Published in final edited form as:

Int J Stroke. 2010 April ; 5(2): 74–79. doi:10.1111/j.1747-4949.2010.00411.x.

Production and validation of Putonghua- and Cantonese-Chinese language National Institutes of Health Stroke Scale Training and Certification Videos

R. T. F. Cheung^{1,2,3,*}, P. D. Lyden^{4,5}, T. H. Tsoi⁶, Y. Huang⁷, M. Liu⁸, S. F. K. Hon³, R. Raman^{4,9}, and L. Liu⁹

¹ Department of Medicine (Neurology), and Research Centre of Heart, Brain, Hormone and Healthy Aging, University of Hong Kong, Hong Kong

² Acute Stroke Services, Hong Kong West Cluster Hospitals, Hong Kong

³ Department of Medicine (Neurology), Queen Mary Hospital, Hong Kong

⁴ Department of Neurosciences, UCSD School of Medicine, San Diego, CA, USA

⁵ Department of Neurology, Veterans Administration Medical Center, San Diego, CA, USA

⁶ Department of Medicine, Pamela Youde Nethersole Eastern Hospital, Hong Kong

⁷ Department of Neurology, Peking University First Hospital, Beijing, China

⁸ Department of Neurology, West China Hospital, Sichuan University, Chengdu, China

⁹ Department of Family and Preventive Medicine, UCSD School of Medicine, San Diego, CA, USA

Abstract

Background and purposes—The National Institutes of Health Stroke Scale (NIHSS) is an integral part of acute stroke assessment. We report our experience with new Putonghua- and Cantonese-Chinese language NIHSS (PC-NIHSS and CC-NIHSS) training and certification videos.

Methods—A professional video production company was hired to create the training and certification videos for both PC-NIHSS and CC-NIHSS. Two training and certification workshops were held in Chengdu and Beijing, and two workshops in Hong Kong. The instruction, training and group A certification videos were presented to workshop attendees. Un-weighted κ statistics were used to measure the agreement among raters, and the inter-rater agreements for PC-NIHSS and CC-NIHSS videos were compared with those of original English language NIHSS (E-NIHSS) videos.

Results—The pass rates using PC-NIHSS and CC-NIHSS videos were 79% and 82%, respectively. All possible responses on individual scale items were included. Facial palsy and limb ataxia (13%) showed poor agreement, nine (60%) to 10 (67%) items showed moderate agreement ($0.4 < \kappa < 0.75$), and three (20%) to four (27%) items showed excellent agreement. When compared with E-NIHSS videos, the agreements on best gaze, visual fields, facial weakness and aphasia were less for PC-NIHSS videos, and the agreements on commands for level of consciousness and visual fields were less for CC-NIHSS videos. Nevertheless, there was no difference between PC-NIHSS or CC-NIHSS and E-NIHSS videos in the agreement total score.

*Correspondence: Raymond T.F. Cheung, Lee Man-Chiu Endowed Professor in Neuroscience, Professor of Neurology and Director of Acute Stroke Services, Department of Medicine, 4/F Professorial Block, Queen Mary Hospital, 102 Pokfulam Road, Hong Kong. rtcheung@hkucc.hku.hk.

There are no conflicts of interest to disclose.

Conclusions—Compared with E-NIHSS videos, PC-NIHSS and CC-NIHSS videos show good content validity and inter-rater reliability. Availability of these videos may facilitate the proper use of NIHSS among physicians and nurses in Putonghua- or Cantonese-speaking communities.

Keywords

Cantonese; Chinese; Mandarin; Putonghua; reliability; stroke assessment; stroke outcome measures; stroke scales

Introduction

Multicentre clinical trials on novel stroke therapies require a reliable, valid and time-efficient rating scale to measure stroke deficits (1–8). The NIHSS was developed for use in clinical trials, notably the landmark clinical stroke trials on recombinant tissue plasminogen activator (9). NIHSS contains 15 items, including level of consciousness, eye movement, visual field, facial movement, limb power, coordination, sensation, language, speech and neglect (3,10). The scale items are scored according to severity using predefined, weighted scores. NIHSS has become an integral part of the acute stroke assessment in both research and clinical settings (11–13).

