

## RESEARCH PAPER

# Measuring functional recovery in stroke patients: the responsiveness of ACTIVLIM-Stroke

Charles Sèbiyo Batcho,<sup>1,2</sup> Ildéphonse Nduwimana,<sup>3</sup> Jean Olouka,<sup>3</sup> Véronique Halkin,<sup>4</sup> Toussaint Godonou Kpadonou,<sup>3</sup> Jean-Louis Thonnard<sup>1</sup>

<sup>1</sup>Institute of Neuroscience, Université catholique de Louvain, Brussels, Belgium

<sup>2</sup>Centre interdisciplinaire de recherche en réadaptation et intégration sociale (CIRIS), Université Laval, Quebec, Canada

<sup>3</sup>Department of Rehabilitation Medicine, Faculty of Health Sciences, Université d'Abomey-Calavi, Cotonou, Benin

<sup>4</sup>Department of Rehabilitation Medicine, Centre Hospitalier Universitaire Ambroise Paré, Mons, Belgium

## Correspondence to

Dr Charles Sèbiyo Batcho, Institute of Neuroscience, Université catholique de Louvain, 53 Avenue Mounier, COSY-B1.53.04, Brussels B-1200, Belgium; [sebiyo.batcho@uclouvain.be](mailto:sebiyo.batcho@uclouvain.be)

Received 8 November 2013

Revised 8 March 2014

Accepted 12 March 2014

Published Online First

9 April 2014

## ABSTRACT

**Background and purpose** To determine the efficacy of the ACTIVLIM-Stroke questionnaire in detecting changes in functional status of stroke patients.

**Method** Sixty-eight Benin and Belgian stroke patients participated in this 2-month longitudinal study, involving baseline and follow-up assessments. Outcome measures combined ACTIVLIM-Stroke questionnaire, Barthel Index (BI), 6-minute-walk test (6MWT) and modified Rankin scale (MRS). Responsiveness of ACTIVLIM-Stroke was investigated through different methodological approaches and compared with BI, 6MWT and MRS. Statistical analyses were performed using the paired t tests, effect size (ES) and correlation tests.

**Results** ACTIVLIM-Stroke detected changes in the whole sample ( $p < 0.001$ ,  $ES = 0.78$ ) and even in a subgroup of patients ( $p < 0.001$ ,  $ES = 0.29$ ) that were classified as stable according to the MRS. Moreover, ACTIVLIM-Stroke permitted the classification of patients into more discriminative groups, including those showing an important improvement ( $p < 0.001$ ,  $ES = 1.87$ ), a slight but non-clinically meaningful improvement ( $p < 0.001$ ,  $ES = 0.38$ ), and no improvement ( $p = 0.1$ ,  $ES = 0.11$ ), demonstrating its high sensitivity to change.

Furthermore, there were concordant relationships between ACTIVLIM-Stroke change and any observed changes in BI, 6MWT and MRS scores ( $r \geq 0.50$ ,  $p < 0.001$ ), confirming the external responsiveness of ACTIVLIM-Stroke.

**Conclusions** ACTIVLIM-Stroke showed good responsiveness and can detect accurately clinical changes in the functional status of stroke patients. The BI and the 6MWT were also responsive and may provide complementary information while investigating change in functional status. However, in addition to being highly sensitive to change, ACTIVLIM-Stroke presents significant methodological advantages for quantifying functional changes in stroke patients.

## INTRODUCTION

Stroke is prevalent worldwide and is a leading cause of death and disability.<sup>1,2</sup> The effects of stroke on the victim's quality of life are multifaceted and generally more pronounced in the physical dimension, on cognition, language and mood. As described by the International Classification of Functioning, Disability and Health framework, efficient management of poststroke patients requires global assessment, which accounts for impairments, activity limitations and restriction of participation, while also considering personal and environmental factors.<sup>3</sup> This global assessment supposes the

measurement of manifest and latent relevant variables, the latter often requiring the use of patient-reported outcome measures. Acquiring accurate information regarding functional recovery after stroke is essential, as it serves as a baseline for healthcare planning and facilitates therapeutic guidelines. Functional recovery after stroke is generally evaluated using scales that are usually incorporated into clinical trials and are routinely used during clinical practice.<sup>4</sup>

ACTIVLIM-Stroke is a self-reported Rasch-built questionnaire that measures the ability of stroke patients to perform daily living activities. It comprises 20 items and considers various aspects of activity limitations.<sup>5</sup> It offers the advantage of cross-cultural validity between African and European contexts. Its validation in these contexts makes ACTIVLIM-Stroke a powerful tool for advancing stroke rehabilitation in multicentric studies. It also shows excellent psychometric properties, including validity, reliability, unidimensionality and invariance. However, its ability to reveal change over time has not yet been determined.

