

# Stroke

American Stroke  
Association<sup>SM</sup>

JOURNAL OF THE AMERICAN HEART ASSOCIATION

A Division of American  
Heart Association



## **ACTIVLIM-Stroke: A Crosscultural Rasch-Built Scale of Activity Limitations in Patients With Stroke**

Charles Sèbiyo Batcho, Alan Tennant and Jean-Louis Thonnard

*Stroke* 2012, 43:815-823: originally published online January 5, 2012

doi: 10.1161/STROKEAHA.111.638965

Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75214

Copyright © 2012 American Heart Association. All rights reserved. Print ISSN: 0039-2499. Online  
ISSN: 1524-4628

The online version of this article, along with updated information and services, is  
located on the World Wide Web at:

<http://stroke.ahajournals.org/content/43/3/815>

Subscriptions: Information about subscribing to *Stroke* is online at  
<http://stroke.ahajournals.org/subscriptions/>

Permissions: Permissions & Rights Desk, Lippincott Williams & Wilkins, a division of Wolters  
Kluwer Health, 351 West Camden Street, Baltimore, MD 21202-2436. Phone: 410-528-4050. Fax:  
410-528-8550. E-mail:  
[journalpermissions@lww.com](mailto:journalpermissions@lww.com)

Reprints: Information about reprints can be found online at  
<http://www.lww.com/reprints>

# ACTIVLIM-Stroke: A Crosscultural Rasch-Built Scale of Activity Limitations in Patients With Stroke

Charles Sèbiyo Batcho, PT; Alan Tennant, PhD; Jean-Louis Thonnard, PhD

**Background and Purpose**—This study describes the development of a Rasch-built scale measuring activity limitations in stroke patients, named ACTIVLIM-Stroke.

**Method**—This new Rasch-built measure was constructed based on stroke patients' perceptions of difficulty in performing daily activities. Patients were recruited from inpatient and outpatient rehabilitation departments in Belgium and Benin. A 73-item questionnaire was completed by 204 participants. A random subsample of 83 subjects was given the questionnaire a second time. Data were analyzed using RUMM2030 software.

**Results**—After successive Rasch analyses, the ACTIVLIM-Stroke questionnaire, a unidimensional and linear 20-item measure of activity limitations, was constructed. All 20 items fulfilled Rasch requirements (overall and individual item fit, category discrimination, invariance, local response independence, and nonredundancy in item difficulty). This simple patient-based scale encompasses a large range of activities related to self-care, transfer, mobility, manual ability, and balance. The ACTIVLIM-Stroke questionnaire exhibited high internal validity, excellent internal consistency, and good crosscultural validity. The test–retest reliability of item difficulty hierarchy (intraclass correlation coefficient=0.99) and patient location (intraclass correlation coefficient=0.92) were both excellent. Furthermore, it showed good external construct validity using correlations with the Functional Independence Measure motor and the Barthel Index and a higher discriminating capacity than either of these widely used indices.

**Conclusions**—The ACTIVLIM-Stroke questionnaire has good psychometric qualities and provides accurate measures of activity limitations in patients with stroke. It is recommended for evaluating clinical and research interventions in patients with stroke, because it provides a higher discrimination and might be more sensitive to change. (*Stroke*. 2012;43:815-823.)

**Key Words:** activity limitations ■ measurement ■ Rasch model ■ scales ■ stroke recovery

The increasing use of patient-reported outcomes has encouraged the development of several questionnaires to evaluate individual functioning based on the International Classification of Functioning, Disability and Health framework. The International Classification of Functioning, Disability and Health, a classification of health and health-related domains, describes individual functioning in 3 domains: (1) body functions and anatomic structures; (2) activity; and (3) participation.<sup>1</sup> Problems in each domain are, respectively, impairments, activity limitations, and participation restrictions. In the activity domain, the International Classification of Functioning, Disability and Health defines activity limitations as the difficulties a person might have in executing daily activities. Activity limitation is a behavior that is a combination of motor function, compensatory behavior of individuals, and personal (eg, age, lifestyle, motivation) and environmental (eg, architectural characteristics, ground type) factors. Therefore, limitation of activity cannot be measured directly

but can be inferred from an individual's perception of the difficulty of performing activities.

Despite the wide range of instruments currently available, only few are identified as meeting rigorous, evidence-based modern psychometric standards for a rating scale.<sup>2,3</sup> Earlier tools were developed following traditional standards of measurement science, concentrating on key aspects such as reliability and validity. More recently, a greater emphasis has been given to more powerful diagnostic approaches, which examine a wider range of attributes such as response category functioning and differential item functioning (DIF). Among the new approaches, the Rasch measurement model is the most commonly used.<sup>4</sup> Over the last 15 years, Rasch analysis has been widely used in health science.<sup>5–9</sup> Some Rasch-built scales such as the ABILHAND scale (a measure of manual ability),<sup>10</sup> the ABILOCO scale (a measure of locomotion ability),<sup>11</sup> and the EG Motor Index (a measure of mobility)<sup>12</sup> assess the functioning of patients with stroke. The Stroke

Received September 14, 2011; final revision received November 15, 2011; accepted November 29, 2011.

From the Institut of Neuroscience (C.S.B., J.-L.T.), Université catholique de Louvain, Brussels, Belgium; and the Department of Rehabilitation Medicine (A.T.), Faculty of Medicine and Health, University of Leeds, Leeds, UK.

Correspondence to Jean-Louis Thonnard, Institut of Neuroscience, Université catholique de Louvain, Brussels, Belgium, 53 avenue Mounier, COSY-B1.53.04, B-1200 Brussels, Belgium. E-mail jean-louis.thonnard@uclouvain.be

© 2012 American Heart Association, Inc.