A formal training and certification process, using validated videos in English, is required to ensure the proper application of NIHSS. Original videos were created in 1988 and made into videotapes for training and certification (3). New training and certification videos were created in 2003 and saved on DVD (10). In addition, real-time, online training and certification for NIHSS have been available since 2005 (<http://nihss-english.trainingcampus.net>). The target audiences for this accredited programme of NIHSS training and certification are stroke care providers and researchers. The instructions have been translated and dubbed from English into Italian and Portuguese, Spanish, German, French, Russian, Bulgarian, Czech, Dutch and Romanian have been completed or planned (<http://www.nihstrokescale.org>). Nevertheless, the set of tools in English for assessment of language and speech may not be appropriate for languages other than English, voice-over of the patients' verbal responses is artificial, and interpretation of impairments in language and speech may be different between English and languages other than English.

China has one-fifth of the world's population and approximately 1.05 billion Chinese speak Mandarin either as a first or second language (<http://www.nvtc.gov/lotw/months/november/worldlanguages.htm>). Mandarin is the official language of mainland China and Taiwan as well as an official language of Singapore. Mandarin is officially known as Putonghua in mainland China because the word 'Putonghua' means 'common language'. Around 71 million people speak Cantonese as their first language in Guangdong province, Guangzhou city (Canton), Hong Kong and Macau, as well as in expatriate Chinese communities globally (<http://www.nvtc.gov/lotw/months/may/Cantonese.htm>). To facilitate and encourage the proper use of NIHSS among physicians and nurses in communities speaking either Putonghua or Cantonese, we produced a new set of Putonghua-Chinese language NIHSS (PC-NIHSS) training and certification videos and a new set of Cantonese-Chinese language NIHSS (CC-NIHSS) training and certification videos, and compared with the original English language NIHSS (E-NIHSS) videos regarding the level of inter-rater reliability. Here we report our experience on NIHSS training and certification using new PC-NIHSS or CC-NIHSS videos.

Methods

A professional video production company was appointed in June 2006 to create new Chinese language versions of NIHSS training and certification videos in both Putonghua and Cantonese. First, original scripts of E-NIHSS training and certification videos on DVDs were translated from English into Putonghua- or Cantonese-Chinese (10). Second, language assessment (item 9) requires a patient to describe the cookie jar picture, name the objects from the object card, and read the sentences from the sentence card. To assess aphasia in a Putonghua-speaking patient, the object card was modified to include objects familiar to Chinese speakers, and avoid objects that would be unknown outside of Western countries, e.g. 'cactus' and 'hammock'. Appropriate sentences in Chinese were included in the sentence card. Evaluation of articulation (item 10) requires an English-speaking patient to read words that include all phonemes used in English language (<http://www.nihstrokescale.org>). The word card was modified extensively to include phonemes commonly used by native Putonghua speakers, and exclude phonemes used only by English speakers. These cards were previously used in a multicentre stroke trial conducted in mainland China (14). Similar modifications were made to these cards to allow proper assessment of aphasia in a Cantonese-speaking patient; these cards were previously used in a multicentre stroke trial conducted in Hong Kong (15).

Third, filming of CC-NIHSS videos took place in the Neurology Wards at Queen Mary Hospital and Pamela Youde Eastern Hospital based in Hong Kong over 5 days from July 2006 to September 2006. Filming of PC-NIHSS videos took place in Neurology Wards of Peking University First Hospital, Beijing, China from 25–29 September 2006. Similar to the production of E-NIHSS videos on DVDs, stroke patients were selected so that every choice in every scale item would be illustrated at least once and that their stroke severity was similar to that of E-NIHSS videos. In general, two cameras captured the video from both the front and side with scrupulous attention to sound and lighting to optimise video illustration of the proper technique and findings (3,10). Consent for videotaping was obtained from all patients.