Turner *et al*<sup>6</sup> reported that an instrument measuring health status should be responsive to changes. Consequently, responsiveness, which reflects the ability of a questionnaire to detect changes over time, is an important psychometric quality and an essential criterion for instrument selection when measuring stroke patient function. Therefore, studies that evaluate the responsiveness of ACTIVLIM-Stroke measure are warmly required.

This study investigates the responsiveness of ACTIVLIM-Stroke in comparison with selected well-known outcome measures, such as the Barthel Index (BI) and the 6-minute walk test (6MWT).

## METHODS

### Study design

We performed a 2-month longitudinal study involving baseline and follow-up assessments. Sixty-eight stroke patients (mean age  $58.52 \pm 12.33$  years old; poststroke duration  $6.86 \pm 11.63$  months; sex 53% female and 47% male) were recruited from the Rehabilitation Departments in Benin and Belgium. Subjects were eligible for inclusion if they had experienced a stroke event at least 2 weeks prior to examination, were undergoing rehabilitation or not, scored 24 or more on the Mini Mental State Examination and freely agreed to participate. Functional status information was collected at baseline and follow-up time points.



CrossMark

**To cite:** Batcho CS, Nduwimana I, Olouka J, *et al.* *J Neurol Neurosurg Psychiatry* 2014;**85**:1337–1342.

The study was approved by the Ethics Committee of the Université catholique de Louvain in Belgium and the local ethics committees of the participating caregiver centres and hospitals in Benin. Each participant signed an informed consent form prior to inclusion in the study.

### Patient assessment

Functional status was measured using a variety of instruments, combining patient-reported and clinician-reported anchors with observed performance. The primary outcome measure was ACTIVLIM-Stroke. Secondary outcome measures were the BI and the 6MWT. The BI is one of the most widely used generic and ordinal scales that measure severity of disability in daily life. It measures a range of activity domains such as communication, eating, self-care, mobility and transfer through 10 items.<sup>7</sup> The 6MWT is a performance test that evaluates walking endurance; it consists in recording the distance (m) a subject walked in 6 min.<sup>8</sup> This test is demonstrated to reflect people functional status.<sup>9 10</sup> Participants performed the 6MWT by walking as quickly as possible on a flat ground square path with a 50–85 m perimeter marked on the floor. They were allowed to stop and rest as necessary. Patients were informed of the time after 2, 4 and 5 min, and the distance walked was recorded to the nearest metre. The higher the score after each outcome measure, the better the functional status of the patient. As such, a positive difference between the second and first scores reflects a positive clinical change.

In addition to the BI and 6MWT, the modified Rankin scale (MRS), which is an assessment of overall disability, was employed to characterise the change in health status of the participants over the 2-month period. This clinician-rated scale categorises the severity of disability based on observation, ranking patients in seven different levels from 'no symptoms at all' to 'dead'. An increasing MRS score indicates worsening status.<sup>11</sup>

### Responsiveness investigation

Analyses were conducted using three approaches: global, group and individual. The global approach consisted of statistically comparing patient baseline scores with those reported at the 2-month follow-up. The group approach allowed investigation of responsiveness indices by patient groups classified as improved, stable or deteriorated according to their MRS scores over the study period. Finally, the individual approach evaluated changes in the functional status of each patient by applying the following formula<sup>12</sup>:

$$t - \text{score} = \frac{m_2 - m_1}{\sqrt{(SE2)^2 + (SE1)^2}}$$

$m_1$  and  $m_2$  represented the patient's functional status measures, respectively, at baseline (time 1) and follow-up (time 2). SE1

and SE2 correspond to their respective associated SE of measurements at time 1 and time 2. The t score approximately follows a standardised normal distribution such that a patient with a t score above 1.96 showed important improvement, while a patient with a t score below -1.96 was significantly worsening.<sup>12</sup> Patients were divided into five classes according to their t score significance limits: important deterioration (t score < -1.96), slight deterioration (-1.96 ≤ t score < 0), no change (t score = 0), slight improvement (0 < t score ≤ 1.96) and important improvement (t score > 1.96).