Stroke is available at <http://stroke.ahajournals.org>

DOI: 10.1161/STROKEAHA.111.638965

Impact Scale, developed in 1999,<sup>13</sup> has been refined in 2003 using Rasch analysis.<sup>14</sup> However, most stroke-specific Rasch-built scales evaluate only some aspects of activity limitations as defined by the International Classification of Functioning, Disability and Health.<sup>1</sup> ABILHAND, ABILOCO, and the EG Motor Index assess specific aspects of activity limitations and are important in trials designed to evaluate the effect of particular interventions focused on a specific skill, for example manual ability or mobility. However, they cannot be used as a comprehensive measure of whole activity limitation. The Stroke Impact Scale is a broad assessment tool of physical function and not a measure specific to activity limitations, because it includes items from the body function domain (eg, “bladder and bowel control”), activity domain (eg, “move from a bed to a chair,” “bathe yourself”), and participation domain (eg, “go shopping”).

A full assessment of a stroke patient’s functional ability should consider the broad range of activity limitations as a whole variable, as did Vandervelde et al<sup>15</sup> in the ACTIVLIM scale, a Rasch-built measure of activity limitations in children and adults with neuromuscular disorders. However, the use of the ACTIVLIM questionnaire for patients with stroke would require a validation in that diagnosis. Moreover, given the increase in multicenter international studies, so facilitating comparison of the outcome of across different populations, crosscultural validated outcome measures are also required. Consequently, this current study aimed to calibrate and validate the ACTIVLIM questionnaire for patients with stroke from Europe (Belgium) and Africa (Benin).

## Methods

Data were collected from French language-speaking communities in European (Belgium) and African (Benin) patients with stroke. This study was approved by the Ethics Committee of the Université catholique de Louvain in Belgium and the local ethics committees of the participating caregiver centers and hospitals in Benin. Patients signed an informed consent form before being included.

## Subjects

Patients were recruited from rehabilitation departments, including patients with stroke currently undergoing rehabilitation and those discharged. Patients who had been discharged were identified from patient registers at the recruitment centers. The study was restricted to patients presenting no major cognitive deficit that could potentially prevent them from completing a self-report questionnaire ( $\geq 24$  of 30 on the Mini-Mental State Examination).<sup>16,17</sup>

## Patient Assessment and Outcome Measures

In addition to demographic and clinical data, assessment included the ACTIVLIM-Stroke questionnaire, the Functional Independence Measure (FIM),<sup>18</sup> and the Barthel Index (BI).<sup>19</sup>

## ACTIVLIM-Stroke Questionnaire

A preliminary list of 81 items generated by Vandervelde et al<sup>15</sup> was submitted to physical therapists, occupational therapists, and medical doctors involved in stroke rehabilitation. They were asked to identify which items were not relevant for patients with stroke, resulting in deletion of 5 items. Three other items that concerned specific lifestyle aspects with no direct correspondence to 1 of the countries studied were removed from the original list. For example, getting on an escalator was deleted because escalators are not common in Benin. A set of 73 items was submitted to both European and African patients with stroke. Patients were asked to provide their perceived

difficulty in performing each activity if completed without technical or human assistance. The response format was a 3-level scale labeled and scored as impossible (0), difficult (1), or easy (2). Unfamiliar activities were recorded as missing responses.

## BI and the FIM

The BI and FIM are observer-rated generic measures of disability widely used in rehabilitation.<sup>20</sup> They are accepted as functional-level assessment tools evaluating the functional status of patients throughout the rehabilitation process. The FIM comprises 18 items,<sup>18</sup> and Linacre et al<sup>21</sup> found that these items define 2 statistically and clinically different indicators: (1) FIM–motor, which assesses disability in motor functions (13 items); and (2) FIM–cognitive, which assesses disability in cognitive functions (5 items). The BI is a 10-item scale assessing different aspects of functional ability for self-care and daily activities.<sup>19</sup> The FIM, although limited, is commonly used by clinicians and researchers as indicated by recent review studies.<sup>22</sup> In a more recent meta-analysis study (2011), it appears that FIM–motor scale and the BI are still used as the main outcome in some randomized controlled trials.<sup>23</sup>

## Data Analysis

### Rasch Analysis and Item Selection

The Rasch analysis tests whether data from a scale satisfy the rules for constructing interval scale measurement.<sup>24</sup> Based on a mathematical model, it estimates person ability and item difficulty by examining a matrix of these items on a common scale comparing individual response patterns with the response pattern of the entire sample.<sup>4</sup> In other words, it is a probabilistic model that converts ordinal scores into interval measures and, in the process, examines other key attributes such as unidimensionality, invariance, sample targeting of a scale, the appropriateness of response format, hierarchy of item difficulty, and the local independence of items within a scale. Consequently, Rasch analysis enables evaluation of the internal construct validity of a scale, and this is judged through overall fit statistics including item fit, person fit, and total  $\chi^2$  probability, which evaluates the extent to which the scale fits the Rasch model. The reliability of the scale was examined using the Person Separation Index, which indicates the extent to which the questionnaire distinguishes distinct ability levels. Rasch analysis and its applications and advantages are described in detail elsewhere.<sup>8,25–30</sup>

Patients’ responses to the ACTIVLIM-Stroke questionnaire and selected personal factors were analyzed using RUMM2030 software for Rasch analysis,<sup>31</sup> under an unrestricted partial credit model. During successive analyses, the following criteria were used for item selection: (1) missing responses: items presenting a missing response rate of  $\geq 20\%$  were removed before analyzing the entire item set; (2) category discrimination: for each item, the 3-level response format was applied defining 2 thresholds of increasing order. Threshold 1 (t1) was between the categories “impossible” and “difficult” and was expected to be followed by threshold 2 (t2) between the categories “difficult” and “easy.” Subjects with higher ability should score higher than subjects with lower ability, indicating correct category discrimination. When the categories were not discriminated as expected, reversed thresholds were observed and these items were deleted; (3) item fit to the model: individual item fit was examined through fit residuals and  $\chi^2$  statistics. Residuals indicated the deviations of items from the expected model score. Only items with residuals within the range  $\pm 2.5$  were considered as fitting model expectations, and others were removed. Because significant  $\chi^2$  probability (below the Bonferroni adjusted value) indicates misfit, these items were removed; (4) DIF: 4 personal factors were used dichotomously to check the invariance of the item difficulty hierarchy: age ( $\leq 55$  years old,  $>55$  years old), sex (male, female), affected side (left, right), and country (Belgian, Benin). The age cutoff of 55 years was based on the median value of our sample’s age distribution, which was 55. Items with DIF were deleted; (5) local dependency: within a scale, local dependency of items affects the test score because it inflates the scale in a particular direction.<sup>32</sup> When items are highly correlated, a patient’s response to 1 item will influence the

