



The adaption and evaluation of a Latvian version of the National Institutes of Health Stroke Scale

Kristaps Jurjāns^{1,2}, Ilja Noviks^{1,2},
Dīana Volčeka¹, Linda Zandersone¹,
Kristīne Meilerte¹, Evija Miglāne^{1,2},
Ainars Stepens¹ and Andrejs Millers^{1,2}

Abstract

Objectives: To determine the validity and reliability of a Latvian version of the National Institutes of Health Stroke Scale (LV-NIHSS) for evaluating Latvian stroke patients.

Methods: The adaption of the LV-NIHSS followed standard methods used for the adaption and validation of clinical assessment tools. The scale validity was tested by comparison with the Glasgow Coma Scale (GCS) and the modified Rankin scale (mRs). The reliability of the LV-NIHSS was evaluated by intra-rater and inter-rater agreement using intra-class correlation coefficient (ICC) analysis.

Results: A total of 296 stroke patients and 101 control subjects were evaluated. The mean age of the overall study population was 73.6 years (range, 37 – 94 years; 227 [57.2%] were female). The mean LV-NIHSS score of the patients with stroke was 8.4 ± 6.2 . In terms of construct validity of the LV-NIHSS, it correlated with the GCS ($r = -0.571$) and mRs ($r = 0.755$). In terms of the reliability of the LV-NIHSS, the inter-rater agreement had an ICC of 0.99 and the intra-rater agreement had an ICC of 0.99.

Conclusion: The adaption of LV-NIHSS was successful and the evaluation showed that the scale was valid and reliable for evaluating Latvian stroke patients.

Keywords

National Institutes of Health Stroke Scale, stroke, validity, reliability, Latvia

Date received: 27 April 2016; accepted: 26 July 2016

¹Department of Neurology and Neurosurgery, Riga Stradins University, Riga, Latvia

²Department of Neurology, Pauls Stradins Clinical University Hospital, Riga, Latvia

Corresponding author:

Kristaps Jurjāns, Department of Neurology, Pauls Stradins Clinical University Hospital, Riga, 13 Pilsonu Street, LV-1002, Riga, Latvia.

Email: kristaps.jurjans@gmail.com



Introduction

The National Institutes of Health Stroke Scale (NIHSS) is a 15-item English-language clinical evaluation instrument that is widely used to assess neurological deficit and objectively quantify the impairment caused by a stroke.¹ The scale is often used on the first day of hospitalization and retested when the patient is discharged. The validity and reliability of the original NIHSS has been proven in multiple clinical studies and is widely documented in the literature.²⁻⁴

According to available data from the National Health Service of Latvia, the incidence of stroke is increasing in Latvia as it is in the rest of Europe due to an ageing population.⁵ Since 2000, seven stroke units have been established in Latvia. The use of the English-language NIHSS in Latvia is prevented by the language barrier, so a valid and reliable Latvian version of the NIHSS is required to enable physicians in Latvia to assess Latvian-speaking patients with stroke. The adaption of the NIHSS to provide a Latvian version needs to follow standard methods for the adaption and validation of the NIHSS,⁶ as has been done in other countries, including Spain,⁷ India,⁸ Korea⁹ and China.¹⁰ This aim of this present study was to adapt the English-language NIHSS to produce a Latvian version (LV-NIHSS), which was then evaluated for validity and reliability in a population of Latvian patients with stroke.

Patients and methods

Study design

The study was undertaken in two phases: phase 1 was the adaption of the English-language NIHSS to produce a Latvian version; and phase 2 evaluated the validity and reliability of the LV-NIHSS in Latvian-speaking patients with stroke. Verbal informed consent was required from each study participant or their relatives before they could be included in the study.

The study was approved by the Ethics Committee of Riga Stradins University, Riga, Latvia.