Analysis of internal responsiveness to change was performed by comparing the follow-up data with baseline data using a paired t test and effect size (ES) where applicable. The ES is calculated as the ratio between the mean change and the SD of the baseline score.<sup>13</sup> Analysis of external responsiveness was evaluated by examining the relationships and concordance between ACTIVLIM-Stroke change and any observed changes in the BI and 6MWT scores. The magnitude of change was interpreted according to Cohen's classification, which describes change as non-significant (ES < 0.2), small (0.2 ≤ ES < 0.5), moderate (0.5 ≤ ES < 0.8) and large (ES ≥ 0.8).<sup>14</sup>

### RESULTS

In terms of internal responsiveness, the mean change in the 68-subject cohort was calculated and the statistical significance was tested based on a paired t test and a signed rank test, respectively, for linear and ordinal-type instruments (table 1). ACTIVLIM-Stroke revealed a significant positive change (p < 0.001), indicating that the overall sample experienced a general health improvement between time 1 and time 2. The ES was 0.78, which corresponded to a large clinical change. The clinicians' ratings of overall disability using the MRS showed that 35 patients (mean poststroke delay of 2.01 ± 2.6 months) improved and that 32 patients (mean poststroke delay of 10.3 ± 7.7 months) were stable, while one deteriorated. As only one patient was rated as deteriorated, the responsiveness analysis by group approach was undertaken with the improved and stable groups (table 2). In the improved group, ACTIVLIM-Stroke confirmed a significant and large positive change (p < 0.001; ES = 0.95). In the stable group, ACTIVLIM-Stroke also detected significant changes (p < 0.001, ES = 0.29); however, this change was considered minor given the small ES.

Analyses based on an individual approach (figure 1) revealed that 8 had a t score between -1.96 and 0, 4 had a t score of 0, 25 had a t score between 0 and 1.96, and 31 patients had a t score above 1.96. This distribution showed that in the group of patients qualified as stable by clinicians according to MRS change score, some had experienced small improvement that was not detectable by MRS. As such, patients were classified into three different groups according to the magnitude of their change as follows: (1) importantly improved (n = 31), (2) non-

**Table 1** Responsiveness analysis following a global approach (n = 68)

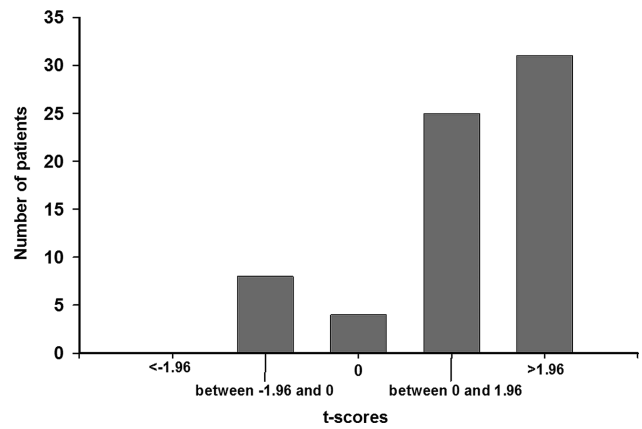
	Time 1	Time 2	Change			ES
			Mean ± SD	95% CI of mean	p Value	
ACTIVLIM-Stroke (logit), mean ± SD	0.77 ± 2.09	2.41 ± 2.19	1.64 ± 1.73	1.22 to 2.06	<0.001	0.78
6-minute walk test (m), mean ± SD	127.57 ± 126.02	192.24 ± 157.46	64.67 ± 107.43	38.66 to 90.67	<0.001	0.51
Barthel Index, median (P25; P75)	80 (60–90)	92.5 (80–100)	–	–	<0.001	–
Modified Rankin scale, median (P25; P75)	3 (2–4)	2 (1.5–3)	–	–	<0.001	–

ES, effect size; P, percentile; time 1, baseline; time 2, follow-up.