**Table 1. Item Calibration and Individual Item Fit of the ACTIVLIM-Stroke 20-Item Rasch-Built Questionnaire**

Items	Location (logit)	SE (logit)	Fit Statistics			
			Fit Residual	$\chi^2$	Degrees of Freedom	$\chi^2$ Probability*
1. Opening a door	-1.95	0.21	-0.74	1.14	2	0.56
2. Brushing one's teeth	-1.85	0.21	0.20	4.42	2	0.11
3. Putting a key in a lock	-1.71	0.20	0.31	3.30	2	0.19
4. Ringing a doorbell	-1.33	0.19	1.30	1.25	2	0.54
5. Getting up from a chair	-1.09	0.17	-1.80	5.53	2	0.06
6. Using the toilet	-0.64	0.16	-1.03	0.35	2	0.84
7. Turning in bed	-0.53	0.16	0.00	1.96	2	0.38
8. Getting out of bed	-0.47	0.15	-1.31	3.90	2	0.14
9. Taking off a t-shirt	-0.45	0.15	-0.70	0.05	2	0.97
10. Picking up something from the floor	-0.30	0.15	-1.00	2.10	2	0.35
11. Getting out of a car	-0.16	0.15	-1.33	2.43	2	0.30
12. Putting the dishes in the cupboard	-0.03	0.16	-0.33	1.67	2	0.43
13. Standing for a long time without support	0.43	0.14	-0.11	0.49	2	0.78
14. Taking a shower	0.50	0.14	-1.42	3.97	2	0.14
15. Putting on socks	0.83	0.14	0.34	0.14	2	0.93
16. Walking upstairs	1.06	0.13	-0.07	2.89	2	0.24
17. Sweeping or vacuuming	1.51	0.15	-0.22	6.26	2	0.04
18. Walking more than one kilometer	1.81	0.13	-1.74	2.79	2	0.25
19. Tying one's laces	2.05	0.13	-0.15	3.59	2	0.17
20. Carrying a heavy load	2.33	0.14	0.94	1.40	2	0.50

\*Bonferroni-adjusted probability value (level: 0.0025).

response to another. We examined residual correlations between items considering correlations of  $\leq 0.2$  as acceptable.

### Scale Validity and Reliability

The external construct validity of the questionnaire was assessed by examining its association with other valid measures. This was based on the concept of convergent validity, which expects good correlation between scales measuring the same aspect. Correlations among FIM-motor, BI, and the ACTIVLIM-Stroke questionnaire were tested. The crosscultural validity of the ACTIVLIM-Stroke questionnaire was examined based on DIF analysis and the invariance of the items' hierarchy across countries.<sup>33,34</sup> Furthermore, we examined the test-retest reliability of the questionnaire in a subsample of 83 subjects who underwent a second evaluation within 1 to 4 weeks.

## Results

### Sample Characteristics

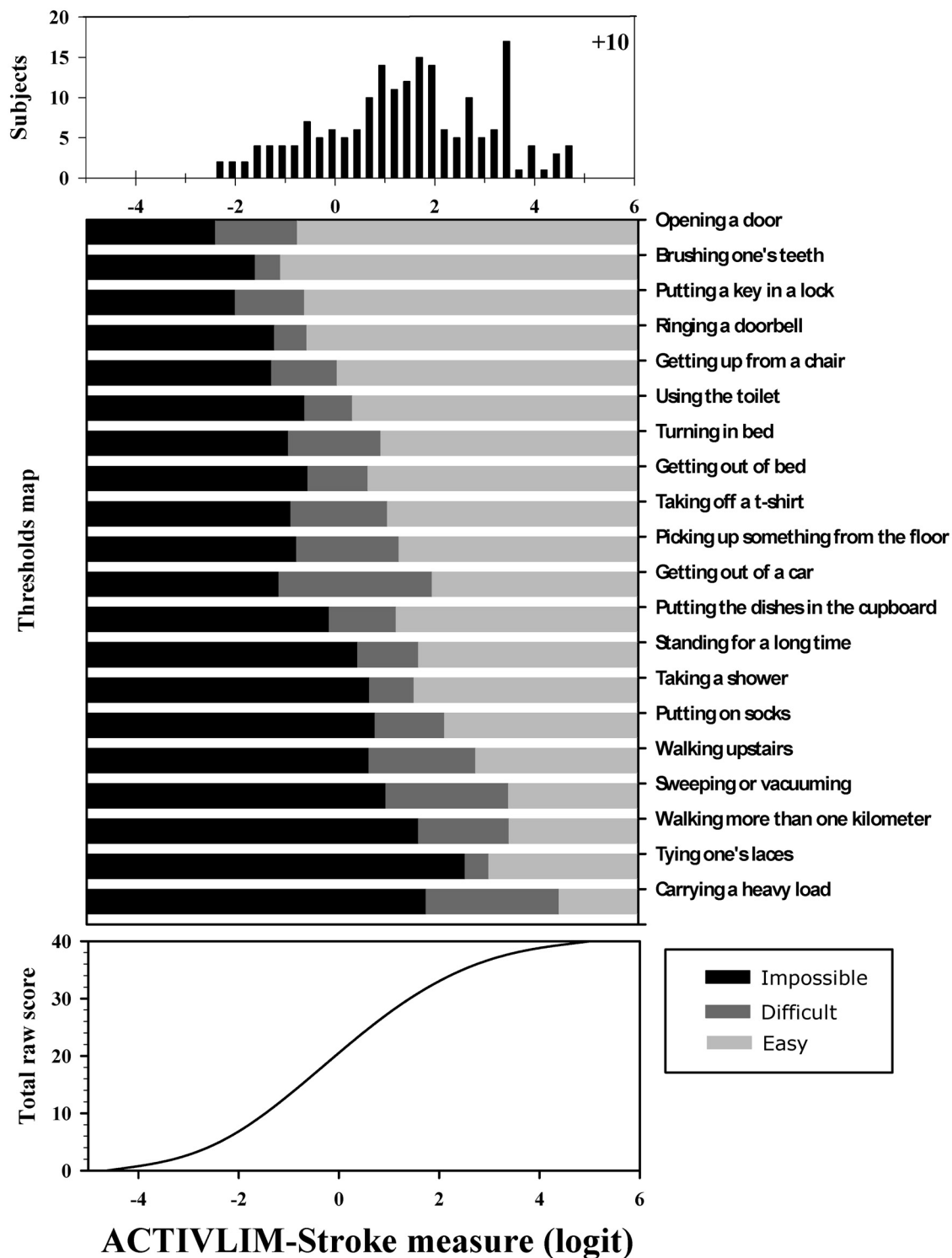
Data from 204 patients with stroke were obtained, all of whom had a stroke occurring at least 1 month previously (108 recruited in Benin and 96 in Belgium). Their mean age  $\pm$  SD was  $57.1 \pm 13.4$  years, their time since stroke was  $21.5 \pm 24.4$  months; their average body mass index was  $25.7 \pm 4.7$  kg/m<sup>2</sup>. The sex distribution was 130 (63.7%) males and 74 (36.3%) females with 113 (55.4%) left paretic side and 91 (44.6%) right paretic side affected; 162 (79.4%) had ischemic stroke and 26 (12.7%) hemorrhagic stroke. The type of stroke was unknown in 16 (7.8%) patients. Their handedness before stroke was 185 (90.7%) right and 19 (9.3%) left and the rehabilitation process was ongoing in 169 (82.8%) patients.

