Interrater Reliability of the Pediatric National Institutes of Health Stroke Scale (PedNIHSS) in a Multicenter Study

Rebecca N. Ichord, MD; Rachel Bastian, BA; Lisa Abraham, MD; Rand Askalan, MD, PhD; Susan Benedict, MD; Timothy J. Bernard, MD; Lauren Beslow, MD; Gabrielle deVeber, MD; Michael Dowling, MD, PhD, MSCS; Neil Friedman, MBChB; Heather Fullerton, MD, MAS; Lori Jordan, MD, PhD; Li Kan, MD; Adam Kirton, MD; Catherine Amlie-Lefond, MD; Daniel Licht, MD; Warren Lo, MD; Chalmer McClure, MD, PhD; Steve Pavlakis, MD; Sabrina E. Smith, MD, PhD; Marilyn Tan, MD; Scott Kasner, MD, MSCE; Abbas F. Jawad, PhD

Background and Purpose—Stroke is an important cause of death and disability among children. Clinical trials for childhood stroke require a valid and reliable acute clinical stroke scale. We evaluated interrater reliability (IRR) of a pediatric adaptation of the National Institutes of Health Stroke Scale.

Methods—The pediatric adaptation of the National Institutes of Health Stroke Scale was developed by pediatric and adult stroke experts by modifying each item of the adult National Institutes of Health Stroke Scale for children, retaining all examination items and scoring ranges of the National Institutes of Health Stroke Scale. Children 2 to 18 years of age with acute arterial ischemic stroke were enrolled in a prospective cohort study from 15 North American sites from January 2007 to October 2009. Examiners were child neurologists certified in the adult National Institutes of Health Stroke Scale. Each subject was examined daily for 7 days or until discharge. A subset of patients at 3 sites was scored simultaneously and independently by 2 study neurologists.

Results—IRR testing was performed in 25 of 113 a median of 3 days (interquartile range, 2 to 4 days) after symptom onset. Patient demographics, total initial pediatric adaptation of the National Institutes of Health Stroke Scale scores, risk factors, and infant characteristics in the IRR subset were similar to the non-IRR subset. The 2 raters’ total scores were identical in 60% and within 1 point in 84%. IRR was excellent as measured by concordance correlation coefficient of 0.97 (95% CI, 0.94 to 0.99); intraclass correlation coefficient of 0.99 (95% CI, 0.97 to 0.99); precision measured by Pearson r of 0.97; and accuracy measured by the bias correction factor of 1.0.

Conclusions—There was excellent IRR of the pediatric adaptation of the National Institutes of Health Stroke Scale in a multicenter prospective cohort performed by trained child neurologists. (Stroke. 2011;42:613-617.)

Key Words: childhood ischemic stroke outcome stroke scale validation

Ischemic stroke affects 1.2 to 7.9 per 100 000 children aged 1 month to 18 years annually in Europe and North America and ranks among the top 10 causes of death. Long-term motor and cognitive deficits interfering with activities of daily life and academic attainment affect 40% to 60% of survivors. There are no proven strategies for acute management or prevention of childhood stroke other than blood transfusion for children with sickle cell anemia. Progress in defining factors that determine outcome and designing clinical trials are hindered by the lack of a validated and reliable clinical stroke scale.

Received November 28, 2010; final revision received December 22, 2010; accepted December 23, 2010.

From the Department of Neurology (R.N.I., R.B., L.B., D.L., S.E.S.), Children’s Hospital Philadelphia, University of Pennsylvania, Philadelphia, PA; the Department of Pediatrics (L.A.), Schenectady Neurological Consultants PC, Schenectady, NY; Department of Pediatrics (R.A., G.d.V.), Division of Neurology, Hospital for Sick Children, Toronto, Ontario Canada; the Department of Pediatrics (S.B.), Division of Neurology, Primary Children’s Hospital, University of Utah, Salt Lake City, UT; the Section of Child Neurology (T.J.B.), Denver Children’s Hospital, University of Colorado, Aurora, CO; the Department of Neurology (M.D.), Children’s Medical Center Dallas, University of Texas Southwestern, Dallas, TX; Center Pediatric Neurology/Desh S60 (N.F.), Neurological Institute, Cleveland Clinic, Cleveland, OH; the Department of Neurology (H.F.), University of California at San Francisco, San Francisco, CA; the Department of Neurology (L.J.), Johns Hopkins University, Baltimore, MD; the Division of Pediatric Neurology (L.K.), Schneider Children’s Hospital, New Hyde Park, NY; the Department of Pediatric Clinical Neurosciences (A.K.), Alberta Children’s Hospital, University of Calgary, Calgary, Alberta, Canada; the Department of Neurology (C.A.-L.), Children’s Hospital of Wisconsin, Medical College of Wisconsin, Milwaukee, WI; the Department of Neurology (W.L.), Nationwide Children’s Hospital, Ohio State University, Columbus, OH; the Department of Neurology (C.M.), Loma Linda University, Loma Linda, CA; the Department of Developmental Medicine and Child Neurology (S.P.), Maimonides Infants Children’s Hospital, Brooklyn NY; the Department of Neurosciences (M.T.), University of the Philippines, Manila, Philippines; the Department of Neurology (S.K.), University of Pennsylvania, Philadelphia, PA; and the Department of Pediatrics (A.F.J.), Children’s Hospital of Philadelphia, University of Pennsylvania, Philadelphia, PA.