Phase 1 –adaption of the English-language NIHSS into Latvian

The first phase of the study was the adaptation of the original English-language version of the NIHSS into Latvian. The adaptation followed standard methods proposed by the International Quality of Life Assessment Project and were similar to those recommended by other authors.⁷⁻¹⁰ The adaptation was performed in four steps: (i) step 1 – the first intermediate Latvian version was produced by two independent physicians who undertook two forward translations of the original English-language NIHSS; (ii) step 2 – the forward translations were tested using a back translation of the first intermediate LV-NIHSS version by a certified translator. The back translation was compared with the original English-language version of the NIHSS in order to assess conceptual equivalence and detect possible linguistic misunderstandings or misinterpretations. Following this comparison, a second intermediate Latvian version of the NIHSS was prepared; (iii) step 3 – the second intermediate Latvian version was reviewed by a committee of Latvian-speaking neurologists, representing the largest hospitals in Latvia, to assess comprehension and acceptability of the translated NIHSS. Each member of the committee reviewed the LV-NIHSS and made individual additions and corrections. Any discrepancies that arose were resolved by consensus to produce a third intermediate Latvian version of the NIHSS; (iv) step 4 – two independent proof readers evaluated the third intermediate Latvian version to correct any spelling, grammar, or other mistakes. The corrected version became the final preliminary LV-NIHSS. Figure 1 presents a pictorial diagram, a naming sheet and a list

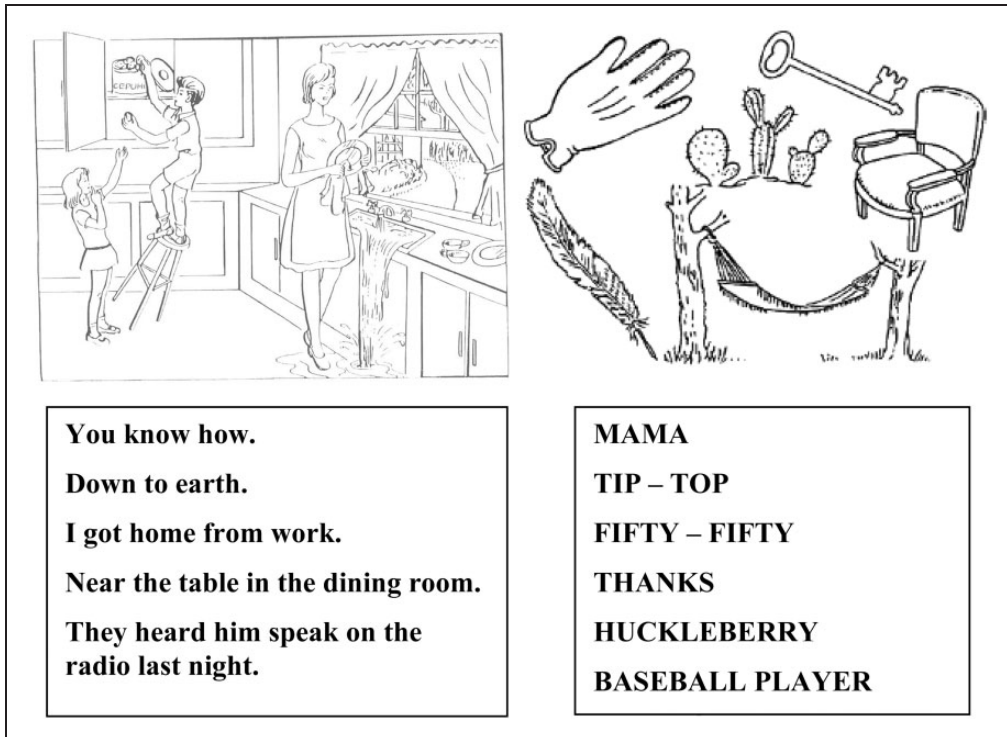


Figure 1. Pictorial diagram for testing items 9 (aphasia) and 10 (dysarthria) in the English-language version of the National Institutes of Health Stroke Scale.

of sentences that are used in the English-language NIHSS for testing language, visual extinction, and inattention. A grammatical and semantic adaptation was created with some modifications made to the original drawings so that they would be recognizable by the Latvian population (Figure 2).