### Item Selection and Metric Quality of the ACTIVLIM-Stroke Questionnaire

From the original 73 items, 3 items presented a missing response rate  $\geq 20\%$  (eg, "putting on a seatbelt") and were removed. The remaining data were fitted to the Rasch model and 8 items had reversed threshold (eg, "using a touchtone phone"), 19 items did not fit the model, 7 items showed DIF (1 for sex, 2 for age, and 4 for country), and 16 presented local dependency with other items. Consequently, 20 remaining items were selected for the ACTIVLIM-Stroke questionnaire.

Analysis of the final 20-item scale demonstrated excellent overall fit statistics for both items (mean  $\pm$  SD,  $-0.44 \pm 0.86$ ) and persons (mean  $\pm$  SD,  $-0.27 \pm 0.83$ ), where fit statistics are standardized to a mean of 0 and a SD of 1. The nonsignificant item-trait interaction ( $\chi^2 = 49.63$ , 40 degrees of freedom;  $P = 0.14$ ) indicated that the hierarchical ranking of the items did not vary across the trait, satisfying the required property of invariance of the scale for patients at different levels of activity limitation. An independent *t* test confirmed the unidimensionality of the scale, where comparisons of estimates derived from different sets of items showed only 3.92% of tests outside the range of  $-1.96$  to  $1.96$ .<sup>35</sup> The CI for the binomial test of proportions was 0.9% to 6.9%. The reliability (Person Separation Index) of the final scale was 0.88.

Table 1 presents the calibration of the final 20-item scale. Items are ordered according to difficulty level from easiest (opening a door:  $-1.95$  logits) to most difficult (carrying a heavy bag:  $2.33$  logits). The table also reports item estimates,

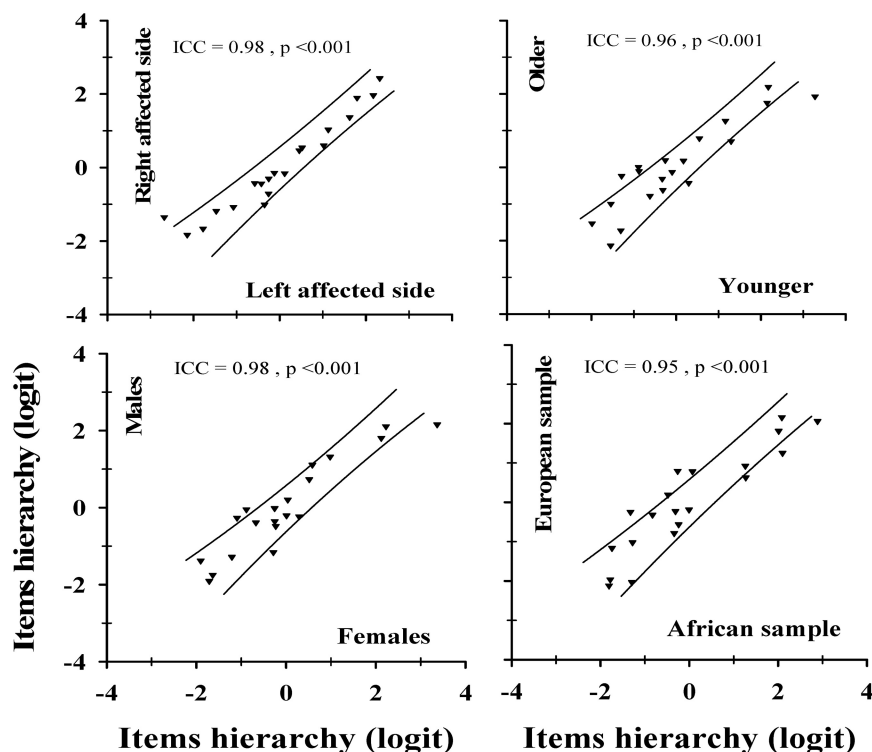


**Figure 1.** Top panel shows patient measures distribution. Approximately 95% of subjects had ACTIVLIM-Stroke measures from  $-2.4$  to  $4.8$  logits. Ten participants answered they could perform all activities easily. Middle panel is the threshold map indicating patient's expected response for each item as a function of ability. Zero is by convention set as the average item difficulty. A patient with an ability of 0 logit would be expected to perform without difficulty the 5 easiest activities, to perform with some difficulties the 7 average activities, and to fail to perform the 8 most difficult activities. Bottom panel represents the relationship between ordinal scores and the ACTIVLIM-Stroke linear measures in logit. This relationship is linear between total raw scores of 5 and 34.