NIHSS examinations were videotaped from a total of 30 native Putonghua-speaking patients and another 30 native Cantonese-speaking patients. For PC-NIHSS videos, 12 cases were used in the instruction and demonstration sections, and remaining 18 cases were chosen for certification purpose. Similar case selection was done for CC-NIHSS videos. Thus, both PC-NIHSS and CC-NIHSS videos contain three groups of certification cases with six patients per group, being balanced for severity and stroke (10).

Three bilingual stroke neurologists certified previously and experienced in E-NIHSS independently scored all items of the 18 certification cases of both PC-NIHSS and CC-NIHSS. Upon disagreement on the score of an item, the video segment was reviewed to generate the acceptable score/s of the item. For PC-NIHSS videos, a training and certification workshop was held in Chengdu, China, 27 August 2007 and another one was held in Beijing, China on 1 September 2007. In addition, two training and certification workshops were held in Hong Kong in June 2008 using CC-NIHSS videos. After a brief introduction, the training and group A certification videos were presented over two sessions on the same day of the workshop. Certification was granted to the raters according the standard scoring rules (3).

Numerical data were described in mean±SD and categorical variables were given as numbers with percentages in parentheses. Demographic and other information of raters who passed the certification were compared with those who failed using Student's *t*-test or χ^2 -test. Multiple logistic regression analysis identified the independent predictors for a pass in the certification workshops using PC-NIHSS or CC-NIHSS videos. To allow for comparison to E-NIHSS videos, unweighted κ statistics for multiple examiners were used to study the inter-rater

agreement on individual items (16) of PC-NIHSS or CC-NIHSS videos. Our interpretation of the magnitude of κ coefficient is as follows:

$\kappa < 0.40$ defines poor agreement,

$0.40 \leq \kappa \leq 0.75$ defines moderate agreement, and

$\kappa > 0.75$ defines excellent agreement.

The 95% bootstrap confidence interval (CI) was also generated (17). The κ coefficient of each item of PC-NIHSS or CC-NIHSS videos was compared with that of E-NIHSS videos on DVDs (10). If κ coefficient for an item of E-NIHSS videos is included within the 95% CI of κ coefficient of the corresponding item of PC-NIHSS or CC-NIHSS videos, we would conclude that there is no statistically significant difference between the two κ coefficients. Intraclass correlation coefficient (ICC) was used to study the agreement on the total score (18). An ICC value close to 1.0 indicates excellent agreement. The 95% bootstrap CI was also generated (17). ICC of E-NIHSS videos was compared with that of PC-NIHSS or CC-NIHSS videos using Fisher's transformation (19).

Results

Altogether 310 doctors and nurses from 34 acute hospitals attended the two workshops in mainland China, and 295 raters submitted their scores for group A cases of PC-NIHSS videos. Item scores of group A cases were available for analyses from 295 raters, including 188 physicians, 78 residents and 29 nurses. We received 26 543 responses from an expected maximum of 27 900 responses (15 items $\times$ 6 patients $\times$ 310 attendees), giving a response rate of 95%. Using group A cases of PC-NIHSS videos, the pass rate was 79%. Demographic and other important information of these 295 raters is summarised in Table 1. Univariate analyses showed that passing the certification using PC-NIHSS videos was associated with medical background and previous experience in E-NIHSS. Multiple logistic regression analysis revealed occupation as the independent predictor ($P = 0.008$) with nurses having a lower pass rate (odds ratio 0.244, 95% CI 0.096–0.623).

For the workshops conducted in Hong Kong using CC-NIHSS videos, 294 nurses and 26 physicians from 14 acute public hospitals (271; 85%), six convalescent/rehabilitation hospitals (33; 10%), and other institutions (16; 5%) attended. The item scores from 290 nurses and 24 physicians were available for analyses. The response rate was 98% (28 259 responses from an expected maximum of 28 800 responses). Using CC-NIHSS training and certification videos, the overall pass rate was 82% (Table 2). Univariate analyses showed that passing the certification using CC-NIHSS videos was associated with an age < 30 years and a shorter duration of working experience. Nevertheless, multiple logistic regression analysis did not identify any independent predictor.

