Role for Telemedicine in Acute Stroke

Feasibility and Reliability of Remote Administration of the NIH Stroke Scale

Saad Shafqat, MD, PhD; Joseph C. Kvedar, MD; Mary M. Guanci, RN, MSN; Yuchiao Chang, PhD; Lee H. Schwamm, MD

***Background and Purpose*—**Immediate access to physicians experienced in acute stroke treatment may improve clinical outcomes in patients with acute stroke. Interactive telemedicine can make stroke specialists available to assist in the evaluation of patients at multiple urban or remote rural facilities. We tested whether interrater agreement for the NIH Stroke Scale (NIHSS), a critical component of acute stroke assessment, would persist if performed over a telemedicine link.

***Methods*—**One bedside and 1 remote NIHSS score were independently obtained on each of 20 patients with ischemic stroke. The bedside examination was performed by a stroke neurologist at the patient's bedside. The remote examination was performed by a second stroke neurologist through an interactive high-speed audio-video link, assisted by a nurse at the patient's bedside. Kappa coefficients were calculated for concordance between bedside and remote scores.

***Results*—**Remote assessments took slightly longer than bedside assessments (mean 9.70 versus 6.55 minutes, $P<0.001$). NIHSS scores ranged from 1 through 24. Based on weighted κ coefficients, 4 items (orientation, motor arm, motor leg, and neglect) displayed excellent agreement, 6 items (language, dysarthria, sensation, visual fields, facial palsy, and gaze) displayed good agreement, and 2 items (commands and ataxia) displayed poor agreement. Total NIHSS scores obtained by bedside and remote methods were strongly correlated ($r=0.97$, $P<0.001$).

***Conclusions*—**The NIH Stroke Scale remains a swift and reliable clinical instrument when used over interactive video. Application of this technology can bring stroke expertise to the bedside, regardless of patient location. **(*Stroke*. 1999;30:2141-2145.)**

Key Words: reliability ■ stroke assessment ■ stroke, ischemic ■ telemedicine

Therapeutic intervention in acute ischemic stroke requires urgent patient evaluation to ensure that patients receive the best available care within critical time windows. Expertise in acute stroke treatment with thrombolytic agents is concentrated in medical centers with dedicated stroke programs or in individual physician practices that cover multiple hospitals. Results of the NINDS Recombinant Tissue Plasminogen Activator (rt-PA) Stroke Trial[1] and the Trial of Prourokinase in Acute Cerebral Thromboembolism (PROACT II)[2] have established specific eligibility criteria for therapeutic intervention in acute stroke and require intervention within 3 or 6 hours, respectively. For patients eligible for intravenous rt-PA, immediate decisions regarding treatment are necessary and must be based on clinical assessment as well as interpretation of subtle CT imaging abnormalities. Delay in treatment due to physician travel time must be minimized. For patients best served by intra-arterial thrombolytic therapy, expeditious transfer to specialized stroke facilities will be required. Clinical decision making may be facilitated by input from remotely located stroke specialists who can review CT images online (teleradiology) and perform clinical assessments through interactive audio and video (interactive telemedicine).[3] With advances in the rapid transmission of real-time, full-motion video on broad bandwidth networks, this technology enables the continuous availability of stroke specialists to participate in the care of acute stroke patients at multiple remote sites.

Because measurement of the degree of neurological deficit is a critical first step in acute stroke evaluation, telemedicine in this setting requires audio-video data transmission with enough speed and clarity to ensure adequate neurological evaluation. The National Institutes of Health Stroke Scale (NIHSS) is a validated, 13-item examination tool for measuring stroke deficit with established interrater reliability.[4–8] It contains most of the neurological elements needed for clinical decision making and has been used as an entry criterion in

Received April 9, 1999; final revision received June 17, 1999; accepted July 13, 1999.

From the Departments of Neurology (S.S., M.M.G., L.H.S.), Dermatology (J.C.K.), and Medicine (Y.C.), Massachusetts General Hospital, and Center for Telemedicine, Partners Healthcare Inc (S.S., J.C.K., L.H.S.), Boston, Mass.

