Abstract

Introduction: The National Institutes of Health Stroke Scale (NIHSS), originally designed in the United States of America, contains items on dysphasia and dysarthria that are deemed culturally unsuitable for the Singapore context. We compared the error rates of dysphasia objects, dysphasia phrases and dysarthria words between the original and alternative items in a cohort of Singaporean subjects without dysphasia or dysarthria.

Methods: In this prospective study, 140 English-speaking Singaporean subjects without impairments of dysphasia or dysarthria had an assessment of NIHSS items 9 and 10 using the original and alternative items. Paired analyses were conducted for comparison of error rates.

Results: The error rates were high for four original dysphasia objects (Hammock: 62.9%, Cactus: 38.6%, Feather: 23.6%, Glove: 20.7%) and significantly lower for alternative items (Snail: 5%, Horse: 1.4%, Hanger: 1.4%, Car: 0%) ($P < 0.001$). For dysphasia phrases and dysarthria words, the error rates were low and there were no differences in error rates between the original and alternative items.

Conclusion: There are cultural issues with several dysphasia objects in the original NIHSS as evidenced by the high error rates, which were lowered with more culturally suitable alternatives. This study formed a basis to derive a more suitable version of the NIHSS for English-speaking subjects in Singapore.

Keywords: Culture, dysarthria, dysphasia, National Institutes of Health Stroke Scale, Singapore
Individuals without dysphasia or dysarthria should have no or minimal errors on NIHSS items 9 and 10. We compared the error rates of dysphasia objects, dysphasia phrases and dysarthria words between the original and alternative items, and investigated the factors associated with error rates with the original NIHSS items in a cohort of Singaporean subjects without dysphasia or dysarthria.

**METHODS**

This was a prospective observational study. Patients from the neurology inpatient and outpatient service at a tertiary hospital in Singapore were screened at random and recruited from January 2016 to August 2016. We recruited subjects without dysphasia or dysarthria based on the following criteria: (a) exclusion of stroke from clinical assessment; (b) of stroke-prone age (≥60 years); and (c) able to converse in English. Subjects with physical barriers to speech (e.g., cleft palates), known or suspected cognitive impairment or who were non-communicative for other reasons (e.g., hearing impaired) were excluded.

The selection of alternative items for NIHSS was done by consensus of a multidisciplinary team consisting of a senior speech therapist, a neurologist, an advanced practice nurse and a medical student. For dysphasia objects, Hammock and Cactus were considered low-frequency objects (less common) by consensus and alternative low frequency objects were identified [Table 1]. Two higher-frequency (more common) objects were also identified and tested. Illustrations of alternative objects were retrieved from snodgrass’ standardised set of 260 pictures. For dysphasia phrases and dysarthria words, alternatives were selected to retain the number of syllables and phonetic structure.

The presence or absence of dysphasia and dysarthria in subjects was determined by the managing medical team. Each subject was administered items 9 and 10 of NIHSS using the original and alternative dysphasia objects, dysphasia phrases and dysarthria words in the same encounter by one assessor [Table 1]. All assessors had NIHSS certification by the American Stroke Association (Colorado, USA).[14] In a sample of 21 patients recruited in the first month, two assessors performed independent assessment to determine the inter-rater reliability. The inter-rater reliability for composite error rate of 26 items (17 original and 9 alternatives) between the two independent assessors for the sample of 21 cases was 0.984 (P < 0.001). Paired comparison of error rates between the original and alternative items for dysphasia objects, dysphasia phrases and dysarthria words was performed using SPSS version 23.0 (IBM Corp, Armonk, NY, USA). The alpha level for significance was set at 0.05.

**RESULTS**

Of 294 patients screened, 140 patients met the inclusion criteria [Figure 1]. The median age of the patients was 66 (interquartile range [IQR] 63–71) years and 54.3% were male.