Phase 2 – validity and reliability testing of the LV-NIHSS

The goal of the second phase of the study was to evaluate the psychometric properties of the LV-NIHSS. Its reliability was evaluated by a total of four physicians all of whom were certified by The National Institutes of Health (NIH) for using the original English-language version of the NIHSS. The undertaking of the second

phase of the study was supervised by Professor A Millers, the head of the Department of Neurology, Pauls Stradins Clinical University Hospital, Riga, Latvia.

The prospective phase of the study enrolled ischaemic stroke patients who were hospitalized in the Department of Neurology, Pauls Stradins Clinical University Hospital, Riga, Latvia between November 2014 and January 2015. Acute ischaemic stroke patients were included if the diagnosis was confirmed by a computed tomography scan. Exclusion criteria were as follows: (i) haemorrhagic stroke; (ii) invasive manipulations (artery stenting or thrombectomy < 2 days before enrolment); (iii) unstable Glasgow Coma Scale (GCS) parameters.¹¹ The patient's age, sex, stroke localization, GCS score and modified Rankin Scale (mRs) score were recorded.¹²

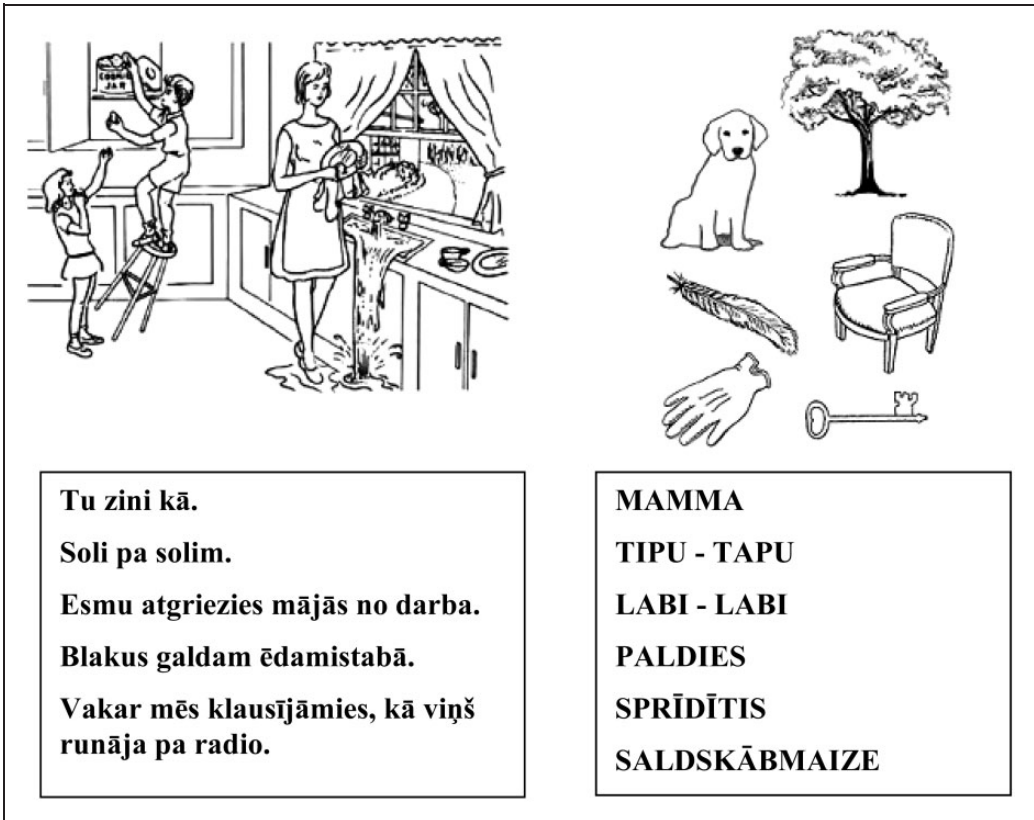


Figure 2. Pictorial diagram for testing items 9 (aphasia) and 10 (dysarthria) in the adapted Latvian version of the National Institutes of Health Stroke Scale.