Correspondence to Lee H. Schwamm, MD, Department of Neurology, VBK 915, Massachusetts General Hospital, 55 Fruit St, Boston, MA 02114. E-mail lschwamm@partners.org

***Stroke* is available at http://www.strokeaha.org**

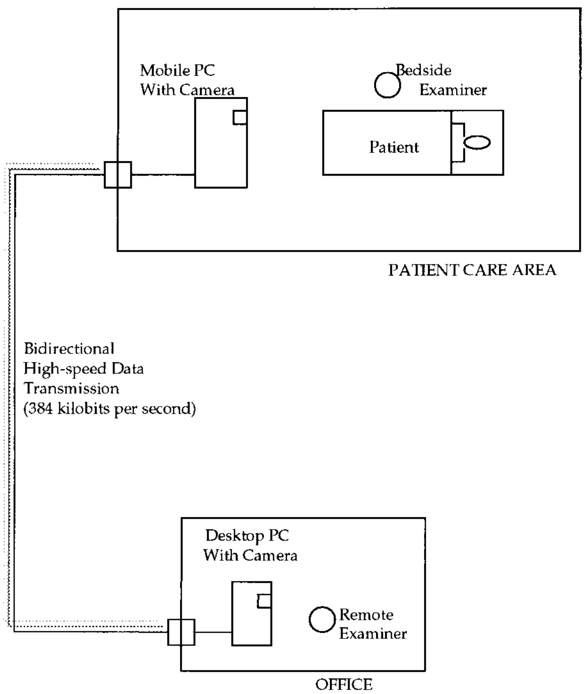

Figure 1. Schematic representation of the telemedicine link between patient and remote examiner.

trials of acute stroke therapy. To assess the feasibility of remote stroke evaluations, we tested the hypothesis that interrater reliability for the NIHSS would extend to a telemedicine paradigm.

Subjects and Methods

The study protocol was approved by our institutional review board. All patients (or their next of kin, if appropriate) gave informed consent to participate. The study subjects included 20 hospitalized adult patients with persistent deficits following acute ischemic stroke. On designated study days, all patients admitted with ischemic stroke were approached for possible inclusion in the study. We excluded patients who were clinically unstable as well as those who were intubated or had undergone an orthopedic procedure that prevented them from receiving a numerical score on each item of the NIHSS. There were 2 examiners in the study, a bedside neurologist and a remote neurologist paired with a bedside nurse. Both examiners were certified in administration of the NIHSS. The bedside neurologist examined the patient at the bedside in a conventional manner; the remote neurologist performed a telemedicine-enabled neurological examination remotely, assisted by a nurse at the patient's bedside. Thus, each patient received a remote and a bedside NIHSS score, which were statistically examined for concordance (see below). The remote and bedside neurologists had no prior clinical knowledge of the patients, and each was kept blinded to the examinations and scores of the other. For the initial 10 patients, the remote assessment was performed first; for the remaining 10, the order was reversed.

The telemedicine link is shown in schematic form in Figure 1. All examinations were performed with patients on stretchers, in front of a VTEL SmartStation TC 1000 (VTEL Corporation) and viewed remotely on a Pentium-based desktop personal computer with a proprietary Codec card (VTEL Corporation) and 21-inch color monitor (Mitsubishi Electric Corporation) set at 800×600 pixel resolution. The connection platform included a user interface, Windows 95–based conference control interface software, a PCI Codec card, an ISA TRI-BRI Promptus I-Mux card, and an integrated PTZ Sony camera (all from VTEL Corporation). An external microphone anchored by tape to the patient's pillow captured audio information. Data were transmitted over 3 parallel integrated services digital network (ISDN) lines at speeds up to 384 kilobits per second. The video transmission used standard compression algorithms with full CIF and was viewed at 30 frames per second. Two preset camera angles were used in all remote examinations: a zoom close-up for assessing facial weakness and eye movements and a wide angle for all other items (Figure 2). To minimize the need for shifting camera views, wide-angle items were performed first in the standard sequence,[5] followed by assessment of facial weakness and eye movements with close-up zoom. For every patient, both the bedside and the remote neurologists scored each item of the NIHSS and recorded the duration of each examination. For the language and dysarthria items, standard full-size NIHSS materials were used (cookie jar picture, naming sheet, word list, and sentence list). After completion of the NIHSS examinations, information on infarct location (categorized by vascular territory) and infarct size (categorized as small, medium, and large based on involvement of <one

Figure 2. Close-up zoom and wide-angle views of NIHSS administration over the telemedicine link. The close-up view is a montage of 4 individual screens. Some resolution is lost in transfer of screen images to print.