**Table 1. Original and alternative items for dysphasia objects, dysphasia phrases and dysarthria words.**

| Original | Alternatives |
|----------|--------------|
| **Dysphasia objects** | |
| Low frequency: | Low frequency: |
| 1. Cactus | 1. Horse |
| 2. Hammock | 2. Snail |
| High frequency: | |
| 1. Car | 2. Hanger |
| **Dysphasia phrases** | |
| 1. Down to earth | 1. Put it back |
| 2. They heard him speak on the radio last night | 2. We heard him speak at the wedding last night |
| **Dysarthria words** | |
| 1. Tip-Top | 1. Tick-Tock |
| 2. Huckleberry | 2. Strawberry Cake |
| 3. Baseball Player | 3. Football Player |

**Figure 1: Consolidated Standards of Reporting Trials diagram of patients of stroke-prone age without dysphasia or dysarthria.**

The ethnic distribution of Chinese, Malay, Indian and other ethnicities was 70.0%, 10.7%, 15.0% and 4.3%, respectively. The highest educational level profile was as follows: 0.7% with no formal education, 12.9% at primary level, 53.6% at secondary level, 22.1% at preuniversity/vocational level, and 10.7% at university level and above.

The error rates for all original and alternative objects, phrases and words were obtained [Table 2]. The error rates for dysphasia objects in the original NIHSS that were
identified as culturally unsuitable by the healthcare worker survey were 62.9% for Hammock and 38.6% for Cactus. High error rates were also found in items that were not originally identified in the healthcare worker survey as culturally unsuitable, namely 23.6% for Feather and 20.7% for Glove. For the alternative items, the error rates were 5% for Snail, 1.4% for Horse, 1.4% for Hanger and 0% for Car. In the comparison made between similar-frequency original and alternative dysphasia objects, there were error rates for all alternative objects \( (P < 0.001) \) [Table 3]. The error rates for dysphasia phrases and dysarthria words were low for original and alternative items, with the highest error rate at 6.4% for ‘Tip-Top’ [Table 2]. There were no significant differences in error rates between the original and alternative items for dysphasia phrases and dysarthria words [Table 3]. 

In this sample of patients without dysphasia or dysarthria, having three or more errors on the original NIHSS for all components of dysphasia objects, dysphasia phrases and dysarthria words was associated with education level \( (P = 0.003) \) [Table 4]. There were no associations with age \( (P = 0.489) \), ethnicity \( (P = 0.396) \) and gender \( (P = 0.881) \).

### Table 2. Error rates of dysphasia objects, dysphasia phrases and dysarthria words.

| Dysphasia objects | Original | Alternatives |
|-------------------|----------|--------------|
| Hammock\(^{a}\)   | 62.9     |              |
| Cactus\(^{a}\)    | 38.6     |              |
| Feather           | 23.6     |              |
| Glove             | 20.7     |              |
| Key               | 0.7      |              |
| Chair             | 0.0      |              |
| Alternatives      |          |              |
| Snail             | 5.0      |              |
| Hanger            | 1.4      |              |
| Horse             | 1.4      |              |
| Car               | 0.0      |              |

| Dysphasia phrases | Original | Alternatives |
|-------------------|----------|--------------|
| Down to earth\(^{a}\) | 2.9     |              |
| They heard him speak on the radio last night\(^{a}\) | 1.4   |              |
| I got home from work | 3.6     |              |
| Near the table in the dining room | 2.1     |              |
| You know how      | 1.4      |              |
| Alternatives      |          |              |
| Put it back       | 2.1      |              |
| We heard him speak at the wedding last night | 1.4 |              |

| Dysarthria words  | Original | Alternatives |
|-------------------|----------|--------------|
| Tip-Top\(^{a}\)  | 6.4      |              |
| Baseball Player\(^{a}\) | 2.1 |              |
| Huckleberry\(^{a}\) | 1.4    |              |
| Thanks            | 1.4      |              |
| Fifty–Fifty       | 0.7      |              |
| Mama             | 0.0      |              |
| Alternatives      |          |              |
| Tick-Tock         | 2.9      |              |
| Strawberry Cake   | 2.1      |              |
| Football Player   | 0.0      |              |

\(^{a}\)Item identified as culturally unsuitable from the Singapore healthcare worker survey.