Table 3 shows the range of scores obtained from each item of all six cases of PC-NIHSS videos. Except for a score of 3 on item 1a, all possible responses on individual scale items were included in group A certification cases of PC-NIHSS videos. The total score of these six cases, in mean $\pm$ SD, ranged from 1.8 ± 1.2 to 29.3 ± 2.6 , and, in median, from 2–29. (Table 4)

For PC-NIHSS videos, facial palsy and limb ataxia (13%) showed poor agreement, nine items (60%) showed moderate agreement, and four items (27%) showed excellent agreement. When compared with E-NIHSS videos, the agreements on best gaze, visual fields, facial weakness or aphasia were less well with PC-NIHSS videos; the agreement on consciousness questions was better in PC-NIHSS videos (Table 5). Intraclass correlation coefficient of total score of PC-NIHSS videos was 0.97 (95% CI 0.64–0.98), and was not statistically different from E-NIHSS videos (0.94; $P = 0.494$) (10). For CC-NIHSS videos, facial palsy and limb ataxia (13%)

showed poor agreement, 10 items (67%) showed good agreement, and three items (20%) showed excellent agreement. When compared with E-NIHSS videos, the agreements on consciousness commands and visual fields were less well in CC-NIHSS videos (Table 5). ICC of total score of CC-NIHSS videos was 0.95 (95% CI 0.77–0.97), and this was not statistically different from that of E-NIHSS videos ($P = 0.906$) (10).

Discussion

All possible responses on individual scale items, except for a score of 3 on item 1a, are included in PC-NIHSS group A cases. The rules to give a score of 3 on item 1a are nonambiguous; this item score is included in a PC-NIHSS nongroup A case. Except for a score of 2 on item 5a and a score of 4 on items 5b and 6b, all possible responses on all individual scale items are included in group A cases of CC-NIHSS videos. Such omissions are acceptable as all possible responses are illustrated when motor items of all four limbs are considered together. Thus, group A certification cases of both PC-NIHSS and CC-NIHSS videos have content validity.

In terms of inter-rater reliability of PC-NIHSS and CC-NIHSS videos, the agreements are good or excellent on 13 out of 15 items, but poor on facial palsy (item 4) and limb ataxia (item 7). These findings replicate those of E-NIHSS videos: facial palsy and limb ataxia (13%) have poor agreement; 11 items (73%) have good agreement; and two items (13%) have excellent agreement (10). When κ coefficients of individual scale items are compared between PC-NIHSS and E-NIHSS videos, the agreements on best gaze, visual fields, facial weakness and aphasia are less well but the agreement on consciousness questions is better in PC-NIHSS group A cases. Similarly, the agreement in commands for the level of consciousness and visual fields may be less well in the group A cases of CC-NIHSS videos than that of the 18 certification cases of E-NIHSS videos. At present we have no explanation for such differences. Groups B and C cases of PC-NIHSS and CC-NIHSS videos should be evaluated in future workshops to allow comparison of κ coefficients of scale items between the groups B and C cases and E-NIHSS certification cases. There is no difference in ICC of total score between PC-NIHSS and E-NIHSS videos as well as between CC-NIHSS and E-NIHSS videos.

Fewer nurses than doctors attended the workshops using PC-NIHSS videos as assessment of stroke deficits, such as NIHSS, is performed by medical doctors but not nurses in Mainland China. In contrast, nurses in Hong Kong are willing to perform assessments such as Glasgow coma scale and NIHSS. Raters of the workshops in Mainland China are younger than those in Hong Kong. Nevertheless, workshop attendees are typical stroke care providers in Mainland China and Hong Kong, and most raters have had working experience in the hospital and with stroke patients.