TABLE 1. Patient Characteristics and Total NIHSS Scores

Patient	Sex	Age, y	Infarct Vascular Distribution	Infarct Size	Days After Stroke	NIHSS Scores		Time, min	
						Bedside	Remote	Bedside	Remote
1	M	71	PICA	Small	2	2	1	4	6
2	M	54	PICA	Small	6	4	4	6	10
3	F	68	MCA	Small	15	2	2	4	8
4	M	63	PCA	Medium	21	6	6	6	10
5	M	69	MCA	Medium	60	5	6	6	8
6	M	82	Vertebral	Small	73	4	2	6	8
7	F	54	Lenticulostriate	Small	22	5	5	6	10
8	M	78	ACA-MCA borderzone	Medium	26	1	3	6	6
9	F	83	MCA	Medium	35	2	2	8	10
10	F	71	Lenticulostriate	Small	6	2	2	5	6
11	M	86	Basilar	Small	4	2	2	6	8
12	M	72	Vertebral	Small	22	3	5	6	8
13	F	74	MCA	Large	35	19	17	6	18
14	F	72	ACA-MCA-PCA borderzone	Large	21	12	15	12	10
15	M	66	MCA	Large	16	24	22	8	12
16	M	72	MCA	Medium	4	8	9	7	15
17	F	62	MCA	Medium	8	4	6	8	10
18	M	79	MCA	Large	15	15	17	8	12
19	F	79	PCA	Large	6	15	12	6	12
20	M	54	MCA	Medium	17	8	7	7	7

ACA indicates anterior cerebral artery; MCA, middle cerebral artery; PCA, posterior cerebral artery; and PICA, posterior inferior cerebellar artery.

third, $<$two thirds, or $>$two thirds of the affected vascular territory, respectively) was collected for each patient.

Interrater agreement between bedside and remote scores was measured by computing weighted κ statistics for each item of the NIHSS and by calculating the interrater correlation coefficient for the total NIHSS scores. In contrast to the unweighted κ statistic, the weighted κ method assigns weights to disagreements based on the magnitude of discrepancy[9] and may be a more accurate measure for a discrete ordinal scale such as the NIHSS. Published standards for interpretation of the κ statistic were used (values >0.75, excellent agreement beyond chance; values between 0.40 and 0.75, fair to good agreement beyond chance; values <0.40, poor agreement beyond chance).[9] Mean examination times for bedside versus remote assessments were compared using a 2-tailed Student *t*-test.

Results

As seen in Table 1, the study patients included 12 men and 8 women with a mean±SD age of 70.5±9.5 years (range 54 to 86 years) who had experienced an ischemic stroke requiring hospitalization within the preceding 73 days (median 16.5 days). Excluding system start-up time (typically 1 to 2 minutes), bedside assessments required a mean of 6.55 minutes (range 4 to 12 minutes) and remote assessments a mean of 9.70 minutes (range 6 to 18 minutes). This difference was statistically significant ($P<0.001$).

The NIHSS scores obtained ranged from 1 through 24. Bedside and remote examiners did not differ on any patient by >3 points. There was a strong linear correlation between total bedside and total remote scores (interrater correlation coefficient=0.97, $P<0.001$).

Based on weighted κ scores, 4 NIHSS items (orientation, motor arm, motor leg, and neglect) displayed excellent agreement beyond chance, 6 items (language, dysarthria, sensation, visual fields, facial palsy, and gaze) displayed good agreement beyond chance, and 3 items (level of arousal, commands, and ataxia) displayed poor to no agreement beyond chance. Table 2 compares κ coefficients for each scale item in our study with those reported from 2 other studies[4,5] of NIHSS interrater reliability.