The online-only Data Supplement is available at http://stroke.ahajournals.org/cgi/content/full/STROKEAHA.110.607192/DC1.

Correspondence to Rebecca N. Ichord, MD, Department of Neurology, Children’s Hospital of Philadelphia, Colket Translational Research Building 10th Floor, 3501 Civic Center Blvd, Philadelphia, PA 19104. E-mail ichord@email.chop.edu

© 2011 American Heart Association, Inc.

Stroke is available at http://stroke.ahajournals.org DOI: 10.1161/STROKEAHA.110.607192

613
Previously published cohort studies of acute clinical presentation or long-term outcome in childhood stroke have not used standardized, validated, and reliable measures of clinical stroke severity at stroke onset. The National Institutes of Health Stroke Scale (NIHSS) is a quantitative measure of stroke-related acute neurological deficit, which has proven intra- and interrater reliability (IRR) and predictive validity for outcome among adults. Neurological examination of children requires adjustment according to the maturation of the child’s neurological and cognitive function and ability to comprehend instructions. The objective of this study was to evaluate IRR of a pediatric modification of the NIHSS, the Pediatric NIHSS (PedNIHSS), administered by child neurologists in children with acute arterial ischemic stroke.

Methods

Study Design
This was a multicenter, prospective consecutive cohort study designed to evaluate the PedNIHSS for IRR, the relationship to acute infarct volume, and predictive validity for functional outcome at 3 and 12 months after stroke onset. Outcome prediction and relationship to infarct volume will be reported separately. Patients were enrolled from 15 sites in the United States and Canada from January 2007 through October 2009. This study was approved by the Institutional Review Board or Institutional Ethics Board at all sites.

Subjects
Patients were identified as potential study subjects and enrolled by the site neurologist or his or her designee at each site. Informed consent was obtained from the parent or guardian. Children enrolled met the following criteria: age at stroke onset ≥2 years and <19 years, presentation within 96 hours after symptom onset with acute neurological deficit of any duration consistent with focal brain ischemia in an arterial distribution, and MRI or CT performed within 96 hours of symptom onset demonstrating acute infarction in an arterial territory corresponding to the clinical deficit. Exclusion criteria included: acute traumatic brain injury; primary intracerebral or intraventricular hemorrhage; meningitis or encephalitis; status epilepticus (continuous clinical or electrographic seizure activity for ≥30 minutes); severe metabolic, toxic, or global hypoxic–ischemic encephalopathy; pre-existing severe neurological deficit due to malignancy, congenital brain malformation, neurodegenerative disease, metabolic encephalopathy, severe residual deficits from perinatal, or postnatal acquired encephalopathy; or stroke after craniotomy for any neurological procedure.

Development of the PedNIHSS
The PedNIHSS was developed by a panel of pediatric and adult stroke experts by a consensus review process in which each item of the NIHSS was reviewed and modified for age-dependent variations in comprehension and participation in the examination item and age appropriateness of testing materials (language items, picture, commands). All items were adapted to an age-appropriate format, whereas the scoring strategy and scoring ranges for all items administered in the adult NIHSS were retained in the PedNIHSS. See the online supplement (available at http://stroke.ahajournals.org) for details of modifications for each item, and item-by-item guide for administration in children aged 2 to 18 years used in this study. Pilot testing of the pediatric modifications was conducted by 4 study neurologists in 15 patients at 2 study sites (Children’s Hospital of Philadelphia and Toronto Hospital for Sick Children) and supported the validity and reliability of the PedNIHSS in a larger multicenter study.

Data Collection
We collected data on demographics, medical history, clinical presentation, treatment, stroke risk factors, stroke-related diagnostic studies, and hospital course. All data were entered in a central database at the clinical and data coordinating center at Children’s Hospital of Philadelphia. Neuroimaging studies including head CT, brain MRI, brain and cervical MR angiography, CT angiograms, and catheter angiograms were deidentified and sent to a central imaging repository at the clinical and data coordinating center. Admission case report forms were reviewed by the clinical and data coordinating center staff (Principal Investigator and study coordinator) for completeness and adherence to inclusion/exclusion criteria. Admission neuroimaging was reviewed by the study Principal Investigator (R.I.) to confirm the diagnosis of arterial ischemic stroke. All cases in which study inclusion was questioned were adjudicated by a rotating panel of 3 site neurologists and a study neuroradiologist. Acute diagnostic and treatment decisions were made according to standard clinical care protocols at each site.