**Table 2** Responsiveness analysis by group approach\*

	Improved (n=35, poststroke delay=2.01±2.6 months)						Stable (n=32, poststroke delay=10.3±7.7 months)					
	Time 1			Time 2			Time 1			Time 2		
	Mean±SD	95% CI of mean	p Value	ES	Time 1	Time 2	Mean±SD	95% CI of mean	p Value	ES	Time 1	Time 2
ACTIVLIM-Stroke (logit), mean±SD	2.66±1.75	2.06 to 3.26	<0.001	0.29	0.46±2.23	3.12±2.16	0.54±0.81	0.25 to 0.84	<0.001	0.29	1.59±1.97	1.04±1.89
6-minute-walk test (m), mean±SD	111.45±123.69	68.96 to 153.93	<0.001	0.13	119.8±117.66	231.25±166.73	15.21±55.44	-4.78 to 35.20	0.13	150.90±139.84	135.69±137.84	
Barthel Index, median (P25; P75)	-	-	<0.001	-	70 (56.25; 85)	95 (85; 100)	-	-	<0.001	90 (75; 95)	85 (67.5; 95)	
Modified Rankin scale, median (P25; P75)	-	-	<0.001	-	3 (3-4)	2 (1-2.75)	-	-	0.99	2 (2-3.5)	2 (2-3.5)	

\* Classification of patients as stable or improved was based on their change measured by clinicians using modified Rankin scale. ES, effect size; mo., months; P, percentile; time 1, baseline; time 2, follow-up.

**Figure 1** Distribution of patients (n=68) according to the magnitude of ACTIVLIM-Stroke change as indicated by their t-scores.

clinically meaningfully improved (n=25) and (3) not improved (n=12) (table 3).

Correlations in external responsiveness were observed between scores changes from all instruments. There was a strong positive correlation between ACTIVLIM-Stroke change and BI change ( $p=0.62$ ,  $p<0.001$ ), MRS change ( $p=0.64$ ,  $p<0.001$ ) and 6MWT change ( $r=0.50$ ,  $p<0.001$ ). Moreover, measurements obtained from 6MWT, BI and MRS confirmed the overall improvement detected by ACTIVLIM-Stroke in the global approach analysis (table 1). BI and 6MWT also demonstrated a large positive change in the group of patients classified as improved according to the MRS (table 2). However, in the stable group, only the 6MWT confirmed no significant change ( $p=0.13$ ;  $ES=0.11$ ). The BI similarly detected a small significant improvement, similar to ACTIVLIM-Stroke, confirming the higher sensitivity of ACTIVLIM-Stroke compared with MRS (table 2). Furthermore, a strong concordance was observed between individual self-reported changes based on ACTIVLIM-Stroke and all other instruments (table 3). Indeed, in the groups of patients that reported important individual improvements, all instruments showed significant ( $p<0.001$ ) and clinically meaningful changes with ESs of 1.87 and 1.09, respectively, for ACTIVLIM-Stroke and the 6MWT. In the group that reported a non-clinically meaningful improvement, all instruments were concordant as they revealed a significant change ( $p\leq 0.03$ ). Furthermore, ACTIVLIM-Stroke ( $ES=0.38$ ) and 6MWT ( $ES=0.36$ ) confirmed that these changes were clinically small. Accordingly, in the group that reported no individual improvement, none of these instruments detected statistical significant change ( $p\geq 0.1$ ;  $ES=0.11$ ). Moreover, the higher the poststroke delay was, the lower the ACTIVLIM-Stroke mean change was (figure 2 and table 3), demonstrating a relationship between functional recovery and poststroke duration.