the associated SE of estimation, and the fit statistics including the standard fit residual,  $\chi^2$ , and  $\chi^2$  probability. The individual item fit residual (range,  $-1.80$  to  $1.30$ ) and the  $\chi^2$  probability (range,  $0.04$ – $0.97$ ) indicated that all 20 items

showed good item fit and contributed to the definition of a unidimensional scale of activity limitations. Figure 1 depicts the structure and targeting of the ACTIVLIM-Stroke questionnaire. The top panel shows the distribution of patients





**Figure 2.** Invariance of the ACTIVLIM-Stroke measure tested by comparing the difficulty hierarchy of items in dichotomous subgroups of patients in 4 criteria: affected side, age, sex, and country. More difficult items are in the top right. Solid lines indicate 95% CI of ideal invariance, so items within the solid lines were ranked with the same hierarchy in both patient subgroups. Minor exceptions of items lying outside the 95% CI are seen but these did not demonstrate significant differential item functioning (DIF).

ACTIVLIM-Stroke measures. Approximately 95% of subjects, with ACTIVLIM-Stroke measures ranging from  $-2.4$  to  $4.8$  logits, reported activity limitations with a fine perception of item difficulties, answering impossible, difficult, or easy. No participant found all items impossible and only 10 answered they could perform all activities easily. The scale is therefore well targeted with no considerable floor or ceiling effect. The middle panel shows the expected response to a given item as a function of patient functional ability. Zero is by convention set as the average item difficulty. A patient with an ability of 0 logit would be expected to perform without difficulty the 5 easiest activities, to perform with some difficulties the 7 average activities, and to fail to perform the 8 most difficult activities. The range of difficulties of the 20 items of ACTIVLIM-Stroke questionnaire fit the distribution of the functional abilities of patients with stroke. ACTIVLIM-Stroke measures were then obtained by converting total raw scores into linear measures. The bottom panel illustrates the relationship between ordinal total score and linear measures. This relationship was approximately linear between total raw scores of 5 and 34, in which a unitary increment in total ordinal score was equal to nearly 0.2 logit. Outside this central range, however, the same unitary increment in total ordinal score increased up to 0.9 logit, indicating that a unitary progression in total score accounted for an increasing amount of functional ability measure.

### Invariance of the Items Hierarchy

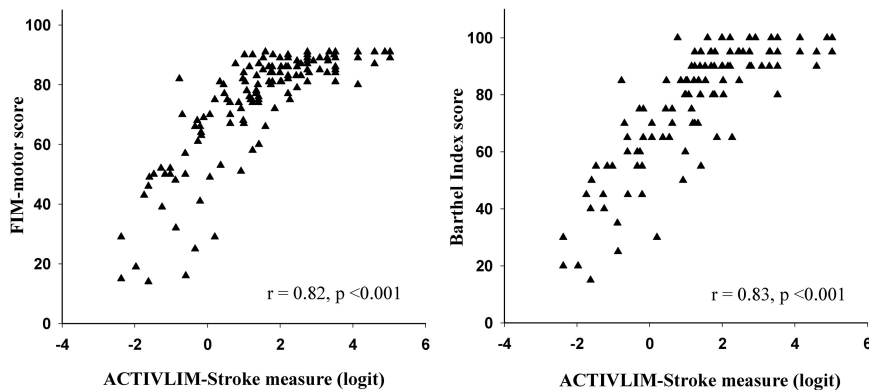
Figure 2 presents differential item functioning plots as a comparison of the difficulty hierarchy of the items as rated by different subgroups according to affected side, age, sex, and country. Most items were within the 95% CI of the identity line, indicating that different subgroups perceived the item

hierarchy similarly. Some minor exceptions of items lying outside the 95% CI were seen, but these did not show significant DIF. Additional statistics, based on intraclass correlation coefficient, confirmed the invariance of the ACTIVLIM-Stroke scale. Indeed, intraclass correlation coefficient between females and males, younger ( $\leq 55$  years old) and older ( $> 55$  years old), left and right affected side, and European and African samples was  $\geq 0.95$  ( $P < 0.001$ ). The ACTIVLIM-Stroke questionnaire demonstrated good invariance of item hierarchy because item difficulty was consistently estimated across groups.

### Validity and Reliability of the ACTIVLIM-Stroke Questionnaire

The ACTIVLIM-Stroke measures were highly and significantly correlated with the BI ( $r = 0.83$ ,  $P < 0.001$ ,  $n = 119$ ) and the FIM-motor ( $r = 0.82$ ,  $P < 0.001$ ,  $n = 140$ ). These correlations demonstrated a good convergent validity of the ACTIVLIM-Stroke questionnaire (Figure 3) and highlighted the low discriminating capacity and high ceiling effect of the FIM-motor and BI. Indeed, 38% ( $n = 77$ ) of participants had FIM-motor scores  $> 80$  of 91, indicating a high level of independence, whereas their ACTIVLIM-Stroke measures ranged from 0.34 to 5.03 logits (Figure 3). Similarly, 27% ( $n = 56$ ) of participants had BI scores  $> 85$  of 100, whereas their activity measures ranged from 0.77 to 5.03 logits (Figure 3). These results confirmed that the widely used FIM and BI do not appear to be sufficiently sensitive to discriminate higher levels of independence.

The test-retest of the ACTIVLIM-Stroke questionnaire in 83 patients showed excellent reliability. Patients' measures (intraclass correlation coefficient = 0.92,  $P < 0.001$ ) and item hierarchy (intraclass correlation coefficient = 0.99,  $P < 0.001$ )



**Figure 3.** Relationships between patient activity measures and the Functional Independence Measure (FIM, left panel) and the Barthel Index (BI, right panel) demonstrated good convergent validity of the ACTIVLIM-Stroke questionnaire but highlighted a ceiling effect of FIM and BI.  $r$  indicates Spearman correlation coefficient.

had very good reproducibility (Figure 4). Item hierarchy and most patient measures were within the 95% CI of the identity line, indicating that: (1) the item difficulty hierarchy was invariant across time; and (2) the patients' estimation of their activity limitations was consistent over time.