PedNIHSS Scoring and Reliability Testing
Study examiners included the site neurologists, who are board-certified child neurologists, and child neurology trainees supervised by the site neurologist, all experienced in the care of children with acute ischemic stroke. Each examiner was certified in use of the NIHSS using the standard online training program of the American Stroke Association (www.asatrainingcampus.org). Each site neurologist was given additional training by the study Principal Investigator and provided with a detailed set of instructions regarding modifications of the NIHSS for the PedNIHSS. The site neurologist performed the PedNIHSS once daily from the day of admission through Day 7 after admission or discharge, whichever occurred first. Evaluation of IRR was based on simultaneous scoring by 2 examiners for a subset of patients at each of 3 sites (Children’s Hospital of Philadelphia, Children’s Hospital of Pittsburgh, and Toronto Hospital for Sick Children). The hospital day chosen for the IRR examination was determined by the site neurologist as the earliest time after hospital admission that 2 examiners were available to perform simultaneous examination. The primary examiner was the site neurologist, and the secondary examiner was designated by the site neurologist from among child neurologists or child neurology trainees at that site who regularly participated in the care of patients with stroke. Each secondary examiner was trained using the online NIHSS training and certification and underwent comparable orientation and instruction in the use of the PedNIHSS as the primary site neurologist. The primary and secondary examiners each independently and simultaneously assigned scores to each item of the PedNIHSS at the time the primary examiner conducted any of that patient’s study examinations during the first week.

Statistical Analysis
The sample size needed for the assessment of IRR was based on obtaining an intraclass correlation coefficient of at least 0.80 for the PedNIHSS total scores. A sample size of 25 subjects, each rated independently by 2 neurologists using the PedNIHSS, was estimated to have 85% power to reject the null hypothesis that intraclass correlation coefficient = 0.50 using an F-test with a 2-sided α of 0.05. PASS was used for sample size estimation (NCSS, Kaysville, UT). Statistical analyses were performed using SAS, SPSS, and MedCalc statistical packages (SAS Institute, Cary, NC; SPSS Science, Chicago, IL; MedCalc Software, Mariakerke, Belgium). Descriptive analyses were performed using means with SDs or medians with interquartile ranges for continuous variables and frequency distributions and proportions for categorical variables. All continuous measures were evaluated for normality, and those variables differing markedly from normality were considered candidates for Box-Cox transformations. For the PedNIHSS scoring data, the mean, median, and frequency distribution for responses were computed for each individual item. A summed total score was computed and its distribution examined for this sample. A 2-tailed probability value <0.05 was considered statistically significant. For the analysis of
PedNIHSS reliability, internal consistency (Cronbach α) of the total scale was computed. IRR was assessed through the intraclass correlation coefficient calculated by the analysis of variance between total summed scores. We analyzed agreement on each item of the PedNIHSS for the 2 raters using weighted κ statistics. A κ < 0.40 defined poor reliability, 0.40 < κ < 0.75 defined moderate reliability, and κ ≥ 0.75 defined excellent reliability. The measures of precision ρ and accuracy bias correction factor for the total score pairs obtained from both raters were estimated. Scoring bias and the limits of agreement of the 2 raters’ total PedNIHSS score were estimated using Bland and Altman methods. In this method, score differences between the 2 raters were plotted against the average reading of the 2 raters. The mean differences in the total scores represent the bias in reading among the 2 rates. The 95% limits of agreement between the 2 raters were estimated by difference ± SD of the differences.

**Results**

**Subject Characteristics**

One hundred thirteen patients were enrolled, of which 25 (22%) underwent simultaneous examinations by 2 investigators and were included in the IRR analysis. Simultaneous examinations for IRR were performed at a median interval of 3 days after symptom onset (interquartile range, 2 to 4 days) and were completed in 4 patients (16%) on hospital Day 1, 6 (24%) on Day 2, 6 (24%) on Day 3, 3 (12%) on Day 4, 2 (8%) on Day 5, and 4 (16%) on Day 6. Patients in the IRR subgroup were compared with non-IRR patients with respect to age, sex, time from symptom onset to presentation to healthcare providers, median total PedNIHSS score on initial examination, infarct location, and primary stroke risk factor and were found to be similar (Table 1).

**Interrater Reliability**

There was an excellent association between scores from both raters, as seen in Figures 1 and 2, with a Spearman correlation coefficient of 0.97 (P < 0.001). The distribution of differences between the 2 raters’ total scores is shown in Table 2. Scores were identical in 60% of cases and within 1 point in 84%. Bias in scoring estimated using Bland and Altman methods was very small at 0.1 (Figure 3), and disagreement between raters was random. The concordance correlation coefficient was 0.97 (95% CI, 0.94 to 0.99). The precision measured by Pearson ρ was 0.97, and accuracy measured by the bias correction factor was 1.0. Reliability measured by intraclass correlation coefficient was excellent (0.99; 95% CI, 0.97 to 0.99). Internal consistency measured by Cronbach α estimated to be 0.99. Analysis of IRR for each examination item of the PedNIHSS demonstrated that agreement was excellent or moderate for all items (Table 3).