## DISCUSSION

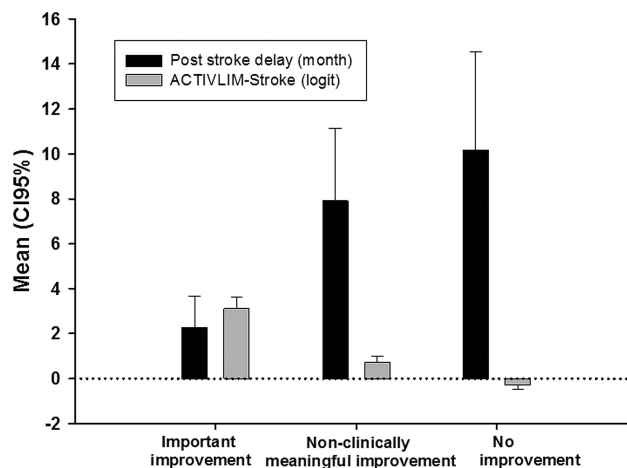
This study evaluated the ability of the ACTIVLIM-Stroke questionnaire to measure relevant changes in stroke patients in comparison with two widely used outcome measures: the BI and the 6MWT. Results showed that ACTIVLIM-Stroke is very sensitive to change and can accurately measure minor and clinically relevant changes in patients over time.

Our analyses showed that changes in ACTIVLIM-Stroke were associated with changes in BI and 6MWT. Data obtained from ACTIVLIM-Stroke, BI and 6MWT all demonstrated that the functional status of the whole sample significantly improved

**Table 3** Reliability of the responsiveness analysis by individual approach\*

	Time 1	Time 2	Change		p Value	ES
			Mean±SD	95% CI of mean		
Important improvement (n=31, psd=2.3±3.8 months)	ACTIVLIM-Stroke (logit), mean±SD	3.06±1.75	3.12±1.35	2.63 to 3.62	<0.001	1.87
	6-minute walk test (m), mean±SD	191.25±137.58	97.25±120.45	53.07 to 141.44	<0.001	1.09
	Barthel Index, median (P25; P75)	70 (47.5; 85)	90 (81.25; 100)	—	—	<0.001
Non-clinically meaningful improvement (n=25, psd=7.9±7.6 months)	ACTIVLIM-Stroke (logit), mean±SD	1.91±2.23	0.72±0.69	0.44 to 1.01	<0.001	0.38
	6-minute walk test (m), mean±SD	146.66±126.41	45.17±100.55	3.67 to 86.68	0.03	0.36
	Barthel Index, median (P25; P75)	85 (73.75; 91.25)	95 (83.75; 96.25)	—	—	<0.001
No improvement (n=12, psd=10.2±6.8 months)	ACTIVLIM-Stroke (logit), mean±SD	2.06±2.65	-0.28±0.32	-0.49 to -0.08	0.1	0.11
	6-minute walk test (m), mean±SD	174.57±184.09	21.09±52.36	-12.18 to 54.36	0.19	0.11
	Barthel Index, median (P25; P75)	92.5 (62.5; 97.5)	95 (70; 97.5)	—	—	0.18

\*Classification of patients in different groups (important change; slight change and no significant change) was based on their individual self-reported change as measured by the calculations of t score. ES, effect size; P, percentile; psd, poststroke delay; time 1, baseline; time 2, follow-up; WI, walk test.



**Figure 2** Groups of patients according to individual self-reported change.

after 2 months. ACTIVLIM-Stroke seemed to be more responsive than BI or 6MWT as it revealed a higher ES. However, this global approach may mask important information as data were pooled together, allowing no distinction between the patients that made significant improvements and those who did not. As such, despite the overall significant improvement, there were wide variations in the magnitude of change among the 68 stroke patients analysed. To rectify this, patients were classified into improved or stable groups based on the MRS as determined by clinicians. This group approach analyses revealed that patients who were classified as improved showed significant improvement with ACTIVLIM-Stroke, BI and 6MWT. In the stable group, only 6MWT showed no significant changes. ACTIVLIM-Stroke and BI detected significant, yet small changes in this group, suggesting that they may be more sensitive than MRS and 6MWT. Although MRS is a valuable instrument for assessing stroke patient global outcomes,<sup>15</sup> yet it demonstrates a lack of sensitivity to small changes.<sup>16</sup> This can be explained by the fact that it is a single-item scale that allows only a broad description of a patient's overall status through a seven-level response format (from 'no symptoms at all' to 'dead'). The ACTIVLIM-Stroke and the BI comprised 20 and 10 non-redundant items, respectively. As such, they capture more details than the MRS, permitting more discrimination and consequently more sensitivity to small changes. In addition, the 6MWT, which was designed to reflect functional recovery in chronic disabled people,<sup>9 10</sup> is a specific test of walking capacity. It measures only part of physical performance (ie, locomotion) and does not evaluate other subdomains of functional status and, therefore, may not be sensitive to small changes occurring in a subdomain other than locomotion. This could justify why 6MWT is less sensitive than the ACTIVLIM-Stroke questionnaire. Similarly, BI exhibits less sensitivity compared with ACTIVLIM-Stroke. This result can be due to the fact that BI measures basic ADL with less detail compared with other commonly used activity of daily living (ADL) scales such as the Functional Independence Measure (FIM) or the Frenchay Index. It is possible that FIM appears to be as sensitive as the ACTIVLIM-Stroke; however, this hypothesis still needs to be investigated.