### Scale Description

The new ACTIVLIM-Stroke questionnaire is a 20-item scale in which all items were free of local dependency; showed good individual item fit; indicated no DIF for sex, age, affected side, or country; and exhibited no redundancy in item difficulty. Each item has 3 response possibilities (impossible, difficult, easy). The possible total linear scores provide a wide range from  $-4.65$  to  $5.03$  logits. This large measurement range widely explores difficulties of performing daily activities in poststroke patients. For clinical practice, a conversion table is particularly useful. Table 2 presents the conversion of the ACTIVLIM-Stroke raw score (0 to 40) to Rasch measure (logit) and as a centile metric score. This will allow clinicians to quickly transform a patient's raw score to interval measures as long as the patient answers all items.

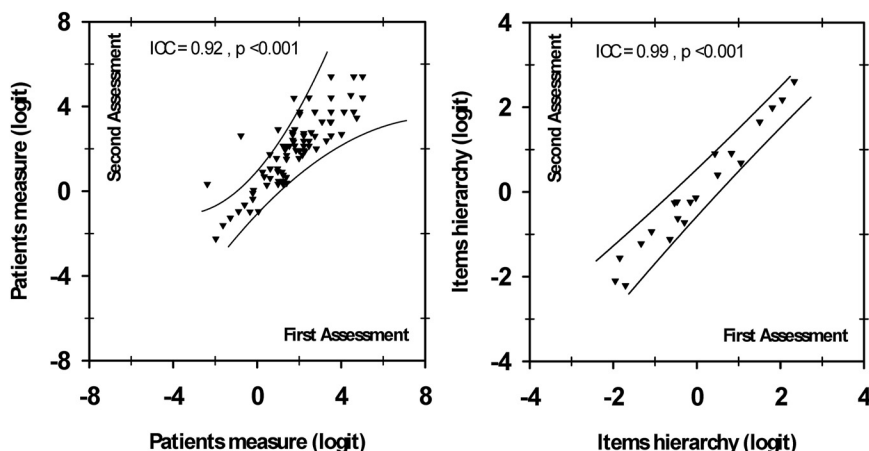
### Discussion

Our study presents a new scale, the ACTIVLIM-Stroke questionnaire, measuring activity limitations in patients with stroke. We describe the calibration and the validation of the scale in a large population of patients with stroke. This new Rasch-built measure was constructed based on both European

and African stroke patients' perception of difficulty in performing daily activities. It provides a practical tool for accurate assessment of activity limitations in clinical settings and in community-based rehabilitation programs. This patient-reported 20-item questionnaire encompasses a large range of activities related to self-care, transfer, mobility, manual ability, and balance, yet remains strictly unidimensional. Thus, the ACTIVLIM-Stroke gives a unidimensional simple summed score of activity limitation at the ordinal level with a transformation to interval scaling where required. It has high internal validity, excellent internal consistency, and good invariance by group. Furthermore, it shows good external construct validity and excellent test-retest reliability.

The Rasch model is based on the principle of unidimensionality, which is a prerequisite to the summation of any set of items.<sup>36,37</sup> Consequently, items fitting the Rasch model such as the ACTIVLIM-Stroke questionnaire items contribute to the construct definition of the scale. As expected from a clinical perspective, activities that require higher energy expenditure<sup>38</sup> such as "walking more than 1 km" or "carrying a heavy load" as well as activities requiring higher bimanual dexterity,<sup>10</sup> for example "tying one's laces," were rated as more difficult.

Analyses of the relationships between patient ACTIVLIM-Stroke scores and other widely used scales also supported the construct validity of the scale (Figure 3). The higher the FIM-motor and BI scores, the higher the ACTIVLIM-Stroke measures. This confirmed that the ACTIVLIM-Stroke mea-



**Figure 4.** Test-retest reliability of the ACTIVLIM-Stroke as reproducibility of patient measures (left panel) and item hierarchy (right panel) between the first and the second assessments. In both panels, the 95% CI of the ideal invariance is a solid line containing patient measure and item hierarchy. ICC indicates intraclass correlation coefficient.

**Table 2. Conversion Table From Ordinal Scores to Interval Scores of the ACTIVLIM-Stroke Measures**

Ordinal Scores (Raw Sum Scores)	Interval Scores	
	Rasch Measures (logit)	Centile Metric Measures (%)
0	-4.6	0
1	-3.8	8.3
2	-3.3	14.0
3	-2.9	17.9
4	-2.6	20.9
5	-2.4	23.5
6	-2.2	25.7
7	-2.0	27.7
8	-1.8	29.5
9	-1.6	31.2
10	-1.5	32.9
11	-1.3	34.5
12	-1.2	36.0
13	-1.0	37.5
14	-0.9	38.9
15	-0.7	40.3
16	-0.6	41.7
17	-0.5	43.1
18	-0.3	44.5
19	-0.2	45.9
20	-0.1	47.3
21	0.1	48.7
22	0.2	50.1
23	0.3	51.5
24	0.5	53.0
25	0.6	54.5
26	0.8	56.0
27	0.9	57.6
28	1.1	59.2
29	1.2	60.9
30	1.4	62.7
31	1.6	64.5
32	1.8	66.5
33	2.0	68.6
34	2.2	71.0
35	2.5	73.5
36	2.8	76.5
37	3.1	80.0
38	3.5	84.5
39	4.1	90.8
40	5.0	100

asures are consistent with patient activity level. Previous studies have pointed out limitation to classical test theory-based scales.<sup>5,7,39–41</sup> We also found that the FIM–motor and the BI were less sensitive than the ACTIVLIM-Stroke questionnaire in chronic stroke patients with high ability level, confirming the appropriateness of the selected items of the

new Rasch-built scale. A possible reason for the high discrimination capacity of the ACTIVLIM-Stroke questionnaire is that it evaluates whether a patient can complete a task without any assistance and, if so, the perceived difficulty, whereas the FIM–motor and BI evaluate a patient on performing a given task, irrespective of the task difficulty perception.