Whereas a third to half of raters have previously performed NIHSS assessment on their patients, only a tenth were formally trained using E-NIHSS videos; language barrier is a possible explanation. A high pass rate of 80% is obtained using PC-NIHSS or CC-NIHSS videos, suggesting an advantage of using the native language in training and certification of stroke assessment scales. Another study involving naïve users will be required to show whether training with PC-NIHSS or CC-NIHSS videos would facilitate certification among healthcare workers speaking Putonghua or Cantonese when compared with E-NIHSS training videos. Nevertheless, previous formal training using E-NIHSS videos is not associated with a pass. A plausible explanation is the small number of raters with formal training.

Both PC-NIHSS and CC-NIHSS videos have good content validity and inter-rater reliability comparable to those of E-NIHSS videos. Similar to E-NIHSS videos, PC-NIHSS and CC-NIHSS videos will be saved on DVDs and made available online for single user to view at home or in office as well as for multiple users to view in a training and certification workshop.

Availability of PC-NIHSS and CC-NIHSS videos should lead to wider use of NIHSS among stroke care providers in Putonghua- and Cantonese-speaking communities. If the current approach turns out to be superior to simple translations of E-NIHSS videos from English into non-English languages, there may be a need to create new NIHSS training and certification videos in other widely spoken languages.

Acknowledgments

This paper was supported by S. K. Yee Medical Foundation, Hong Kong (project no. 205257).

References

1. Brott T, Adams HP Jr, Olinger CP, et al. Measurements of acute cerebral infarction: a clinical examination scale. *Stroke* 1989;20:864–70. [PubMed: 2749846]
2. Goldstein LB, Bartels C, Davis JN. Interrater reliability of the NIH Stroke Scale. *Arch Neurol* 1989;46:660–2. [PubMed: 2730378]
3. Lyden P, Brott T, Tilley B, et al. Improved reliability of the NIH stroke scale using video training. *Stroke* 1994;25:2220–6. [PubMed: 7974549]
4. Albanese MA, Clarke WR, Adams HP Jr, Woolson RF. Ensuring reliability of outcome measures on multicenter clinical trials of treatments for acute ischemic stroke: the program developed for the trial of ORG 10172 in acute stroke treatment (TOAST). *Stroke* 1994;25:1746–51. [PubMed: 8073453]
5. Goldstein L, Samsa G. Reliability of the National Institutes of Health Stroke Scale. *Stroke* 1997;28:307–10. [PubMed: 9040680]
6. Kasner SE, Chalela JA, Luciano JM, et al. Reliability and validity of estimating the NIH Stroke Scale score from medical records. *Stroke* 1999;30:1534–7. [PubMed: 10436096]
7. Lyden PD, Lu M, Jackson C, et al. Underlying structure of the National Institutes of Health Stroke Scale: results of a factor analysis. *Stroke* 1999;30:2347–54. [PubMed: 10548669]
8. Lyden PD, Lu M, Levine S, Brott TG, Broderick J. A modified national institutes of health stroke scale for use in stroke clinical trials. Preliminary reliability and validity. *Stroke* 2001;32:1310–7. [PubMed: 11387492]
9. The National Institutes of Neurological Disorders and Stroke rt-PA Stroke Study Group. Tissue plasminogen activator for acute ischemic stroke. *N Engl J Med* 1995;333:1581–7. [PubMed: 7477192]
10. Lyden P, Raman R, Liu L, et al. NIHSS training and certification using a new digital video disk is reliable. *Stroke* 2005;36:2446–9. [PubMed: 16224093]
11. Adams HP Jr, del Zoppo G, Alberts MJ, et al. Guidelines for the early management of adults with ischemic stroke: a guideline from the American Heart Association/American Stroke Association Stroke Council, Clinical Cardiology Council, Cardiovascular Radiology and Intervention Council, and the Atherosclerosis Peripheral Vascular Disease and Quality of Care Outcomes in Research Interdisciplinary Working Groups. *Stroke* 2007;38:1655–711. [PubMed: 17431204]
12. The European Stroke Organization (ESO) Executive Committee and the ESO Writing Committee. Guidelines for management of ischaemic stroke and transient ischaemic attack 2008. *Cerebrovasc Dis* 2008;25:457–507. [PubMed: 18477843]
13. Hacke W, Kaste M, Bluhmki E, et al. Thrombolysis with alteplase 3 to 4.5 hours after acute ischaemic stroke. *N Engl J Med* 2008;359:1317–29. [PubMed: 18815396]
14. Huang Y, Cheng Y, Wu J, et al. Cilostazol as an alternative to aspirin after ischaemic stroke: a randomised, double-blind, pilot study. *Lancet Neurol* 2008;7:494–9. [PubMed: 18456558]
15. Wong KS, Chen C, Ng PW, et al. Low-molecular-weight heparin compared with aspirin for the treatment of acute ischaemic stroke in Asian patients with large artery occlusive disease: a randomised study. *Lancet Neurol* 2007;6:407–13. [PubMed: 17434095]
16. Fleiss, JL. *Statistical Methods for Rates and Proportions*. 2. New York: John Wiley & Sons; 1981.
17. Efron, B.; Tibshirani, RJ. *An Introduction to the Bootstrap*. New York: Chapman & Hall/CRC; 1993.
18. Shoukri, MM. *Measures of Interobserver Agreement*. Boca Raton: Chapman & Hall/CRC; 2004.
19. Zar, JH. *Biostatistical Analysis*. 4. New Jersey: Prentice Hall; 1999. p. 390-2.

