality-of-life research. How meaningful is it? *Pharmacoeconomics, 18*(5), 419–423.

26. Crum, R. M., Anthony, J. C., Bassett, S. S., & Folstein, M. F. (1993). Population-based norms for the Mini-Mental State Examination by age and educational level. *JAMA, 269*(18), 2386–2391.

27. Bohannon, R. W., & Smith, M. B. (1987). Interrater reliability of a modified Ashworth scale of muscle spasticity. *Physical Therapy, 67*(2), 206–207.

28. van Hout, B., Janssen, M. F., Feng, Y. S., Kohlmann, T., Busschbach, J., Golicki, D., et al. (2012). Interim scoring for the EQ-5D-5L: Mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value Health, 15*(5), 708–715. doi:10.1016/j.jval.2012.02.008.

29. Tsuchiya, A., Ikeda, S., Ikegami, N., Nishimura, S., Sakai, I., Fukuda, T., et al. (2002). Estimating an EQ-5D population value set: The case of Japan. *Health Economics, 11*(4), 341–353. doi:10.1002/hec.673.

30. Council, M. R. (1976). *Aids to examination of the peripheral nervous system*. London: Her Majesty's Stationary Office.

31. Swash, M., Brown, M. M., & Thakkar, C. (1995). CT muscle imaging and the clinical assessment of neuromuscular disease. *Muscle and Nerve, 18*(7), 708–714. doi:10.1002/mus.880180706.

32. Tzvetanov, P., & Rousseff, R. T. (2003). Median SSEP changes in hemiplegic stroke: Long-term predictive values regarding ADL recovery. *NeuroRehabilitation, 18*(4), 317–324.

33. Paternostro-Sluga, T., Grim-Stieger, M., Posch, M., Schuhfried, O., Vacariu, G., Mittermaier, C., et al. (2008). Reliability and validity of the Medical Research Council (MRC) scale and a modified scale for testing muscle strength in patients with radial palsy. *Journal of Rehabilitation Medicine, 40*(8), 665–671. doi:10.2340/16501977-0235.

34. Fugel-Meyer, A. R., Jaasko, L., Leyman, I., Ollson, S., & Steglind, S. (1975). The poststroke hemiplegic patient. 1. A method for evaluation of physical performance. *Scandinavian Journal of Rehabilitation Medicine, 7*, 13–31.

35. Platz, T., Pinkowski, C., van Wijck, F., Kim, I. H., di Bella, P., & Johnson, G. (2005). 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 multicentre study. *Clinical Rehabilitation, 19*(4), 404–411.

36. Hsieh, Y. W., Wu, C. Y., Lin, K. C., Chang, Y. F., Chen, C. L., & Liu, J. S. (2009). Responsiveness and validity of three outcome measures of motor function after stroke rehabilitation. *Stroke, 40*(4), 1386–1391. doi:10.1161/STROKEAHA.108.530584.

37. Cohen, M. E., & Marino, R. J. (2000). The tools of disability outcomes research functional status measures. *Archives of Physical Medicine and Rehabilitation, 81*(12 Suppl 2), S21–S29.

38. Granger, C. V. (1998). The emerging science of functional assessment: Our tool for outcomes analysis. *Archives of Physical Medicine and Rehabilitation, 79*(3), 235–240.

39. Chau, N., Daler, S., Andre, J. M., & Patris, A. (1994). Inter-rater agreement of two functional independence scales: The Functional Independence Measure (FIM) and a subjective uniform continuous scale. *Disability and Rehabilitation, 16*(2), 63–71.

40. Dodds, T. A., Martin, D. P., Stolov, W. C., & Deyo, R. A. (1993). A validation of the functional independence measurement and its performance among rehabilitation inpatients. *Archives of Physical Medicine and Rehabilitation, 74*(5), 531–536.

41. Wallace, D., Duncan, P. W., & Lai, S. M. (2002). Comparison of the responsiveness of the Barthel Index and the motor component of the Functional Independence Measure in stroke: The impact of using different methods for measuring responsiveness. *Journal of Clinical Epidemiology, 55*(9), 922–928.