Thus, the ACTIVLIM-Stroke questionnaire might be recommended based on its higher discrimination and might be more sensitive to change in patients with chronic stroke.

The reliability (Person Separation Index) of the ACTIVLIM-Stroke was 0.88, indicating that an appropriate degree of precision will be observed when clinicians or researchers use the questionnaire to evaluate either a group of patients or a single subject.<sup>8</sup> The ACTIVLIM-Stroke also showed excellent test–retest reliability. The intraclass correlation coefficients were 0.99 for item hierarchy and 0.92 for patient measure. These coefficients are sufficiently high to conclude that the questionnaire is reproducible over time and thus provides reliable measures. Additionally, the crosscultural validity of the ACTIVLIM-Stroke questionnaire was tested by checking its invariance across European and African patients with stroke. The absence of DIF supports the crosscultural validity of the new scale.

In addition to (1) being a Rasch-built scale with excellent clinimetric properties; and (2) exhibiting higher ability to discriminate different level of activity limitations in patients with stroke, the ACTIVLIM-Stroke questionnaire has potential add-value compared with existing scales. Indeed, the 20 items of the ACTIVLIM-Stroke questionnaire describe common activities for poststroke patients from Benin and Belgium providing an equal basis for patients' evaluation in both countries. The crosscultural validity of the scale supports its appropriateness and, therefore, its usefulness in multicenter studies.

The 20 items of the ACTIVLIM-Stroke questionnaire are clinically relevant because they represent situations that are regularly experienced by patients with stroke in their everyday life. Activities related to self-care, transfer, mobility, manual ability, and balance are essential human needs and represent a challenge for patients with stroke and their caregivers during rehabilitation.<sup>42–44</sup> For wide and easy use of Rasch-built instruments, some have proposed nomograms that allow the translation of raw sum scores to logits or centile metric score.<sup>24,45</sup> We provide a conversion table of ACTIVLIM-Stroke ordinal summed scores to interval measures. Consequently, even if clinicians are not familiar with Rasch analysis, they will be able to transform a patient's total raw score into a linear measure. However, the conversion table is appropriate only if the patients answer all items. If responses are missing, online analysis is available ([www.rehab-scales.org](http://www.rehab-scales.org)) to directly convert ACTIVLIM-Stroke raw scores into linear Rasch measures of activity limitations, taking missing values into account.

Finally, the ACTIVLIM-Stroke questionnaire meets the key scale assessment criteria provided by the 2 most widely used guideline documents for psychometric standards for rating scales established by the Scientific Advisory Commit-



tee of the Medical Outcomes Trust<sup>46</sup> and the US Food and Drug Administration.<sup>47</sup> Except for responsiveness, which will be evaluated in further studies, this questionnaire satisfies several metric properties such as a conceptual and measurement model, unidimensionality, targeting, reliability, validity, interpretability, respondent and administration burden, alternative forms, and cultural adaptation.

## Conclusions

The ACTICLIM-Stroke questionnaire assesses activity limitations in patients with stroke and is focused on the International Classification of Functioning, Disability and Health activity domain. Rasch analysis enabled robust calibration and validation of this new scale through rigorous selection of 20 items. The questionnaire presents good psychometrics qualities and provides more precise and accurate measures of activity limitations in populations with high activity level than existing traditional questionnaires. The ACTICLIM-Stroke questionnaire was constructed based on both Belgium and Benin stroke patients' perception of difficulty in performing daily activities. A conversion table from ordinal scores to interval scores reinforces the practicality and usefulness of the ACTIVLIM-Stroke questionnaire in clinical settings as well as in research studies.

## Acknowledgments

We thank the patients who participated in the study. We are grateful to the Ministère de l'Enseignement Supérieur et de la Recherche Scientifique de la Communauté française de Belgique, for granting data collection in Benin.

## Disclosures

None.