**Discussion**

We found excellent IRR of the PedNIHSS administered by child neurologists in a prospective, multicenter study of children aged 2 to 18 years with acute arterial ischemic stroke. Moreover, our finding that IRR was good to excellent for all items of the scale suggests that all items contribute to the overall excellent IRR of total PedNIHSS scores. Subjects examined for IRR were representative of the larger cohort with respect to demographic and clinical factors as well as initial PedNIHSS score. The generalizability of our findings is supported by the fact that subjects in our study were representative of the age and sex distribution, time to initial presentation, and primary stroke risk factors reported in other pediatric stroke cohort studies.

The PedNIHSS closely resembles the adult NIHSS and is the first clinical stroke scale developed and evaluated for

---

**Table 1. Demographic and Clinical Characteristics of Study Subjects**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Non-IRR Cohort (N=88)</th>
<th>IRR Cohort (N=25)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years (mean ±SD)</td>
<td>10.4 ± 5.4</td>
<td>9.7 ± 5.2</td>
<td>0.55</td>
</tr>
<tr>
<td>Sex, no. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>56 (64)</td>
<td>15 (60)</td>
<td>0.74</td>
</tr>
<tr>
<td>Female</td>
<td>32 (36)</td>
<td>10 (40)</td>
<td></td>
</tr>
<tr>
<td>Time to presentation to first health care provider, median hours (IQR)</td>
<td>2.2 (0.9–13)</td>
<td>3.2 (1.0–14)</td>
<td>0.64</td>
</tr>
<tr>
<td>Total PedNIHSS score on initial examination, median (IQR)</td>
<td>7 (3–14)</td>
<td>5 (3–10)</td>
<td>0.25</td>
</tr>
<tr>
<td>Location of stroke: no. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anterior circulation (MCA, ACA)</td>
<td>59 (67)</td>
<td>17 (68)</td>
<td>0.93</td>
</tr>
<tr>
<td>Posterior cerebral artery only</td>
<td>6 (7)</td>
<td>2 (8)</td>
<td>0.84</td>
</tr>
<tr>
<td>Brainstem or cerebellum</td>
<td>15 (17)</td>
<td>5 (20)</td>
<td>0.73</td>
</tr>
<tr>
<td>Multiple vascular territories</td>
<td>8 (9)</td>
<td>1 (4)</td>
<td>0.41</td>
</tr>
<tr>
<td>Primary stroke risk factor: no. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardiac</td>
<td>25 (28)</td>
<td>9 (36)</td>
<td>0.47</td>
</tr>
<tr>
<td>Arteriopathy*</td>
<td>28 (32)</td>
<td>7 (28)</td>
<td>0.72</td>
</tr>
<tr>
<td>Hemoglobinopathy</td>
<td>7 (8)</td>
<td>1 (4)</td>
<td>0.50</td>
</tr>
<tr>
<td>Thrombophilia†</td>
<td>4 (5)</td>
<td>2 (8)</td>
<td>0.50</td>
</tr>
<tr>
<td>Other‡</td>
<td>6 (7)</td>
<td>1 (4)</td>
<td>0.61</td>
</tr>
<tr>
<td>Unknown</td>
<td>18 (20)</td>
<td>5 (20)</td>
<td>0.96</td>
</tr>
</tbody>
</table>

*Arteriopathy includes: non-Hbg SS moyamoya syndromes, intracranial vasculopathy, dissection, vasculitis, postradiation vasculopathy.
†Cases in which thrombophilia is an isolated risk factor.
‡Includes history of minor trauma without dissection, carotid body tumor resection, critical illness after liver transplant, migraine, basilar artery dissecting aneurysm.
IQR indicates interquartile range; MCA, middle cerebral artery; ACA, anterior cerebral artery.