The type of stroke was identified by computed tomography. All patients were divided into two groups based on the localization of the stroke: an anterior circulation stroke or a posterior circulation stroke.

The control group comprised of patients hospitalized due to non-neurological disease at the Pauls Stradins Clinical University Hospital, Riga, Latvia between November 2014 and January 2015. Control subjects who had prior strokes, other central nervous system diseases, or an unstable clinical condition according to the GCS or the mRS were excluded.

The LV-NIHSS was administered by physicians who were certified by the NIH for using the original English-language

version of the NIHSS. The LV-NIHSS was administered following the translated/adapted recommendations suggested by the original NIHSS as follows:¹ (i) the examiner did not coach or help the patient with the assigned task; (ii) the examiner sometimes demonstrated the commands to patients that were unable to comprehend verbal instructions; (iii) the examiner sometimes helped the patient to get into a position to begin the test, but provided no further assistance while the patient was attempting to complete the task; (iv) for each item, the examiner scored the patient's first effort and repeated attempts did not affect the patient's score with the exception of when scoring for language (Item 9) as the patients best effort

was scored. The LV-NIHSS total score was calculated as recommended by the original NIHSS guidelines.¹

The inter-rater reliability of the LV-NIHSS was evaluated by two independent NIH-certified physicians who independently administered the scale to clinically and neurologically stable patients with ischaemic stroke 5–7 days after the onset of the stroke with a mean interval between evaluations of 3 h. The intra-rater reliability of the LV-NIHSS was evaluated by two independent NIH-certified physicians who administered the LV-NIHSS to clinically and neurologically stable patients with ischaemic stroke 5–7 days after the onset of the stroke, and after 24 h, the same physicians repeated the assessment using the LV-NIHSS with the same patients.

The construct validity of the LV-NIHSS was determined by comparison with the GCS and mRs, which provide a global rating of the patient's status and are reliable tools for the examination of patients with neurological diseases.^{13,14} The NIH-certified physicians used the LV-NIHSS and the GCS and mRs to assess each patient within 24 h of hospital admission. Spearman's rank correlation coefficient was calculated between the LV-NIHSS and the GCS and mRs.

Statistical analyses

All statistical analyses were performed using the SPSS® statistical package, version 22.0 (SPSS Inc., Chicago, IL, USA) for Windows®. The demographic and clinical characteristics of the study population are presented. Continuous data are presented as mean ± SD and categorical data as *n* of patients. The inter-rater correlation was calculated by measuring the intra-class correlation coefficient between two independent observers evaluating the same patients consecutively. Intra-rater reliability was calculated by measuring the intra-class

correlation coefficient for two consecutive observations of the same patients within a 24-h interval by the same physician. Spearman's rank correlation coefficient was calculated between the LV-NIHSS and the GCS and mRs score to evaluate the construct validity of the LV-NIHSS. A *P*-value < 0.05 was considered statistically significant.

Results

The second prospective phase of the study evaluated a total of 397 patients; 296 patients with stroke and 101 control subjects. The mean age of the overall study population was 73.6 years (range, 37 – 94 years; 227 [57.2%] were female). The mean LV-NIHSS score of the patients with stroke was 8.4 ± 6.2. A total of 233 patients (78.7%) had an anterior circulation stroke (Table 1).