## References

- World Health Organization. *International Classification of Functioning, Disability and Health: ICF*. Geneva, WHO; 2001.
- Baker K, Cano SJ, Playford ED. Outcome measurement in stroke: a scale selection strategy. *Stroke*. 2011;42:1787–1794.
- Cano SJ, Hobart JC. Watch out, watch out, the FDA are about. *Dev Med Child Neurol*. 2008;50:408–409.
- Rasch G. *Probabilistic Models for Some Intelligence and Attainment Tests*. Chicago: MESA Pr; 1980.
- Tesio L, Simone A, Bernardinello M. Rehabilitation and outcome measurement: where is Rasch analysis going? *Eura Medicophys*. 2007;43:417–426.
- Hobart J, Cano S. Improving the evaluation of therapeutic interventions in multiple sclerosis: the role of new psychometric methods. *Health Technol Assess*. 2009;13:iii, ix–x, 1–177.
- Merkies IS, Hughes RA. The measurement of disease. *J Neurol Neurosurg Psychiatry*. 2010;81:943.
- Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Rheum*. 2007;57:1358–1362.
- Belvedere SL, de Morton NA. Application of Rasch analysis in health care is increasing and is applied for variable reasons in mobility instruments. *J Clin Epidemiol*. 2010;63:1287–1297.
- Penta M, Tesio L, Arnould C, Zancan A, Thonnard JL. The ABILHAND questionnaire as a measure of manual ability in chronic stroke patients: Rasch-based validation and relationship to upper limb impairment. *Stroke*. 2001;32:1627–1634.
- Caty GD, Arnould C, Stoquart GG, Thonnard JL, Lejeune TM. ABILOCO: a Rasch-built 13-item questionnaire to assess locomotion ability in stroke patients. *Arch Phys Med Rehabil*. 2008;89:284–290.
- Engberg A, Garde B, Kreiner S. Rasch analysis in the development of a rating scale for assessment of mobility after stroke. *Acta Neurol Scand*. 1995;91:118–127.
- 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;30:2131–2140.
- Duncan PW, Bode RK, Min Lai S, Perera S. Glycine Antagonist in Neuroprotection Americans Investigators. Rasch analysis of a new stroke-specific outcome scale: the Stroke Impact Scale. *Arch Phys Med Rehabil*. 2003;84:950–963.
- Vandervelde L, Van den Bergh PY, Goemans N, Thonnard JL. ACTIVLIM: a Rasch-built measure of activity limitations in children and adults with neuromuscular disorders. *Neuromuscul Disord*. 2007;17:459–469.
- 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;12:189–198.
- Crum RM, Anthony JC, Bassett SS, Folstein MF. Population-based norms for the Mini-Mental State examination by age and educational level. *JAMA*. 1993;269:2386–2391.
- Granger CV, Hamilton BB, Keith RA, Zielezny M, Sherwin FS. Advances in functional assessment for medical rehabilitation. *Top Geriatr Rehabil*. 1986;1:59–74.
- Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. *Md State Med J*. 1965;14:61–65.
- Houlden H, Edwards M, McNeil J, Greenwood R. Use of the Barthel Index and the Functional Independence Measure during early inpatient rehabilitation after single incident brain injury. *Clin Rehabil*. 2006;20:153–159.
- Linacre JM, Heinemann AW, Wright BD, Granger CV, Hamilton BB. The structure and stability of the Functional Independence Measure. *Arch Phys Med Rehabil*. 1994;75:127–132.
- Chumney D, Nollinger K, Shesko K, Skop K, Spencer M, Newton RA. Ability of Functional Independence Measure to accurately predict functional outcome of stroke-specific population: systematic review. *J Rehabil Res Dev*. 2010;47:17–29.
- Ferrarello F, Baccini M, Rinaldi LA, Cavallini MC, Mossello E, Masotti G, et al. Efficacy of physiotherapy interventions late after stroke: a meta-analysis. *J Neurol Neurosurg Psychiatry*. 2011;82:136–143.
- van der Velde G, Beaton D, Hogg-Johnston S, Hurwitz E, Tennant A. Rasch analysis provides new insights into the measurement properties of the Neck Disability Index. *Arthritis Rheum*. 2009;61:544–551.
- Pallant JF, Tennant A. An introduction to the Rasch measurement model: an example using the Hospital Anxiety and Depression Scale (HADS). *Br J Clin Psychol*. 2007;46:1–18.
- Smith EV Jr, Conrad KM, Chang K, Piazza J. An introduction to Rasch measurement for scale development and person assessment. *J Nurs Meas*. 2002;10:189–206.
- Wright BD, Stone MH. *Best Test Design. Rasch Measurement*. Chicago: MESA Press; 1979.
- Andrich D. *Rasch Models for Measurement*. Newbury Park, CA: Sage Publications; 1998.
- Bond TG, Fox CM. *Applying the Rasch Model. Fundamental Measurement in the Human Sciences*. Mahwah, NJ: Erlbaum Publishers; 2001.
- Penta M, Arnould C, Decruynaere C. *Développer et interpréter une échelle de mesure. Applications du modèle de Rasch*. Hayen, Belgium: Mardaga; 2005.
- Andrich D, Lyne A, Sheridan B, Luo G. *RUMM2030*. Perth, Western Australia: RUMM Laboratory Pty Ltd; 2010.
- Embretson SE, Reise SP. *Item Response Theory for Psychologists*. Mahwah, New Jersey: Lawrence Erlbaum Associates Publishers; 2000.
- Tennant A, Penta M, Tesio L, Grimby G, Thonnard JL, Slade A, et al. Assessing and adjusting for cross-cultural validity of impairment and activity limitation scales through differential item functioning within the framework of the Rasch model: the PRO-ESOR project. *Med Care*. 2004;42:137–48.
- Küçükdeveci AA, Sahin H, Ataman S, Griffiths B, Tennant A. Issues in cross-cultural validity: example from the adaptation, reliability, and validity testing of a Turkish version of the Stanford Health Assessment Questionnaire. *Arthritis Rheum*. 2004;51:14–19.

35. Smith EV Jr. Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *J Appl Meas*. 2002;3:205–231.
36. Streiner D, Norman G. *Health Measurement Scales*. Oxford: Oxford University Press; 1989.
37. Wright BD, Masters GN. *Rating Scale Analysis*. Chicago: MESA Press; 1982.
38. Ainsworth BE, Haskell WL, Whitt MC, Irwin ML, Swartz AM, Strath SJ, et al. Compendium of physical activities: an update of activity codes and MET intensities. *Med Sci Sports Exerc*. 2000;32:S498–504.
39. Schepers VP, Ketelaar M, Visser-Meily JM, Dekker J, Lindeman E. Responsiveness of functional health status measures frequently used in stroke research. *Disabil Rehabil*. 2006;28:1035–1040.
40. Lawton G, Lundgren-Nilsson A, Biering-Sørensen F, Tesio L, Slade A, Penta M, et al. Cross-cultural validity of FIM in spinal cord injury. *Spinal Cord*. 2006;44:746–752.
41. Lundgren-Nilsson A, Grimby G, Ring H, Tesio L, Lawton G, Slade A, et al. Cross-cultural validity of functional independence measure items in stroke: a study using Rasch analysis. *J Rehabil Med*. 2005;37:23–31.
42. Jang SH. The recovery of walking in stroke patients: a review. *Int J Rehabil Res*. 2010;33:285–289.
43. Murray J, Ashworth R, Forster A, Young J. Developing a primary care-based stroke service: a review of the qualitative literature. *Br J Gen Pract*. 2003;53:137–142.
44. McKevitt C, Fudge N, Redfern J, Sheldenkar A, Crichton S, Rudd AR, et al. Self-reported long-term needs after stroke. *Stroke*. 2011;42:1398–1403.
45. van Nes SI, Vanhoutte EK, van Doorn PA, Hermans M, Bakkers M, Kuitwaard K, et al. Rasch-built Overall Disability Scale (R-ODS) for immune-mediated peripheral neuropathies. *Neurology*. 2011;76:337–345.
46. Scientific Advisory Committee of the Medical Outcomes Trust. Assessing health status and quality of life instruments: attributes and review criteria. *Qual Life Res*. 2002;11:193–205.
47. US Food and Drug Administration. Patient reported outcome measures: use in medical product development to support labeling claims. 2009. Available at: [www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf](http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf). Accessed May 24, 2011.