42. Carod-Artal, F. J., Coral, L. F., Trizotto, D. S., & Moreira, C. M. (2008). The stroke impact scale 3.0: Evaluation of acceptability, reliability, and validity of the Brazilian version. *Stroke, 39*(9), 2477–2484. doi:10.1161/STROKEAHA.107.513671.

43. Lin, K. C., Fu, T., Wu, C. Y., Hsieh, Y. W., Chen, C. L., & Lee, P. C. (2010). Psychometric comparisons of the Stroke Impact Scale 3.0 and Stroke-Specific Quality of Life Scale. *Quality of Life Research, 19*(3), 435–443. doi:10.1007/s11136-010-9597-5.

44. De Vet, H. C., Terwee, C. B., Mokkink, L. B., & Knol, D. L. (2011). *Measurement in medicine*. Cambridge: Cambridge University Press.

45. Nichols-Larsen, D. S., Clark, P. C., Zeringue, A., Greenspan, A., & Blanton, S. (2005). Factors influencing stroke survivors' quality of life during subacute recovery. *Stroke, 36*(7), 1480–1484. doi:10.1161/01.STR.0000170706.13595.4f.

46. Sidiras, G., Patsaki, I., Dakoutrou, M., Karatzanos, E., Gerovasili, V., Kouvarakos, A., et al. (2012). *Muscle strength assessment of critically ill patients is associated with functional ability and quality of life at hospital discharge*. Paper presented at the 32nd international symposium on intensive care and emergency medicine, Brussels, Belgium.

47. Abubakar, S. A., & Isezuo, S. A. (2012). Health related quality of life of stroke survivors: Experience of a stroke unit. *International Journal of Biomedical Sciences, 8*(3), 183–187.

48. Husted, J. A., Cook, R. J., Farewell, V. T., & Gladman, D. D. (2000). Methods for assessing responsiveness: A critical review and recommendations. *Journal of Clinical Epidemiology, 53*(5), 459–468.

49. Wright, J. G., & Young, N. L. (1997). A comparison of different indices of responsiveness. *Journal of Clinical Epidemiology, 50*(3), 239–246.

50. Cohen, J. (1988). *Statistical power analysis for the behavior sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.

51. Norman, G. R., Sloan, J. A., & Wyrwich, K. W. (2003). Interpretation of changes in health-related quality of life: The remarkable universality of half a standard deviation. *Medical Care, 41*(5), 582–592. doi:10.1097/01.MLR.0000062554.74615.4C.

52. Lin, K. C., Fu, T., Wu, C. Y., Wang, Y. H., Liu, J. S., Hsieh, C. J., et al. (2010). Minimal detectable change and clinically important difference of the Stroke Impact Scale in stroke patients. *Neurorehabilitation and Neural Repair, 24*(5), 486–492. doi:10.1177/1545968309356295.

53. Duncan, P. W., Wallace, D., Lai, S. M., Johnson, D., Embretson, S., & Laster, L. J. (1999). The stroke impact scale version 2.0. Evaluation of reliability, validity, and sensitivity to change. *Stroke, 30*(10), 2131–2140.

54. Zorowitz, R. D., Smout, R. J., Gassaway, J. A., & Horn, S. D. (2005). Usage of pain medications during stroke rehabilitation: The Post-Stroke Rehabilitation Outcomes Project (PSROP). *Topics in Stroke Rehabilitation, 12*(4), 37–49. doi:10.1310/C7MF-VLR0-CKDL-3C44.

55. Ahlsio, B., Britton, M., Murray, V., & Theorell, T. (1984). Disablement and quality of life after stroke. *Stroke, 15*(5), 886–890.

56. Sagberg, L. M., Jakola, A. S., & Solheim, O. (2013). Quality of life assessed with EQ-5D in patients undergoing glioma surgery: What is the responsiveness and minimal clinically important difference? *Quality of Life Research, 23*(5), 1427–1434. doi:10.1007/s11136-013-0593-4.

57. Pickard, A. S., Neary, M. P., & Cella, D. (2007). Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer. *Health and Quality of Life Outcomes, 5*, 70.