---

**Figure 1.** Scatterplot demonstrating association between total PedNIHSS scores of the primary rater (Y-axis) and secondary rater (X-axis). Open symbols depict data for 2 subjects; closed symbols for single subjects. Scatterplot of total PedNIHSS scores by 2 examiners.
reliability in children. Like with the adult NIHSS, we found the PedNIHSS is readily performed in the acute hospital setting in children of a broad range of ages and stroke severity. The excellent IRR demonstrated in our study compares favorably to that seen in adults examined with the NIHSS. Brott et al demonstrated good IRR in a cohort of 24 adult patients examined by neurologists and neurology house officers with $\kappa$ values of 0.70 to 0.77 for the total score. Interrater agreement for individual examination items in Brott’s study was similar to our findings with the strongest interrater agreement seen in adults for level of consciousness as assessed by level of consciousness questions, best gaze, and motor tasks and least agreement for sensory and language examination items. Goldstein et al further evaluated IRR of a refined version of Brott’s stroke scale in 20 adult patients with acute ischemic stroke and found good agreement with $\kappa$ values for language and motor items of 0.77 to 0.79 and lower $\kappa$ values (0.4 to 0.6) for level of consciousness commands, neglect, and sensory items. In contrast to the adult studies, we found good IRR for facial weakness, dysarthria, and ataxia. The reasons for this difference are uncertain. It is possible that the more observational nature of the examination in children actually enhances reliability compared with the more prescribed and confrontational examination in adults.

There are several limitations in our study. The small sample size relative to age range of subjects precluded separate analysis of IRR in different age groups. We did not evaluate PedNIHSS in children aged <2 years due to the importance we ascribed to the inclusion of language assessment, because the typical child aged <2 years has limited language ability. Additionally, children aged <2 years and neonates often fail to present with focal deficits, potentially requiring a scale with less emphasis on focal sensorimotor deficits. It is possible interactions between raters during examination may have occurred to achieve clarity of findings, resulting in an overestimate of interrater reliability. This effect is likely small because investigator training addressed this possible confounder at study start-up. Examiners in our study were child neurologists and child neurology trainees. This limits the generalizability of our findings on IRR with respect to the examining healthcare provider. Because neurological assessment of the acutely ill child is often challenging, evaluation of the reliability of the PedNIHSS performed by nonneurologists will be needed before this scale can be used by nonneurologists in clinical trials.

Table 2. Distribution of Differences in Total PedNIHSS Scores Between the 2 Examiners

<table>
<thead>
<tr>
<th>Difference (Primary–Secondary Raters' Scores)</th>
<th>Frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>−4</td>
<td>1 (4)</td>
</tr>
<tr>
<td>−2</td>
<td>2 (8)</td>
</tr>
<tr>
<td>−1</td>
<td>2 (8)</td>
</tr>
<tr>
<td>0</td>
<td>15 (60)</td>
</tr>
<tr>
<td>1</td>
<td>4 (16)</td>
</tr>
<tr>
<td>3</td>
<td>1 (4)</td>
</tr>
<tr>
<td>Total</td>
<td>25</td>
</tr>
</tbody>
</table>

*Raters had 100% agreement; no. of categories insufficient to calculate $\kappa$. LOC indicates level of consciousness.
Conclusions
Evaluation of the PedNIHSS in a multicenter cohort study revealed excellent IRR. This is an important first step in developing a valid pediatric acute stroke scale and is of fundamental importance for planning and executing future clinical trials in childhood stroke. Analysis of the relationship of PedNIHSS scores with infarct volume and 3- and 12-month functional outcomes obtained in our study is in progress and will further characterize the validity and use of this instrument.

Acknowledgments
We thank Stefanie Mason and Charlene Jones for database development and data management; and Jorina Elbers, MD, Lori Billinghurst, MD, and Nomazulu Dlamini, MBBS, for performing IRR examinations.

Sources of Funding
R.N.I. received funding from National Institutes of Health (NIH) R01-NS050488 and K23-NS062110; T.J.B. NIH K23-HL096895; L.B. NIH T32-NS007413 and L. Morton Morley Funds of The Philadelphia Foundation; G.d.V. NIH R01-NS062820; M.D. NIH K23-RR024983 and The First American Real Estate Services, Inc; H.F. NIH R01-NS062820 and K02-NS053883; L.J. NIH K23-NS062110; A.K. Heart/Stroke Foundation Canada, Heart/Stroke Foundation Alberta and Children’s Hospital Foundation and Research Institute, and Hotchkiss Brain Institute; D.L. NIH K23-NS52380, Dana Foundation, and June and Steve Wolfson Family Fund for Neurological Research; and S.E.S. NIH K12-NS049453.

Disclosures
R.N.I. and L.J. are consultants and on the Advisory Board of the Berlin Heart Clinical Event Committee.

References
Interrater Reliability of the Pediatric National Institutes of Health Stroke Scale (PedNIHSS) in a Multicenter Study

Stroke. published online February 11, 2011;
Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2011 American Heart Association, Inc. All rights reserved.
Print ISSN: 0039-2499. Online ISSN: 1524-4628

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://stroke.ahajournals.org/content/early/2011/02/11/STROKEAHA.110.607192

Data Supplement (unedited) at:
http://stroke.ahajournals.org/content/suppl/2011/02/17/STROKEAHA.110.607192.DC1

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Stroke can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Stroke is online at:
http://stroke.ahajournals.org/subscriptions/
ONLINE SUPPLEMENT

PedNIHSS Guide and item-by-item description of exam administration.
**PedNIHSS INSTRUCTIONS:** Administer stroke scale items in the order listed. Follow directions provided for each exam item. Scores should reflect what the patient does, not what the clinician thinks the patient can do. **MODIFICATIONS FOR CHILDREN:** Modifications to testing instructions from the adult version for use in children are shown in bold italic with each item where appropriate. Items with no modifications should be administered and scored with children in the same manner as for adults.