To determine the inter-rater reliability of the LV-NIHSS, a total of 200 patients with stroke were evaluated. Excellent inter-rater reliability was shown with the Cronbach's alpha ranging from 0.97 to 1.00 for the 15 individual items of the LV-NIHSS assessment (Table 2). The intra-class correlation coefficient was 0.99 (*P* < 0.001). The inter-rater reliability was excellent independent of the stroke localization as demonstrated by the two groups of patients with stroke (anterior circulation [*n* = 166] versus posterior circulation [*n* = 34]), which both had a Cronbach's alpha of 0.99 and an intra-class correlation coefficient of 0.99 (*P* < 0.001) when evaluated.

A total of 96 patients with stroke and 101 control subjects were included in the assessment of intra-rater reliability. The intra-rater reliability for the 15 individual items of the LV-NIHSS was excellent with the Cronbach's alpha ranging from 0.99 to 1.00 and the intra-class correlation coefficient was 0.99 (*P* < 0.001) (Table 3). The intra-rater reliability was excellent

Table 1. Demographic and clinical characteristics of stroke patients compared with the control group who participated in this study to test the validity and reliability of a Latvian version of the National Institutes of Health Stroke Scale (LV-NIHSS).

Characteristic	Stroke patients <i>n</i> = 296	Control group <i>n</i> = 101
Age, years	75.6 ± 9.5	72.6 ± 8.6
Glasgow Coma Scale score	14.1 ± 1.8	15.0 ± 1.2
Female	172 (58.1%)	55 (54.5%)
LV-NIHSS score	8.4 ± 6.2	0.6 ± 0.2
Anterior circulation stroke	233 (78.7%)	
Posterior circulation stroke	63 (21.3%)	

Data presented a mean ± SD or *n* of patients (%).

Table 2. Inter-rater reliability, comparison of raters and inter-rater agreement (*n* = 200) for individual items and overall scores of a Latvian version of the National Institutes of Health Stroke Scale.

Item		Cronbach's alpha	ICC	<i>P</i> -value
1a	LOC	0.99	0.99	<i>P</i> < 0.001
1b	LOC questions	0.98	0.98	<i>P</i> < 0.001
1c	LOC commands	0.98	0.98	<i>P</i> < 0.001
2	Best gaze	0.99	0.98	<i>P</i> < 0.001
3	Visual	0.98	0.98	<i>P</i> < 0.001
4	Facial Palsy	0.98	0.98	<i>P</i> < 0.001
5a	Motor arm left	0.99	0.98	<i>P</i> < 0.001
5b	Motor arm right	0.99	0.99	<i>P</i> < 0.001
6a	Motor leg left	1.00	1.00	<i>P</i> < 0.001
6b	Motor leg right	0.99	0.99	<i>P</i> < 0.001
7	Limb ataxia	0.99	0.99	<i>P</i> < 0.001
8	Sensory	0.97	0.97	<i>P</i> < 0.001
9	Best language	0.98	0.98	<i>P</i> < 0.001
10	Dysarthria	0.98	0.98	<i>P</i> < 0.001
11	Extinction and inattention	0.98	0.98	<i>P</i> < 0.001
	Total	0.99	0.99	<i>P</i> < 0.001

LOC, level of consciousness; ICC, intra-class correlation coefficient.

independent of the mRs score as patients with an mRs of 1 (*n* = 3) had a Cronbach's alpha of 1.00 and an intra-class correlation coefficient of 1.00. Patients with an mRs of 2 (*n* = 19), 4 (*n* = 19), and 5 (*n* = 36) had a Cronbach's alpha of 0.99 and an intra-class correlation coefficient of 0.99. Patients with an mRs of 3 (*n* = 19) had the lowest values with a Cronbach's alpha of 0.98 and an

intra-class correlation coefficient of 0.98 (*P* < 0.001). The intra-rater reliability was excellent independent of the stroke localization as demonstrated by the two groups of patients with stroke (anterior circulation [*n* = 67] versus posterior circulation [*n* = 29]), which both had a Cronbach's alpha of 0.99 and an intra-class correlation coefficient of 0.99 (*P* < 0.001) when evaluated.