### Table 3. Comparison of error rates of dysphasia objects, dysphasia phrases and dysarthria words between the original and alternative items.

| Error rate (%) \((n = 140)\) | \(P\) |
|------------------------------|------|
| **Dysphasia objects**        |      |
| Low-frequency                 |      |
| Hammock (62.9)                |      |
| Horse (1.4)                   | \(<0.001^{*}\) |
| Cactus (38.6)                 |      |
| Snail (5.0)                   |      |
| High-frequency                |      |
| Feather (23.6)                |      |
| Car (0.0)                     | \(<0.001^{†}\) |
| Glove (20.7)                  |      |
| Hanger (1.4)                  |      |
| **Dysphasia phrases**         |      |
| They heard him speak on the radio last night (1.4) | 1.000 |
| We heard him speak at the wedding last night (1.4) |      |
| Down to earth (2.9)           |      |
| Put it back (2.1)             |      |
| **Dysarthria words**          |      |
| Tip-Top (6.4)                 |      |
| Tick-Tock (2.9)               | \(0.180\) |
| Huckleberry (1.4)             |      |
| Strawberry Cake (2.1)         |      |
| Baseball Player (2.1)         |      |
| Football Player (0.0)         | \(0.250\) |

\(^{*}\)Comparisons of Hammock vs. Horse, Hammock vs. Snail, Cactus vs. Horse, Cactus vs. Snail. \(^{†}\)Comparisons of Feather vs. Car, Feather vs. Hanger, Glove vs. Car, Glove vs. Hanger.

### Table 4. Associations of errors with the original NIHSS.

| Variable                        | \(0–2\) errors \((n = 108)\) | \(\geq3\) errors \((n = 32)\) | \(P\) |
|---------------------------------|-----------------------------|-----------------------------|------|
| Age\(^{a}\) (yr)                | 66.0 (62.3–71.0)            | 66.5 (63.0–75.8)            | \(0.489\) |
| Ethnicity                       |                             |                             | \(0.396\) |
| Chinese                         | 72.2                        | 62.5                        |      |
| Malay                           | 10.2                        | 12.5                        |      |
| Indian                          | 14.8                        | 15.6                        |      |
| Others                          | 2.8                         | 9.4                         |      |
| Gender                          |                             |                             | \(0.881\) |
| Male                            | 54.6                        | 53.1                        |      |
| Female                          | 45.4                        | 46.9                        |      |
| Highest education level         |                             |                             | \(0.003\) |
| No formal education             | 0                           | 3.1                         |      |
| Primary                         | 8.3                         | 28.1                        |      |
| Secondary                       | 52.8                        | 56.3                        |      |
| Preuniversity/vocational        | 25.9                        | 9.4                         |      |
| University                      | 13.0                        | 3.1                         |      |

\(^{a}\)Data presented as median (interquartile range). NIHSS: National Institutes of Health Stroke Scale
DISCUSSION

This study’s findings showed high error rates with dysphasia objects of Item 9 of the original NIHSS among subjects in Singapore without dysphasia or dysarthria, suggesting that the objects are culturally unsuitable. This is in concordance to an Irish study where unfamiliarity of original objects to a non-American population was postulated to have led to high error rates for several items of the original NIHSS. These error rates were lower with alternatives that are more familiar to the local population, showing that modification for the local context will allow for more accurate assessment. In contrast to dysphasia objects, all dysphasia phrases and dysarthria words in the original NIHSS had low error rates, even for those identified as culturally unsuitable by the healthcare worker survey. One possible explanation for this difference is that direct phrase reading and word repetition do not require prior familiarity, unlike object identification. However, it may also imply that the dysphasia phrases and dysarthria words in the original NIHSS are suitable for the Singapore culture.

There was inconsistency between the healthcare worker survey and performance of subjects in this study, which shows that the perception of healthcare workers may not accurately or adequately identify the suitable items in local or cultural context. In particular, for dysphasia objects, the ambiguity of the pictures depicting the objects in the original version could have been a contributory factor for high error rates. Thus, future studies investigating the cultural issues of NIHSS should involve performance by patients rather than by the perception of healthcare workers, and further modified versions should be derived from objective findings in a population of subjects. Based on this study’s findings, we derived a modified NIHSS version for English-speaking subjects in Singapore, modifying the dysphasia objects while keeping the dysphasia phrases and dysarthria words unchanged. This is now being utilised in clinical practice. Further studies are needed to assess this modified NIHSS for its ability to discriminate between unimpaired patients and patients with dysphasia or dysarthria.