<table>
<thead>
<tr>
<th>Item# and Instructions</th>
<th>Scale Definition and Scoring Guide</th>
</tr>
</thead>
</table>
| **1a. Level of Consciousness:** the investigator must choose a response, even if a full evaluation is prevented by such obstacles as an endotracheal tube, language barrier, oro-tracheal trauma/bandages. A 3 is scored only if the patient makes no movement (other than reflexive posturing) in response to noxious stimulation. | 0 = Alert; keenly responsive.  
1 = Not alert, but arousable by minor stimulation to obey, answer, or respond.  
2 = Not alert, requires repeated stimulation to attend, or is obtunded and requires strong or painful stimulation to make movements (not stereotyped).  
3 = Responds only with reflex motor or autonomic effects or totally unresponsive, flaccid, areflexic. |
| **1b. LOC Questions:** The patient is asked the month and his/her age. The answer must be correct - there is no partial credit for being close. Aphasic and stuporous patients who do not comprehend the questions will score 2. Patients unable to speak because of endotracheal intubation, oro-tracheal trauma, severe dysarthria from any cause, language barrier or any other problem not secondary to aphasia are given a 1. It is important that only the initial answer be graded and that the examiner not "help" the patient with verbal or non-verbal cues. **Modified for children, age 2 years and up. A familiar Family Member must be present for this item:** Ask the child "how old are you?" or "How many years old are you?" for question number one. Give credit if the child states the correct age, or shows the correct number of fingers for his/her age. **For the second question, ask the child "where is XX?", XX referring to the name of the parent or other familiar family member present. Use the name for that person which the child typically uses, e.g. "mommy". Give credit if the child correctly points to or gazes purposefully in the direction of the family member. | 0 = Answers both questions correctly.  
1 = Answers one question correctly.  
2 = Answers neither question correctly. |
| **1c. LOC Commands:** The patient is asked to open and close the eyes and then to grip and release the non-paretic hand. **For children one may substitute the command to grip the hand with the command "show me your nose" or "touch your nose".** Substitute another one step command if the hands cannot be used. Credit is given if an unequivocal attempt is made but not completed due to weakness. If the patient does not respond to command, the task should be demonstrated to them (pantomime) and score the result (i.e., follows none, one or two commands). Patients with trauma, amputation, or other physical impediments should be given suitable one-step commands. Only the first attempt is scored. | 0 = Performs both tasks correctly  
1 = Performs one task correctly  
2 = Performs neither task correctly |
| **2. Best Gaze:** Only horizontal eye movements will be tested. Voluntary or reflexive (oculocephalic) eye movements will be scored but caloric testing is not done. If the patient has a conjugate deviation of the eyes that can be overcome by voluntary or reflexive activity, the score will be 1. If a patient has an isolated peripheral nerve paresis (CN III, IV or VI) score a 1. Gaze is testable in all aphasic patients. Patients with ocular trauma, bandages, pre-existing blindness or other disorder of visual acuity or fields should be tested with reflexive movements and a choice made by the investigator. Establishing eye contact and then moving about the patient from side to side will occasionally clarify the presence of a partial gaze palsy. | 0 = Normal  
1 = Partial gaze palsy. This score is given when gaze is abnormal in one or both eyes, but where forced deviation or total gaze paresis are not present.  
2 = Forced deviation, or total gaze paresis not overcome by the oculocephalic maneuver. |
| **3. Visual:** Visual fields (upper and lower quadrants) are tested by confrontation, using finger counting (for children > 6 years) or visual threat (for children age 2 to 6 years) as appropriate. Patient must be encouraged, but if they look at the side of the moving fingers appropriately, this can be scored as normal. If there is unilateral blindness or enucleation, visual fields in the remaining eye are scored. Score 1 only if a clear-cut asymmetry, including quadrantanopia is found. If patient is blind from any cause score 3. Double simultaneous stimulation is performed at this point. If there is extinction patient receives a 1 and the results are used to answer question 11. | 0 = No visual loss  
1 = Partial hemianopia  
2 = Complete hemianopia  
3 = Bilateral hemianopia (blind including cortical blindness) |
### 4. Facial Palsy

Ask, or use pantomime to encourage the patient to show teeth or raise eyebrows and close eyes. Score symmetry of grimace in response to noxious stimuli in the poorly responsive or non-comprehending patient. If facial trauma/bandages, orotracheal tube, tape or other physical barrier obscures the face, these should be removed to the extent possible.