As shown in previous studies,<sup>17 18</sup> an individual approach, such as the one taken in the present study, provides essential information regarding responsiveness, given that individuals in a group do not equally experience the mean change of the group.

Here, an individual approach analysis allowed the investigation of how each of the 68 patients improved after the 2-month time frame. Based on their individual self-reported activity limitations, patients could then be classified not only as improved or stable, but also in more discriminative groups as having importantly or slightly improved, or showing no improvement. This indicates that, compared with the MRS, ACTIVLIM-Stroke has a greater ability to detect change. It is important to note that the individual approach could only be applied with ACTIVLIM-Stroke, which, like any Rasch-built scale, provides the individual's measures with associated SEs. Indeed, using the approach developed by Wright and Stone,<sup>12</sup> the SEs are an essential requirement for investigating statistically significant individual changes.

The degree of agreement between changes measured by different instruments is essential in responsiveness investigations. This study showed that the individual change reported by patients themselves matched the change measured by other instruments, confirming the external responsiveness of ACTIVLIM-Stroke. 6MWT and BI revealed important significant mean changes in the group of patients who reported having importantly improved. They showed only a small mean change in the patient group reporting non-clinically meaningful improvement, while they showed no significant change in the group reporting no improvement over the 2-month period. These results also demonstrated that patient's self-report scores related well to clinician-based scores and observed performances, confirming previous findings.<sup>19 20</sup>

In addition, our selected outcome measures covered a variety of assessment methods. While BI and MRS are clinician-reported anchors, ACTIVLIM-Stroke is a self-reported questionnaire and 6MWT is a performance-based test. Combining all of these outcome measures allowed a multidimensional assessment of change as determined by the patient's own judgement, the clinician's perception and observed performance.

Our results confirmed that ACTIVLIM-Stroke, BI and 6MWT can be used to monitor the functional recovery of patients undergoing rehabilitation. Nonetheless, it appeared that ACTIVLIM-Stroke showed a greater sensitivity to change. BI and 6MWT present some shortcomings compared with ACTIVLIM-Stroke. First, BI has previously been described as less sensitive to change at extremes of disability (floor and ceiling effects).<sup>16 21</sup> Second, being an ordinal scale, BI restricts statistical investigations. For example, analysis should be limited to non-parametric statistics with no possibility to compute mean changes and ESs. Third, both BI and 6MWT do not provide associated SE to individual measures, making statistical investigations of responsiveness at individual level impossible. More importantly, individual approach of clinical change is essential in clinical trial since it has been demonstrated that ES calculations are potentially misleading indicators of rating scale responsiveness at the group comparison level so that, when comparing various instruments, ES computed at group level should be interpreted carefully.<sup>22</sup> Therefore, it is important to analyse change at individual level since changes that have a meaning in groups may not be meaningful in individuals.<sup>23</sup> Moreover, an outcome measure that permits decision making regarding the significance of a change on an individual basis is extremely useful in clinical practice, where each patient constitutes his own control over time. The linear characteristics of ACTIVLIM-Stroke and the fact that it provides individual SEs allow thorough investigations of its responsiveness in global, group and individual approaches. This particular advantage makes ACTIVLIM-Stroke a recommendable tool in clinical trial

and practice. As it is time efficient, this scale is also a reasonable option when tracking changes in stroke patients' functional status and may provide assistance when making decisions regarding the efficacy of interventions or treatments.

