The NIH Stroke Scale Has Limited Utility in Accurate Daily Monitoring of Neurologic Status

The Neurohospitalist
2016, Vol. 6(3) 97-101
© The Author(s) 2015
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1941874415619964
nhos.sagepub.com

Elisabeth B. Marsh, MD¹, Erin Lawrence, RN, MSN¹,
Rebecca F. Gottesman, MD, PhD¹, and Rafael H. Llinas, MD¹

Abstract

Background and Purpose: The National Institute of Health Stroke Scale (NIHSS) is rapid and reproducible, a seemingly attractive metric for the documentation of clinical progress in patients presenting with ischemic stroke. Many institutions have adopted it into daily clinical practice. Unfortunately, the scale may not adequately capture all forms of functional change. We evaluate its utility as a measure of recovery in patients treated with intravenous tissue plasminogen activator (IV tPA) for ischemic stroke. **Methods:** We prospectively evaluated the difference in the rate of improvement based on NIHSS (≥ 4 point change based on previous trials) versus physician-documented subjective and objective measures in 41 patients' status post IV tPA treatment. The NIHSS 24 hours posttreatment, on discharge, and at follow-up were compared to NIHSS on admission using tests of proportions and McNemar tests of paired data. Secondary analyses were performed defining significant improvement as NIHSS changes of 1 to 3 points. **Results:** The mean NIHSS improved from 9 to 6, 24 hours post-tPA. Of the 41 patients, 29 improved by physician documentation, although only 11 of the 29 met the NIHSS criteria ($P < .001$; McNemar $P < .001$). On discharge, 20 of the 41 patients met the NIHSS criteria; however, the proportion "better" by physician documentation (71%) remained significantly higher ($P = .04$; McNemar $P = .004$). The mean postdischarge follow-up NIHSS was 2. Twenty of the 21 patients improved by documentation versus 16 of the 21 by NIHSS ($P = .08$, McNemar $P = .125$). Using NIHSS changes of 1 to 3 increased sensitivity for detecting improvement but remained lower than physician documentation. **Conclusion:** The NIHSS has many advantages; however, it may miss functional changes when used in place of a comprehensive neurological examination to measure improvement poststroke.

Keywords

stroke, tissue plasminogen activator (tPA), outcomes, recovery

Introduction

The ability to efficiently document an accurate, reproducible assessment of neurologic function that can be followed over the course of a hospitalization is critical, particularly in patients presenting with ischemic stroke who must be monitored closely for functional improvement or decline. The National Institute of Health Stroke Scale (NIHSS) is one of the most commonly used measures of both initial stroke severity and response to treatment, particularly in the acute setting, and has many advantages. The scale was designed as a rapid, reproducible screening tool and is effective in this role. Many institutions have also incorporated it into their daily clinical practice as a way to evaluate patient progress over time. However, its utility in this role is less clear. It is well accepted that the NIHSS is not a substitute for a comprehensive neurological examination when considering the localization of a lesion or impact of subtle deficits. This is

predominantly due to the inherently coarse nature of the scale. The NIHSS can underrepresent the size of infarction for both posterior circulation and right hemisphere lesions^{1,2} and is more heavily weighted toward gross motor and language tasks, rendering it less sensitive to milder changes in domains such as distal finger movement or cognition that can have a significant impact on functional outcome.^{3,4}

There is a lack of standardization in the way neurologic function is monitored and recorded across institutions in the

¹ Department of Neurology, The Johns Hopkins School of Medicine, Baltimore, MD, USA

Corresponding Author:

Elisabeth B. Marsh, Department of Neurology, The Johns Hopkins School of Medicine, 600 North Wolfe St, Phipps 446C, Baltimore, MD 21287, USA.
Email: ebmarsh@jhmi.edu

inpatient setting. The NIHSS is used by nursing staff at many hospitals to document clinical status and determine when to notify providers of a change in examination. In prior studies, a change in NIHSS score of 4 or more was the threshold used to define significant deterioration due to symptomatic intracerebral hemorrhage.⁵ In the National Institute of Neurological Disorders and Stroke (NINDS) trial comparing intravenous tissue plasminogen activator (IV tPA) to placebo for the treatment of acute ischemic stroke, the same 4-point cutoff was also used to define clinically significant improvement.⁶ Despite its use in trials, the utility of the NIHSS to monitor and document functional improvement or decline as part of routine monitoring has not been well evaluated, and the ideal threshold for significant change is even less clear.

In this study, we evaluate the performance of the NIHSS in detecting clinical improvement in 41 patients treated with IV tPA for acute ischemic stroke during both the acute hospitalization and at follow-up. Given its limitations, we suspect the NIHSS will be significantly less sensitive than physician documentation at all time points, but the degree of insensitivity will help to determine the trade-off in using a standardized, well-known screening tool for clinical documentation.