Table 1

Demographics and other information of raters of PC-NIHSS videos*

Characteristics	All raters	Passed raters	Failed raters	P value
	N =295	N =233	N =62	
Age (years)	30.1±7.2	30.0±7.1	30.3±7.6	0.740
Age ≥30 years	119 (40.3)	95 (40.8)	24 (38.7)	0.884
Male gender	114 (38.6)	92 (39.5)	22 (35.5)	0.660
Occupation groups				0.001
Physician	188 (63.7)	154 (66.1)	34 (54.8)	
Resident	78 (26.4)	64 (27.5)	14 (22.6)	
Nurse	29 (9.8)	15 (6.4)	14 (22.6)	
Working experience (years)	6.6±7.7	6.4±7.6	7.3±8.0	0.377
Working experience in groups				0.736
No experience	80 (27.1)	65 (27.9)	15 (24.2)	
1–5 years	94 (31.9)	75 (32.2)	19 (30.6)	
>5 years	121 (41.0)	93 (39.9)	28 (45.2)	
Experience with stroke patient				0.601
No experience	34 (11.5)	25 (10.7)	9 (14.5)	
1–5 years	152 (51.5)	123 (52.8)	29 (46.8)	
>5 years	109 (36.9)	85 (36.5)	24 (38.7)	
Experience with NIHSS	128 (43.4)	109 (46.8)	19 (30.6)	0.030
Trained with E-NIHSS videos	24 (8.1)	21 (9.0)	3 (4.8)	0.433

* Numerical variables are in mean±SD and categorical variables in number with percentages in parentheses. The numbers do not add to the total because of missing data. The percentages do not add to 100 because of rounding. Comparisons between raters who passed and those who failed are made using Student's *t*-test or χ^2 -test.

E-NIHSS, English language NIHSS; N, number of raters; NIHSS, National Institutes of Health Stroke Scale; PC-NIHSS, Putonghua-Chinese NIHSS.