It has been argued that the use of more sensitive measurement instruments would increase the evidence on the efficacy of rehabilitation treatment.<sup>16 24</sup> A good health status outcome measure should allow the detection of small changes over time. This property is essential when analysing the efficacy of a treatment or when monitoring the variations in the health status of a cohort of patients. This study confirms that distribution-based methods and anchor-based approaches are complementary in responsiveness investigations.<sup>4 25</sup> It also shows that global scales with single item are much less sensitive to changes in disability than the activity of daily living scales.<sup>16</sup> Our results also highlighted that, in this cohort, acute and subacute stroke patients experienced better improvement than chronic stroke patients (figure 2). This indicated that functional recovery might be more important in the phases immediately after stroke than in the chronic phase. This observation is supported by existing evidence that generally describes functional poststroke recovery as slowing down in chronic phase.<sup>26–28</sup>

Finally, instruments that are responsive in both severe and mild strokes are required for patients' clinical follow-up. Indeed, due to ceiling effect, many available scales are useless in clinical trials that enrol patients with mild disease conditions. In this study, we have demonstrated the high sensitivity to change of ACTIVLIM-Stroke in patients with relatively moderate disability as revealed by the baseline scores, confirming its usefulness as outcome measure for clinical practice and trial. Nevertheless, further studies could analyse the responsiveness of ACTIVLIM-Stroke in patients with more severe disability. It would also be useful to investigate how responsive is the ACTIVLIM-Stroke questionnaire in comparison to other commonly used scales such as the Functional Independence Measure and the Frenchay Activity index.

## CONCLUSION

In summary, different statistical indices and methodological approaches were combined to investigate the comparative responsiveness of ACTIVLIM-Stroke, BI and 6MWT. Our results showed that all these instruments are responsive and can be used to measure clinical changes in the functional status of stroke patients. ACTIVLIM-Stroke presents good internal and external responsiveness and can detect even small changes. This scale appears to be more sensitive than BI and 6MWT, and offers significant methodological and statistical advantages for investigating change.

**Acknowledgement** The authors thank the patients who participated in the study.

**Contributors** CSB, IN, JO and J-LT contributed to the study design and the data analysis. CSB, IN, JO and VH participated in the data collection. CSB performed the statistical analyses and drafted the manuscript. All authors participated in the data interpretation, critically revised the draft of the manuscript for important intellectual content and contributed to the writing. All authors have read and approved the final manuscript.

**Competing interests** None.

**Patient consent** Obtained.

**Ethics approval** Committee of the Université Catholique de Louvain in Belgium and the local ethics committees of the participating caregiver centres and hospitals in Benin.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** Dataset is available from the corresponding author at [sebiyo.batcho@uclouvain.be](mailto:sebiyo.batcho@uclouvain.be)