<table>
<thead>
<tr>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 = Normal symmetrical movement</td>
</tr>
<tr>
<td>1 = Minor paralysis (flattened nasolabial fold, asymmetry on smiling)</td>
</tr>
<tr>
<td>2 = Partial paralysis (total or near total paralysis of lower face)</td>
</tr>
<tr>
<td>3 = Complete paralysis of one or both sides (absence of facial movement in the upper and lower face)</td>
</tr>
</tbody>
</table>

### 5 & 6. Motor Arm and Leg

The limb is placed in the appropriate position: extend the arms (palms down) 90 degrees (if sitting) or 45 degrees (if supine) and the leg 30 degrees (always tested supine). Drift is scored if the arm falls before 10 seconds or the leg before 5 seconds. **For children too immature to follow precise directions or uncooperative for any reason, power in each limb should be graded by observation of spontaneous or elicited movement according to the same grading scheme, excluding the time limits.** The aphasic patient is encouraged using urgency in the voice and pantomime but not noxious stimulation. Each limb is tested in turn, beginning with the non-paretic arm. Only in the case of amputation or joint fusion at the shoulder or hip, or immobilization by an IV board, may the score be "9" and the examiner must clearly write the explanation for scoring as a "9". Score each limb separately.

#### 5a. Left Arm
- 0 = No drift, limb holds 90 (or 45) degrees for full 10 seconds.
- 1 = Drift, limb holds 90 (or 45) degrees, but drifts down before full 10 seconds; does not hit bed or other support.
- 2 = Some effort against gravity, limb cannot get to or maintain (if cued) 90 (or 45) degrees, drifts down to bed, but has some effort against gravity.
- 3 = No effort against gravity, limb falls.
- 4 = No movement
- 9 = Amputation, joint fusion explain:

#### 5b. Right Arm
- 0 = No drift, limb holds 90 (or 45) degrees for full 10 seconds.
- 1 = Drift, limb holds 90 (or 45) degrees, but drifts down before full 10 seconds; does not hit bed or other support.
- 2 = Some effort against gravity, limb cannot get to or maintain (if cued) 90 (or 45) degrees, drifts down to bed, but has some effort against gravity.
- 3 = No effort against gravity, limb falls.
- 4 = No movement
- 9 = Amputation, joint fusion explain:

#### 6a. Left Leg
- 0 = No drift, leg holds 30 degrees position for full 5 seconds.
- 1 = Drift, leg falls by the end of the 5 second period but does not hit bed.
- 2 = Some effort against gravity; leg falls to bed by 5 seconds, but has some effort against gravity.
- 3 = No effort against gravity, leg falls to bed immediately.
- 4 = No movement
- 9 = Amputation, joint fusion explain:

#### 6b. Right Leg
- 0 = No drift, leg holds 30 degrees position for full 5 seconds.
- 1 = Drift, leg falls by the end of the 5 second period but does not hit bed.
- 2 = Some effort against gravity; leg falls to bed by 5 seconds, but has some effort against gravity.
- 3 = No effort against gravity, leg falls to bed immediately.
- 4 = No movement
- 9 = Amputation, joint fusion explain:

### 7. Limb Ataxia

This item is aimed at finding evidence of a unilateral cerebellar lesion. Test with eyes open. In case of visual defect, insure testing is done in intact visual field. The finger-nose-finger and heel-shin tests are performed on both sides, and ataxia is scored only if present out of proportion to weakness. **In children, substitute this task with reaching for a toy for the upper extremity, and kicking a toy or the examiner's hand, in children too young (< 5 years) or otherwise uncooperative for the standard exam item.** Ataxia is absent in the patient who cannot understand or is paralyzed. Only in the case of amputation or joint fusion may the item be scored "9", and the examiner must clearly write the explanation for not scoring. In case of blindness test by touching nose from extended arm position.

<table>
<thead>
<tr>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 = Absent</td>
</tr>
<tr>
<td>1 = Present in one limb</td>
</tr>
<tr>
<td>2 = Present in two limbs</td>
</tr>
</tbody>
</table>

### 8. Sensory

Sensation or grimace to pin prick when tested, or withdrawal from noxious stimulus in the obtunded or aphasic patient. **For children too young or otherwise uncooperative for reporting gradations of sensory loss, observe for any behavioral response to pin prick, and score it according to the same scoring scheme as a "normal" response, "mildly diminished" or "severely diminished" response.** Only sensory loss attributed to stroke is scored as abnormal and the examiner should test as many body areas [arms (not hands), legs, trunk, face] as needed to accurately check for hemisensory loss. A score of 2, "severe or total," should only be given when a severe or total loss of sensation can be clearly demonstrated. Stuporous and aphasic patients will therefore probably score 1 or 0.