Table 3. Intra-rater reliability, comparison of raters and intra-rater agreement ($n = 96$) for individual items of a Latvian version of the National Institutes of Health Stroke Scale.

Item		Crude agreement	Cronbach's alpha	P-value
1a	LOC	98	0.995	$P < 0.001$
1b	LOC questions	97	0.992	$P < 0.001$
1c	LOC commands	98	0.994	$P < 0.001$
2	Best gaze	98	1.000	$P < 0.001$
3	Visual	98	1.000	$P < 0.001$
4	Facial Palsy	100	1.000	$P < 0.001$
5a	Motor arm left	99	0.998	$P < 0.001$
5b	Motor arm right	97	0.993	$P < 0.001$
6a	Motor leg left	98	0.994	$P < 0.001$
6b	Motor leg right	98	0.996	$P < 0.001$
7	Limb ataxia	99	0.994	$P < 0.001$
8	Sensory	100	1.000	$P < 0.001$
9	Best language	99	0.998	$P < 0.001$
10	Dysarthria	99	0.993	$P < 0.001$
11	Extinction and inattention	100	1.000	$P < 0.001$
	Intra-class correlation coefficient	0.999		$P < 0.001$

LOC, level of consciousness.

The construct validity of the LV-NIHSS was evaluated by comparing it with the GCS and the mRs scores. The LV-NIHSS inversely correlated with the GCS ($r = -0.571$, $P < 0.01$). A positive correlation between the LV-NIHSS and the mRs was observed ($r = 0.755$; $P < 0.01$).

Discussion

The NIHSS is a 15-item English-language clinical evaluation instrument that is widely used to assess the degree of neurological impairment.⁴ The frequent usage of the scale in clinical trials and observational studies has been widely documented in the literature.^{2,4,5,6,15} The original scale was developed and validated in English, which is of limited application in Latvia. Therefore, this present study aimed to adapt the NIHSS to produce a Latvian version, which was then tested for its validity and reliability using standard methods.

The present study demonstrated high inter-rater agreement rates for the LV-NIHSS when tested in a Latvian-speaking population. A test is considered reliable if the measurement error (variance) rate is low. The reliability includes reproducibility among different observers and consistency between scale items. The scales reliability is important because the error increases if there is poor reproducibility among separate observers. On the other hand, inter-rater reliability measures the agreement among physicians performing the assessment as it ranges from 0 (no reliability) to 1 (perfect reliability).

Intra-rater reliability was excellent with the Cronbach's alpha ranging from 0.99 to 1.00 with an intra-class correlation coefficient of 0.99 ($P < 0.001$). Agreement was the highest for items evaluating gaze, visual, facial palsy, sensation and extinction and inattention. The lowest agreement was for the item of 1b level of continence questions

(Cronbach's alpha 0.992, $P < 0.001$), motor right arm (Cronbach's alpha 0.993, $P < 0.001$) and dysarthria (Cronbach's alpha 0.993, $P < 0.001$). In this present study, as in a Spanish study,⁷ the raters were NIH-certified physicians, whereas when the Hindi version of the NIHSS was evaluated one of the raters was a research fellow and not a neurologist.⁷ This probably accounts for the higher agreement rates observed in the present study and the Spanish study compared with the Indian study.^{7,8}

The original NIHSS reliability evaluation showed that the results for dysarthria and level of consciousness were rated fair to poor for inter-rater agreement.^{6,16} In comparison, the present study demonstrated that dysarthria had a substantially higher inter-rater agreement (intra-class correlation coefficient 0.98, $P < 0.001$), as was the case for the item assessing the level of consciousness (intra-class correlation coefficient 0.99, $P < 0.001$).

Among over 15 000 individuals who have taken online NIHSS certification, the NIHSS items with a poorer inter-rater reliability include facial palsy ($k = 0.25$), ataxia ($k = 0.15$), level of consciousness ($k = 0.43$), dysarthria ($k = 0.46$) and gaze ($k = 0.44$).¹⁵ These NIHSS items with poor inter-rater reliability have also been identified in the Spanish and Chinese versions of the NIHSS.^{7,10} However, the LV-NIHSS evaluation showed high intra-class correlation coefficients for all items.