## REFERENCES

- 1 Strong K, Mathers C, Bonita R. Preventing stroke: Saving lives around the world. *Lancet Neurol* 2007;6:182–7.
- 2 Owolabi MO. Impact of stroke on health-related quality of life in diverse cultures: The berlin-ibadan multicenter international study. *Health Qual Life Outcomes* 2011;9:81.
- 3 WHO. *The international classification of functioning, disability and health: Icf*. Geneva: WHO, 2001.
- 4 Revicki D, Hays RD, Cella D, et al. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 2008;61:102–9.
- 5 Batcho CS, Tennant A, Thonnard JL. Activlim-stroke: a crosscultural rasch-built scale of activity limitations in patients with stroke. *Stroke* 2012;43:815–23.
- 6 Turner N, Campbell J, Peters TJ, et al. A comparison of four different approaches to measuring health utility in depressed patients. *Health Qual Life Outcomes* 2013;11:81.
- 7 Mahoney FI, Barthel DW. Functional Evaluation: The Barthel Index. *Md State Med J* 1965;14:61–5.
- 8 Guyatt GH, Sullivan MJ, Thompson PJ, et al. The 6-minute walk: a new measure of exercise capacity in patients with chronic heart failure. *Can Med Assoc J* 1985;132:919–23.
- 9 Wegrzynowska-Teodorczyk K, Rudzinska E, Lazorzyc M, et al. Distance covered during a six-minute walk test predicts long-term cardiovascular mortality and hospitalisation rates in men with systolic heart failure: an observational study. *J Physiother* 2013;59:177–87.
- 10 Henricson E, Abresch R, Han JJ, et al. The 6-Minute Walk Test and Person-Reported Outcomes in Boys with Duchenne Muscular Dystrophy and Typically Developing Controls: longitudinal comparisons and clinically-meaningful changes over one year. *PLoS Curr* 2013;doi:10.1371/currents.md.9e17658b007eb79fcd6f723089f79e06
- 11 Bonita R, Beaglehole R. Recovery of motor function after stroke. *Stroke* 1988;19:1497–500.
- 12 Wright BD, Stone MH. *Best test design*. Chicago: Mesa Press, 1979.
- 13 Middel B, van Sonderen E. Statistical significant change versus relevant or important change in (quasi) experimental design: Some conceptual and methodological problems in estimating magnitude of intervention-related change in health services research. *Int J Integr Care* 2002;2:e15.
- 14 Cohen J. *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: L. Erlbaum Associates, 1988.
- 15 Banks JL, Marotta CA. Outcomes validity and reliability of the modified rankin scale: Implications for stroke clinical trials: A literature review and synthesis. *Stroke* 2007;38:1091–6.
- 16 Dromerick AW, Edwards DF, Diringner MN. Sensitivity to changes in disability after stroke: a comparison of four scales useful in clinical trials. *J Rehabil Res Dev* 2003;40:1–8.
- 17 Batcho CS, Durez P, Thonnard JL. Responsiveness of the abillhand questionnaire in measuring changes in rheumatoid arthritis patients. *Arthritis Care Res (Hoboken)* 2011;63:135–41.
- 18 Vandervelde L, Van den Bergh PY, Goemans N, et al. Activity limitations in patients with neuromuscular disorders: A responsiveness study of the activlim questionnaire. *Neuromuscul Disord* 2009;19:99–103.
- 19 Teixeira-Salmela LF, Devaraj R, Olney SJ. Validation of the human activity profile in stroke: A comparison of observed, proxy and self-reported scores. *Disabil Rehabil* 2007;29:1518–24.
- 20 Vandervelde L, Dispa D, Van den Bergh PY, et al. A comparison between self-reported and observed activity limitations in adults with neuromuscular disorders. *Arch Phys Med Rehabil* 2008;89:1720–3.
- 21 Quinn TJ, Langhorne P, Stott DJ. Barthel index for stroke trials: Development, properties, and application. *Stroke* 2011;42:1146–51.
- 22 Hobart JC, Cano SJ, Thompson AJ. Effect sizes can be misleading: is it time to change the way we measure change? *J Neurol Neurosurg Psychiatry* 2010;81:1044–8.
- 23 Redelmeier DA, Tversky A. Discrepancy between medical decisions for individual patients and for groups. *N Engl J Med* 1990;322:1162–4.
- 24 Ferrarello F, Baccini M, Rinaldi LA, et al. Efficacy of physiotherapy interventions late after stroke: A meta-analysis. *J Neurol Neurosurg Psychiatry* 2011;82:136–43.
- 25 Husted JA, Cook RJ, Farewell VT, et al. Methods for assessing responsiveness: A critical review and recommendations. *J Clin Epidemiol* 2000;53:459–68.
- 26 Jorgensen HS, Nakayama H, Raaschou HO, et al. Outcome and time course of recovery in stroke. Part i: Outcome. The copenhagen stroke study. *Arch Phys Med Rehabil* 1995;76:399–405.
- 27 van de Port IG, Kwakkel G, van Wijk I, et al. Susceptibility to deterioration of mobility long-term after stroke: a prospective cohort study. *Stroke* 2006;37:167–71.
- 28 Hankey GJ, Jamrozik K, Broadhurst RJ, et al. Long-term disability after first-ever stroke and related prognostic factors in the perth community stroke study, 1989–1990. *Stroke* 2002;33:1034–40.