<table>
<thead>
<tr>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 = Normal; no sensory loss.</td>
</tr>
<tr>
<td>1 = Mild to moderate sensory loss; patient feels pinprick is less sharp or is dull on the affected side; or there is a loss of superficial pain with pinprick but patient is aware he/she is being touched.</td>
</tr>
<tr>
<td>2 = Severe to total sensory loss; patient is not aware of being touched in the face, arm, and leg.</td>
</tr>
</tbody>
</table>
9. **Best Language:** A great deal of information about comprehension will be obtained during the preceding sections of the examination. *For children age 6 years and up with normal language development before onset of stroke: The patient is asked to describe what is happening in the attached picture, to name the items on the attached naming sheet, to repeat words from the attached list, and to read from the attached list of sentences (Table S1; Fig S1, S2, S3). Comprehension is judged from responses here as well as to all of the commands in the preceding general neurological exam. If visual loss interferes with the tests, ask the patient to identify objects placed in the hand, repeat, and produce speech. The intubated patient should be asked to write. The patient in coma (question 1a=3) will arbitrarily score 3 on this item. The examiner must choose a score in the patient with stupor or limited cooperation but a score of 3 should be used only if the patient is mute and follows no one step commands. For children age 2 yrs to 6 yrs (or older children with premorbid language skills < 6 yr level), score this item based on observations of language comprehension and speech during the examination.*

0 = No aphasia, normal

1 = Mild to moderate aphasia; some obvious loss of fluency or facility of comprehension, without significant limitation on ideas expressed or form of expression. Reduction of speech and/or comprehension, however, makes conversation about provided material difficult or impossible. For example in conversation about provided materials examiner can identify picture or naming card from patient's response.

2 = Severe aphasia; all communication is through fragmentary expression; great need for inference, questioning, and guessing by the listener. Range of information that can be exchanged is limited; listener carries burden of communication. Examiner cannot identify materials provided from patient response.

3 = Mute, global aphasia; no usable speech or auditory

The patient with brain stem stroke who has bilateral loss of sensation is scored 2. If the patient does not respond and is quadriplegic score 2. Patients in coma (Item 1a=3) are arbitrarily given a 2 on this item.
10. **Dysarthria**: If patient is thought to be normal an adequate sample of speech must be obtained by asking patient to read or repeat words from the attached list. If the patient has severe aphasia, the clarity of articulation of spontaneous speech can be rated. Only if the patient is intubated or has other physical barrier to producing speech, may the item be scored “9”, and the examiner must clearly write an explanation for not scoring. Do not tell the patient why he/she is being tested.

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Normal</td>
</tr>
<tr>
<td>1</td>
<td>Mild to moderate; patient slurs at least some words and, at worst, can be understood with some difficulty.</td>
</tr>
<tr>
<td>2</td>
<td>Severe; patient's speech is so slurred as to be unintelligible in the absence of or out of proportion to any dysphasia, or is mute/anarthric.</td>
</tr>
<tr>
<td>9</td>
<td>Intubated or other physical barrier, explain:</td>
</tr>
</tbody>
</table>

11. **Extinction and Inattention (formerly Neglect)**: Sufficient information to identify neglect may be obtained during the prior testing. If the patient has a severe visual loss preventing visual double simultaneous stimulation, and the cutaneous stimuli are normal, the score is normal. If the patient has aphasia but does appear to attend to both sides, the score is normal. The presence of visual spatial neglect or anosagnosia may also be taken as evidence of abnormality. Since the abnormality is scored only if present, the item is never untestable.

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No abnormality.</td>
</tr>
<tr>
<td>1</td>
<td>Visual, tactile, auditory, spatial, or personal inattention or extinction to bilateral simultaneous stimulation in one of the sensory modalities.</td>
</tr>
<tr>
<td>2</td>
<td>Profound hemi-inattention or hemi-inattention to more than one modality. Does not recognize own hand or orient to only one side of space.</td>
</tr>
</tbody>
</table>
Table S1. Language testing items for PedNIHSS:

<table>
<thead>
<tr>
<th>Repetition</th>
<th>Each of 4 word-repetition tasks is presented:</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>Stop</td>
</tr>
<tr>
<td>b.</td>
<td>Stop and go</td>
</tr>
<tr>
<td>c.</td>
<td>If it rains we play inside</td>
</tr>
<tr>
<td>d.</td>
<td>The President lives in Washington</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reading</th>
<th>Each of 3 items is presented for the child to read in Fig 1. Adjust expectations according to child’s age/school level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naming</td>
<td>Pictures are presented and of a clock, pencil, skateboard, shirt, baseball, bicycle (Fig 2).</td>
</tr>
</tbody>
</table>

| Fluency and word finding        | The picture (Fig 3) is presented and the child is asked to describe what he/she sees.                              |

Stop

See the dog run

Little children like to play outdoors

Fig S1. Reading items for PedNIHSS
Fig. S2 Pictures to test naming for Item 9 Best Language of PedNIHSS