Table 2

Demographics and other information of the raters of CC-NIHSS videos*

Characteristics	All raters <i>N</i> =314	Passed raters <i>N</i> =257	Failed raters <i>N</i> =57	<i>P</i> value
Age (years)	37.0±8.5	36.5±8.5	38.9±8.2	0.057
Age ≥30 years	243 (77.4)	193 (75.1)	50 (87.7)	0.039
Male gender	49 (15.6)	36 (14.0)	13 (22.8)	0.098
Occupation groups				0.723
Physician	24 (7.6)	19 (7.4)	5 (8.8)	
Nurse	290 (92.4)	238 (92.6)	52 (91.2)	
Working experience (years)	14.6±8.4	14.2±8.4	16.6±8.1	0.044
Working experience in groups				0.124
1–5 years	49 (15.7)	44 (17.2)	5 (8.9)	
>5 years	263 (84.3)	212 (82.8)	51 (91.1)	
Experience with stroke patient				0.599
No experience	6 (1.9)	4 (1.6)	2 (3.6)	
1–5 years	111 (36.0)	91 (36.0)	20 (36.4)	
>5 years	191 (62.0)	158 (62.5)	33 (60.0)	
Experience with NIHSS	110 (35.0)	92 (35.8)	18 (31.6)	0.546
Trained with E-NIHSS videos	34 (10.8)	31 (12.1)	3 (5.3)	0.133

* Numerical variables are in mean±SD and categorical variables in number with percentages in parentheses. The numbers do not add to the total because of missing data. The percentages do not add to 100 because of rounding. Comparisons between raters who passed and those who failed are made using Student's *t*-test or χ^2 -test.

E-NIHSS, English language NIHSS; *N*, number of raters; NIHSS, National Institutes of Health Stroke Scale; CC-NIHSS, Cantonese-Chinese NIHSS.

Table 3

Distribution of responses by individual items of PC-NIHSS videos*

Item number and name	Total responses on the item	Level of responses, number (%)				
		0	1	2	3	4
1a LOC	1770	1474 (83.3)	274 (15.5)	22 (1.2)	0 (0.0)	—
1b LOC questions	1769	898 (50.8)	288 (16.3)	583 (33.0)	—	—
1c LOC command	1770	1481 (83.7)	286 (16.2)	3 (0.2)	—	—
2 Gaze	1770	1235 (69.8)	354 (20.0)	181 (10.2)	—	—
3 Visual fields	1770	1415 (79.9)	84 (4.7)	132 (7.5)	139 (7.9)	—
4 Facial weakness	1770	678 (38.3)	538 (30.4)	285 (16.1)	269 (15.2)	—
5a Motor left arm	1770	1283 (72.5)	188 (10.6)	7 (0.4)	4 (0.2)	288 (16.3)
5b Motor right arm	1770	629 (35.5)	716 (40.5)	391 (22.1)	30 (1.7)	4 (0.2)
6a Motor left leg	1770	1338 (75.6)	136 (7.7)	3 (0.2)	4 (0.2)	289 (16.3)
6b Motor right leg	1769	602 (34.0)	576 (32.6)	237 (13.4)	73 (4.1)	281 (15.9)
7 Ataxia	1768	538 (30.4)	656 (37.1)	574 (32.5)	—	—
8 Sensory	1767	1145 (64.8)	619 (35.0)	3 (0.2)	—	—
9 Aphasia	1770	756 (42.7)	727 (41.1)	276 (15.6)	11 (0.6)	—
10 Dysarthria	1770	1031 (58.2)	638 (36.0)	101 (5.7)	—	—
11 Extinction	1770	1270 (71.8)	336 (19.0)	164 (9.3)	—	—

* This table shows 15 NIHSS items and level of responses to each item by 295 raters on PC-NIHSS group A cases. Unequal total responses to items are due to missing values. The percentages do not add to 100 because of rounding. Level of responses corresponds to the available responses for each item; some items have three and others have four or five possible responses.

—, response unavailable;

PC-NIHSS, Putonghua-Chinese NIHSS.