Methods

Data Collection

The Johns Hopkins Institutional Review Board approved this study. A waiver of consent was granted, given that data were obtained primarily for quality assurance purposes, and data were prospectively collected for all patients receiving IV tPA for acute ischemic stroke at our tertiary referral, academic stroke center over a 14-month period. Charts were reviewed for variables of interest including patient demographics and stroke risk factors; stroke location and etiology (by Trial of Org 10172 in Acute Stroke Treatment [TOAST] criteria⁷); NIHSS: on admission, 24 hours post-tPA, on discharge, and at outpatient follow-up; and both subjective (documented opinion) and objective (change in neurological examination) improvement documented by the treating physician at each time point. A formal NIHSS was performed for all patients at the time of administration of IV tPA and at outpatient follow-up (1-3 months postdischarge). The physician-documented neurological examination was used to extrapolate the NIHSS 24 hours post-tPA and on discharge when no formal NIHSS was recorded.⁸ In order to avoid bias, 2 separate reviewers (EBM and RHL), blinded to NIHSS documentation (located separately within the chart), reviewed the daily progress notes for evidence of improvement at all time points. Improvement was defined as (1) the treating physician's documented expert opinion, typically reported within the assessment and/or (2) an objectively documented change in comprehensive neurologic examination within the physical examination section. When there was disagreement in whether the patient met the criteria for improvement, resolution was achieved by joint chart review ($n = 6$).

Data Analysis

Changes in NIHSS were calculated at each time point: 24 hours post-tPA, on discharge, and at 1- to 3-month follow-up. For our primary analysis, we evaluated "significant improvement" for each time point by the NIHSS criteria compared to physician documentation using a test of proportions and McNemar test of paired data. We defined significant improvement by NIHSS as a 4-point change in score based on the NINDS r-tPA trial. The sensitivity and specificity for the NIHSS' ability to detect improvement was subsequently calculated for each time point using physician documentation as the gold standard. In a secondary analysis, we redefined the NIHSS criteria using changes of 1, 2, and 3 points. The percentage agreement between NIHSS and provider documentation was calculated for each threshold.

Results

Approximately 250 patients were admitted to our institution with acute ischemia between July 2013 and September 2014. Forty-one (16.4%) received IV tPA, above the national average of 4% to 7%.⁹ All patients were followed over the course of their hospital stay for improvement. The average age of the cohort was 66 years. Sixty-one percent were women and 32% were African American. The mean time to treatment was 155 minutes. Nineteen were located within the right hemisphere and 2 in the posterior circulation. Fifteen were determined to be cardioembolic, 14 large vessel, and 5 lacunar infarcts. The mean NIHSS on presentation was 9 (median: 8) and improved to 6 (median: 5) 24 hours postadministration. Twenty-nine (71%) of the 41 patients were documented to be subjectively "better"; however, the average change in NIHSS score was 2 points, and only 11 (27% of the total population) met the NIHSS criteria for improvement (P vs clinical documentation $<.001$, McNemar $P < .001$; Figure 1). At discharge, the mean NIHSS score was 5 (median: 3). Twenty (49%) patients met criteria for improvement in this case, but this proportion remained significantly lower than those meeting criteria by physician documentation (71%; $P = .04$, McNemar $P = .004$). Twenty-one (51%) of the initial cohort returned for outpatient follow-up after hospital discharge (median time to follow-up = 1 month). The mean NIHSS at follow-up was 2 (median: 3). Ninety-five (20 of 21) percent of patients were subjectively better, whereas 16 (76%) of 21 had a greater than 4 point change in NIHSS score ($P = .08$, McNemar $P = .125$). The sensitivity of the NIHSS to detect improvement compared to physician documentation was 37.9% at 24 hours post-tPA, 68.9% at discharge, and 80% at follow-up, with a specificity of 100% at each time point.

When patients were stratified by lesion location (right hemisphere and posterior circulation vs left hemisphere), there was no difference in results. When patients were stratified by presenting NIHSS (divided at the median value of 8), those with lower scores showed a greater discrepancy in progress documented by NIHSS versus physician documentation for all

Figure 1. The National Institute of Health Stroke Scale (NIHSS) fails to capture a significant percentage of patients with physician-documented improvement after intravenous tissue plasminogen activator (IV tPA).

Table 1. Stratification of Proportion Improved by NIHSS.