The strength of this study is that it is the first to compare the error rates of all components in Item 9 and Item 10 (dysphasia objects, dysphasia phrases and dysarthria words) in the original NIHSS versus alternative items. The main limitation is the lack of blinding of the assessor to the patients’ diagnosis of absence of stroke before administration of NIHSS, which could lead to potential assessor bias. This was mitigated in two ways. First, we ensured that there was good inter-rater reliability in a selected sample. Second, there was a predefined objective procedure for assessment to reduce subjectivity and ambiguity. For example, only one correct response was allowed for identification of dysphasia objects, and every word from the dysphasia phrases had to be correctly read. We acknowledge that NIHSS assessment for items 9 and 10 are based on an overall impression of the performance on all components and not on a numeric threshold of incorrect responses, and in this study, we evaluated the individual components rather than the composite subjective performance, which is difficult to quantify. Furthermore, owing to the limited sample size, the findings of the study may not be representative of the Singapore population.

In conclusion, this study found high error rates for some original dysphasia objects from the NIHSS among Singaporean subjects without dysphasia or dysarthria. These error rates were lowered with alternative objects that are more culturally appropriate. Findings from objective assessment were not consistent with the subjective opinions of healthcare workers. Thus, cultural issues and objective findings are important considerations with regards to the use of the dysphasia objects in the original NIHSS.

Financial support and sponsorship

The study received funding from the Duke-NUS Medical Student Research Fellowship Grant (AM-ETHOS01/FY2015/23-A23) and the National Neuroscience Institute Centre Grant Pilot Studies Research Grant (NCG PA08).

Conflicts of interest

There are no conflicts of interest.

REFERENCES

1. Brott T, Adams HP Jr, Olinger CP, Marler JR, Barsan WG, Biller J, et al. Measurements of acute cerebral infarction: A clinical examination scale. Stroke 1989;20:864-70.
2. Martin-Schild S, Siegler J, Kumar AD, Lyden P. Troubleshooting the NIHSS: Question-and-answer with one of the designers. Int J Stroke 2015;10:1284-6.
3. Oh MS, Yu KH, Lee JH, Jung S, Ko IS, Shin JH, et al. Validity and reliability of a Korean version of the National Institutes of Health Stroke Scale. J Clin Neurol 2012;8:177-83.
4. Cheung RTF, Lyden PD, Tsai TH, Huang Y, Liu M, Hon SF, et al. Production and validation of Putonghua- and Cantonese-Chinese language National Institutes of Health Stroke Scale training and certification videos. Int J Stroke 2010;5:74-9.
5. Dominguez R, Vila JF, Augustovski F, Irazola V, Castillo PR, Rotta Escalante R, et al. Spanish cross-cultural adaptation and validation of the National Institutes of Health Stroke Scale. Mayo Clin Proc 2006;81:476-80.
6. Cincu C, Pontes-Neto OM, Neville IS, Mendes HF, Menezes DF, Mariano DC, 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;27:119-22.
7. Sun TK, Chiu SC, Yeh SH, Chang KC. Assessing reliability and validity of the Chinese version of the stroke scale: Scale development. Int J Nurs Stud 2006;43:457-63.
8. Hussein HM, Abdel Moneim A, Emara T, Abd-Elhamid YA, Salem HH, Abd-Allah F, et al. Arabic cross cultural adaptation and validation of the National Institutes of Health Stroke Scale. J Neurol Sci 2015;357:152-6.
9. Prasad K, Dash D, Kumar A. Validation of the Hindi version of National Institute of Health Stroke Scale. Neurol India 2012;60:40-4.
10. Pezzella FR, Picconi O, De Luca A, Lyden PD, Fiorelli M. Development of the Italian version of the National Institutes of Health Stroke Scale: It-NIHSS. Stroke 2009;40:2557-9.
11. Burns M, Somers K, McElwaine P, Harbison J. National Institutes of Health Stroke Scale object naming test in a non-American English-speaking population. Int J Stroke 2014;9:E35.
12. QueckKK,FabbielaNL.,WoonFF,TayDD,OhCT,NgWM,etal.Culturalissues of the National Institutes of Health Stroke Scale dysphasia and dysarthria components in Singapore-A survey of healthcare workers. Int J Stroke 2016;11:NP93.
13. Snodgrass J, Vanderwart M. A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. J Exp Psychol Hum Learn 1980;6:174-215.
14. NIH Stroke Scale. AHA Learning Center. Available from: https://learn.heart.org/lms/nihss. [Last accessed on 2021 Sep 29].