Table 4

Distribution of responses by individual items of CC-NIHSS videos*

Item number and name	Total responses on the item	Level of responses, number (%)				
		0	1	2	3	4
1a LOC	1884	1307 (69.4)	540 (28.7)	36 (1.9)	1 (0.1)	–
1b LOC questions	1884	340 (18.0)	1397 (74.2)	147 (7.8)	–	–
1c LOC command	1884	1559 (82.7)	257 (13.6)	68 (3.6)	–	–
2 Gaze	1883	994 (52.8)	755 (40.1)	134 (7.1)	–	–
3 Visual fields	1884	1320 (70.1)	374 (19.9)	161 (8.5)	29 (1.5)	–
4 Facial weakness	1884	368 (19.5)	526 (27.9)	716 (38.0)	274 (14.5)	–
5a Motor left arm	1884	1335 (70.9)	236 (12.5)	0 (0.0)	30 (1.6)	283 (15.0)
5b Motor right arm	1884	732 (38.9)	1119 (59.4)	32 (1.7)	1 (0.1)	0 (0.0)
6a Motor left leg	1884	1080 (57.3)	481 (25.5)	9 (0.5)	2 (0.1)	312 (16.6)
6b Motor right leg	1884	647 (34.3)	908 (48.2)	320 (17.0)	9 (0.5)	0 (0.0)
7 Ataxia	1884	948 (50.3)	552 (29.3)	384 (20.4)	–	–
8 Sensory	1884	1551 (82.3)	323 (17.1)	10 (0.5)	–	–
9 Aphasia	1884	433 (23.0)	831 (44.1)	611 (32.4)	9 (0.5)	–
10 Dysarthria	1884	426 (22.6)	1114 (59.1)	344 (18.3)	–	–
11 Extinction	1884	1496 (79.4)	186 (9.9)	202 (10.7)	–	–

* This table shows the 15 NIHSS items and the level of responses to each item by 314 raters on CC-NIHSS group A cases. Unequal total responses to items are due to missing values. The percentages do not add to 100 because of rounding. Level of responses corresponds to the available responses for each item; some items have three and others have four or five possible responses.

–, response unavailable;

CC-NIHSS, Cantonese-Chinese NIHSS.

Table 5

Interobserver agreement on individual scale items for PC-NIHSS or CC-NIHSS videos as compared to E-NIHSS videos*

Item number and name	PC-NIHSS		CC-NIHSS		E-NIHSS κ
	κ	CI	κ	CI	
1a LOC	0.49	0.14–0.64	0.49	0.17–0.73	0.46
1b LOC questions	0.97	0.91–0.99	0.74	0.00–0.98	0.77
1c LOC command	0.93	0.00–0.94	0.77	0.00–0.79	0.92
2 Gaze	0.60	0.04–0.67	0.70	0.06–0.87	0.70
3 Visual fields	0.55	0.00–0.65	0.49	0.02–0.57	0.72
4 Facial weakness	0.31	0.08–0.35	0.28	0.04–0.44	0.38
5a Motor left arm	0.72	0.00–0.92	0.75	0.01–0.87	0.65
5b Motor right arm	0.61	0.26–0.77	0.60	0.03–0.73	0.72
6a Motor left leg	0.64	0.01–0.87	0.74	0.20–0.91	0.64
6b Motor right leg	0.77	0.55–0.84	0.73	0.43–0.82	0.64
7 Ataxia	0.28	0.08–0.42	0.29	0.12–0.38	0.21
8 Sensory	0.85	0.01–0.92	0.87	0.00–0.94	0.73
9 Aphasia	0.42	0.07–0.58	0.49	0.07–0.69	0.64
10 Dysarthria	0.45	0.09–0.56	0.70	0.07–0.87	0.56
11 Extinction	0.59	0.00–0.64	0.55	0.01–0.68	0.57

* Table shows the agreement (unweighted κ) and 95% CI among raters for 15 NIHSS items using PC-NIHSS or CC-NIHSS videos. $\kappa < 0.40$ defines poor agreement, $0.40 \leq \kappa \leq 0.75$ defines moderate agreement, and $\kappa > 0.75$ defines excellent agreement. The agreements on individual items among raters using E-NIHSS videos are given for comparison (10).

E-NIHSS, English language NIHSS; CC-NIHSS, Cantonese-Chinese NIHSS; PC-NIHSS, Putonghua-Chinese NIHSS.