		Post-tPA		Discharge		Follow-Up	
		Improved	P	Improved	P	Improved	P
Entire cohort	Physician reported	0.71	<.001	0.71	.004	0.95	.13
	NIHSS criteria	0.27		0.49		0.76	
NIHSS ≥ 8	Physician reported	0.59	.23	0.59	.85	0.89	1
	NIHSS criteria	0.41		0.62		0.89	
NIHSS < 8	Physician reported	0.84	<.001	0.84	.003	1	.03
	NIHSS criteria	0.11		0.37		0.67	

Abbreviations: NIHSS, National Institute of Health Stroke Scale; tPA, tissue plasminogen activator.

Table 2. Percentage Agreement of Significant Improvement Between NIHSS and Physician Documentation by Change in Score Cutoffs.

Change in Score Cutoff	Post-tPA	Discharge	Follow-Up
4	54%	73%	81%
3	66%	83%	95%
2	78%	83%	100%
1	83%	83%	100%

Abbreviations: NIHSS, National Institute of Health Stroke Scale; tPA, tissue plasminogen activator.

time points (Table 1). When 1- to 3-point changes in NIHSS were used to define significant clinical improvement post-tPA rather than ≥ 4 , the percentage agreement (number agreed upon/number of total cases) of significant improvement between NIHSS criteria and physician documentation improved from ≥ 4 : 54%, to ≥ 3 : 66%, ≥ 2 : 78%, and ≥ 1 : 83%, respectively. Percentage agreements for each cutoff at each time point (post-tPA, discharge, and follow-up) are displayed in Table 2.

Discussion

Given that tPA restores blood flow to the ischemic penumbra through recanalization, one would expect potentially to see high rates of significant improvement in clinical examination 24 hours posttreatment. We therefore chose this cohort as our population in which to evaluate the utility of the NIHSS to follow change in neurologic function. Our results indicate that as expected, relying solely on the NIHSS can underestimate patient improvement and may not be the ideal choice for monitoring, despite its ease of administration. Chart review identified 18 patients who benefitted from IV tPA in the acute setting based on physician opinion and an improved comprehensive neurological examination but did not meet NIHSS criteria (change of ≥ 4 points). The insensitivity of our current outcome metrics to detect improvement may somewhat explain why until recently many of the clinical trials in stroke treatment failed to show statistically significant results.¹⁰⁻¹³ Although the NIHSS has its strengths, it is not necessarily adequate as a measure of daily progress or an outcome tool to determine whether an individual is getting better, and if so

by how much. These overall results are consistent with those seen in the NINDS trial, with an increased number of patients showing improvement by NIHSS criteria at follow-up compared to 24 hours posttreatment and on discharge. Improved sensitivity at follow-up is likely due to the fact that a small amount of brain saved enhances recovery over time, as reorganization occurs until a critical point is reached where improvement is detectable by any metric, regardless of how sensitive.

We hypothesized that the inaccurate representation of the right hemisphere and posterior circulation function may contribute to the insensitivity of NIHSS. The number of posterior circulation events was too small to analyze separately; however, when we evaluated those with right hemisphere and posterior circulation lesions compared to left hemisphere strokes, there was no difference in the sensitivity of the NIHSS to measure improvement in our population. The sensitivity of the scale was, however, dependent on admission NIHSS score. The scale was more likely to show significant improvement in those with scores of ≥ 8 . This likely reflects the fact that patients with severe strokes have a larger potential for improvement, whereas those with lower scores need to recover almost completely in order to meet criteria. This also suggests that the NIHSS may be sensitive enough to detect improvement in those patients with large vessel occlusions who typically present with higher admission NIHSS scores but not as useful in those presenting with smaller lacunar infarcts and lower scores. Unfortunately, despite low scores, individuals with “minor strokes” (NIHSS < 5) continue to experience significant problems that impact their relationships with others, return to work, and reintegration into society.¹⁴ It is important to have the ability to follow their progress as well.

Given the limitations of requiring an NIHSS change of 4 points or more to define significant functional change, we evaluated the impact of using lower thresholds to define improvement. Using values of 3, 2, and 1, we did find higher rates of percentage agreement with physician documentation. The difficulty with using lower cutoffs is the interrater variability of the NIHSS,¹⁵ which may increase the rate of false-positive changes in examination when lower thresholds are used. Therefore, determining the most ideal threshold is difficult and may vary for each patient, and although we focused here on clinical improvement, it may be most relevant when considering functional decline. An additional complicating factor is the nonlinearity of the NIHSS. A change in score from 1 to 3 due to a newly paretic limb represents significant worsening, whereas a change from 18 to 20 in a patient with encephalopathy may not. These factors highlight the relative importance of provider judgment compared to relying solely on numeric values.