In conclusion, the present study showed that the Latvian version of the NIHSS is valid and reliable. This will allow the scale to be widely used in Latvian hospitals to accurately assess all Latvian-speaking stroke patients. Also, general practitioners could be taught to use this scale in their management of stroke survivors.

Acknowledgements

We would like to thank Ramona Valante, Viktorija Perepjolkina, Olga Minibajeva,

Jekaterina Jagodzinska, Inara Abelite and Signeta Schäfer for their contributions.

Declaration of conflicting interests

The authors declare that there are no conflicts of interest.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

1. National Institute of Neurological Disorders and Stroke. NIH Stroke Scale, http://www.ninds.nih.gov/doctors/NIH_Stroke_Scale.pdf (accessed 7 September 2016).
2. Dewey HM, Donnan GA, Freeman EJ, et al. Interrater reliability of the National Institutes of Health Stroke Scale: rating by neurologists and nurses in a community-based stroke incidence study. *Cerebrovasc Dis* 1999; 9: 323–327.
3. Goldstein LB, Bertels C and Davis JN. Interrater reliability of the NIH stroke scale. *Arch Neurol* 1989; 46: 660–662.
4. Brott T, Adams HP Jr, Olinger CP, et al. Measurements of acute cerebral infarction: a clinical examination scale. *Stroke* 1989; 20: 864–870.
5. Budincevic H, Tiu C, Bereczki D, et al. Management of ischemic stroke in Central and Eastern Europe. *Int J Stroke* 2015; 10(Suppl A100): 125–127.
6. Meyer BC, Hemmen TM, Jackson CM, et al. Modified National Institutes of Health Stroke Scale for use in stroke clinical trials: prospective reliability and validity. *Stroke* 2002; 33: 1261–1266.
7. Domínguez R, Vila JF, Augustovski F, et al. Spanish cross-cultural adaptation and validation of the National Institutes of Health Stroke Scale. *Mayo Clin Proc* 2006; 81: 476–480.
8. Prasad K, Dash D and Kumar A. Validation of the Hindi version of National Institute of Health Stroke Scale. *Neurol India* 2012; 60: 40–44.

9. Oh MS, Yu KH, Lee JH, et al. Validity and reliability of a Korean version of the National Institutes of Health Stroke Scale. *J Clin Neurol* 2012; 8: 177–183.
10. Sun TK, Chiu SC, Yeh SH, et al. Assessing reliability and validity of the Chinese version of the stroke scale: scale development. *Int J Nurs Stud* 2006; 43: 457–463.
11. Centers for Disease Control and Prevention. Glasgow Coma Scale, <http://www.cdc.gov/masstrauma/resources/gcs.pdf> (accessed 7 September 2016).
12. Heart & Stroke Foundation. Canadian Partnership for Stroke Recovery. Modified Rankin Scale, <http://www.stroking.ca/pdf/mrs.pdf> (accessed 7 September 2016).
13. Fischer J and Mathieson C. The history of the Glasgow Coma Scale: implications for practice. *Crit Care Nurs Q* 2001; 23: 52–58.
14. Wilson JT, Hareendran A, Hendry A, et al. Reliability of the modified Rankin Scale across multiple raters: benefits of a structured interview. *Stroke* 2005; 36: 777–781.
15. Ghandehari K. Challenging comparison of stroke scales. *J Res Med Sci* 2013; 18: 906–910.
16. Cincura C, Pontes-Neto OM, Neville IS, et al. Validation of the National Institutes of Health Stroke Scale, modified Rankin scale and Barthel Index in Brazil: the role of cultural adaptation and structured interviewing. *Cerebrovasc Dis* 2009; 27: 119–122.