Our study is not without limitations. It is based on the experiences of a single stroke center. Although much of the data were collected prospectively, some values of NIHSS posttreatment and on discharge were retrospectively extrapolated. Extrapolation was based on a complete neurological examination, a method previously shown to have good

accuracy and reproducibility.⁸ There was also a significant number of patients lost to follow-up, resulting in a smaller sample size for our final comparison at postdischarge follow-up. It is possible that patients who followed up in clinic were different than those who did not (ie, had lower NIHSS scores than those lost to follow-up); however, there is no indication that this would have biased our findings when comparing improvement or change in function via NIHSS score to physician documentation, particularly in the acute setting. Additionally, it is possible that the reviewers identified too many patients as having a significant improvement. However, improvement was evaluated independently by 2 blinded reviewers, with good interrater reliability ($\kappa = 0.63$). Finally, our stroke center is in an urban setting, with high rates of both atrial fibrillation and carotid disease. It is possible that our results are not generalizable to every population.

Regardless of the limitations, our findings illustrate the inadequacies of the NIHSS in capturing changes in functional outcome, particularly when used for daily neurologic monitoring in patients with lower NIHSS scores. This has important implications as we consider how best to standardize assessment and documentation practices for monitoring both daily progress and overall short- and long-term outcome. A more sensitive tool is needed that maintains the efficiency and reproducibility of the NIHSS. Testing of one such battery is currently underway at our institution.

Conclusion

The NIHSS is a rapid, reproducible, assessment of severity in the acute stroke setting; but, as a metric for improvement posttreatment, it may inadequately capture functional changes, particularly for daily monitoring in those with lower NIHSS scores.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

References

1. Gottesman RF, Kleinman JT, Davis C, Heidler-Gary J, Newhart M, Hillis AE. The NIHSS-plus: improving cognitive assessment with the NIHSS. *Behav Neurol*. 2010;22(1-2):11-15.
2. Linfante I, Llinas RH, Schlaug G, Chaves C., Warach S, Caplan LR. Diffusion-weighted imaging and National Institutes of Health Stroke Scale in the acute phase of posterior-circulation stroke. *Arch Neurol*. 2001;58(4):621-628.
3. Planton M, Peiffer S, Albucher J, et al. Neuropsychological outcome after a first symptomatic ischaemic stroke with ‘good recovery’. *Eur J Neurol*. 2012;19(2):212-219.
4. Van Zandvoort MJ, Kappelle LJ, Algra A, De Haan EH. Decreased capacity for mental effort after single supratentorial

- lacunar infarct may affect performance in everyday life. *J Neurol Neurosurg Psychiatry*. 1998;65(5):697-702.
5. ECASS, Larrue V, von Kummer R, Muller A, Bluhmki E. Risk factors for severe hemorrhagic transformation in ischemic stroke patients treated with recombinant tissue plasminogen activator. a secondary analysis of the European-Australian Acute Stroke Study (ECASS II). *Stroke*. 2001;32(2):438-441.
 6. The National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group. Tissue plasminogen activator for acute ischemic stroke. *N Engl J Med*. 1995;333(24):1581-1588.
 7. TOAST Adams HP, Bendixen BH, Kappelle LJ, et al. Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. *Stroke*. 1993;24:35-41.
 8. Kasner SE, Chalela JA, Luciano JM, et al. Reliability and validity of estimating the NIH stroke scale score from medical records. *Stroke*. 1999;30(8):1534-1537.
 9. Reeves MJ, Arora S, Broderick JP, et al. Acute stroke care in the US: results from 4 pilot prototypes of the Paul Coverdell National Acute Stroke Registry. *Stroke*. 2005;36(6):1232-1240.
 10. Liebeskind DS. Time and tissue: imaging of early ischemia in FAST-MAG; Presented at: The International Stroke Conference Annual Meeting; 2014; San Diego, CA.
 11. Davis SM, Donnan GA, Parsons MW, et al. Effects of alteplase beyond 3 hours after stroke in the Echoplanar Imaging Thrombolytic Evaluation Trial (EPITHET): a placebo-controlled randomized trial. *Lancet Neurol*. 2008;7(4):299-309.
 12. Broderick JP, Palesch YY, Demchuk AM, et al. Endovascular therapy after intravenous t-PA versus t-PA alone for stroke. *N Engl J Med*. 2013;368(10):893-903.
 13. Kidwell CS, Jahan R, Gornbein J, et al. A trial of imaging selection and endovascular treatment for ischemic stroke. *N Engl J Med*. 2013;368(10):914-923.
 14. Maygers J, Lawrence E, Llinas RH, Marsh EB. Transitions of care: increasing follow-up and decreasing readmission rates after hospitalization for acute ischemic stroke. 2015; Poster presented at: The International Stroke Conference Annual Meeting; 2015; Nashville, TN.
 15. Meyer BC, Lyden PD. The modified national institutes of health stroke scale: its time has come. *J Stroke*. 2009;4(4):267-273.