Discussion

Telemedicine has been defined as "the use of telecommunications technologies to provide medical information and services."[3] Technically, this encompasses all aspects of medicine practiced at a distance, including use of telephone, fax, and electronic mail technology, as well as the use of interactive full-motion video and audio, that brings together patients and providers separated by distance.[10] Levine and Gorman[11] have proposed the term "telestroke" for the use of telemedicine in acute stroke intervention. This model offers several potential advantages for improving the care of stroke patients. The core steps of an acute stroke clinical encounter include rapid neurological assessment, review of brain imaging, and clinical formulation, including assessment of eligibility for thrombolytic therapy or neuroprotective clinical trials. With the potential to facilitate each of these steps,[10–12] telemedicine technology may provide specialists with the data necessary to assist clinicians at the bedside in stroke-related decision-making on patients presenting to distant facilities.

The current study demonstrates that it is feasible to perform remote neurological assessments with widely available and relatively inexpensive (under $10 000 and decreas-

TABLE 2. Kappa Statistics for Interrater Reliability of Individual NIHSS Items

		Prior Studies	
NIHSS Item	Current Study	Goldstein et al[4]	Brott et al[5]
Level of arousal	0†	0.50	0.49
Orientation	0.75	0.64	0.80
Commands	0.29*	0.41	0.58
Motor arm	0.82	0.77	0.85
Motor leg	0.83	0.78	0.83
Language	0.65	0.79	0.64
Dysarthria	0.55	0.32*	0.55
Ataxia	−0.07*	−0.16*	0.57
Sensory	0.48	0.50	0.60
Neglect	0.77	0.61	0.58
Visual fields	0.60	0.57	0.81
Facial palsy	0.40	0.22*	0.57
Best gaze	0.41	0.33*	0.82

*Items with poor interrater agreement.

†Bedside examiner assessed all 20 patients as awake and alert (score of 0), rendering the κ statistic meaningless for this item. Simple interrater agreement was 95%.

ing) telemedicine technology. The majority of studies attempting to validate telemedicine systems have focused on asynchronous processes, such as review of radiological and dermatologic images and of pathological specimens.[13] The NIHSS is a clinical tool with demonstrated interrater reliability that is frequently used to evaluate acute stroke patients and assess their outcomes after therapy.[6,14,15] When we compared itemized NIHSS scores obtained by bedside examination with scores for the same patients obtained remotely through a telemedicine link, the degree of interrater agreement on the various scale items paralleled the interrater agreement previously documented for bedside-bedside comparisons. Previous studies have used the difference of $\geq$4 points on the NIHSS to reflect a clinically significant change beyond interrater variability.[1,8,16] It is worth noting that the remote and bedside examiners in the current study did not differ in the total scores on any patient by $>$3 points (Table 1). As seen in Table 2, the κ coefficients obtained in our study compare favorably with the corresponding values reported by Goldstein et al[4] and Brott et al,[5] who studied the interrater agreement of bedside-administered NIHSS in 20 and 24 patients, respectively. The comparison suggests that certain NIHSS items (including dysarthria, ataxia, gaze, and facial palsy) have consistently inferior interrater agreement while other items (including orientation, motor arm, motor leg and language) have consistently superior interrater agreement. Our telemedicine-based study found fair interrater agreement for best gaze and facial palsy, suggesting that the technology provided adequate spatial and temporal resolution for close-up evaluation of the face and eyes (Figure 2). One of the items in our study (neglect) had greater interrater agreement compared with earlier reports. We also found that 1 of the items (performance of verbal commands) had poor interrater agreement compared with the results of Goldstein et al[4] and Brott et al.[5] However, there was no consistent pattern to the disagreement. The different pattern of interrater agreement in our study may result from the natural variability to be expected in studies of interrater reliability or it may reflect differences inherent in the telemedicine paradigm. The latter possibility suggests that detailed scoring and administration guidelines may have to be adapted for an NIH "telestroke" scale.

In its present form, the NIHSS is amenable to administration in a telemedicine environment. The modest increase in duration of telemedicine-administered NIHSS (average of 3.15 minutes compared with bedside administration) would not be expected to have a significant impact on clinical care. Proper testing of visual fields, sensation, and neglect requires assistance at the bedside. Additionally, while items such as language, ataxia, and performance of verbal commands are testable without assistance, their administration is facilitated by a bedside assistant. The bedside assistant need not have neurological expertise, because the specialist can expertly direct the assistant through interactive video. The adequacy of a nonspecialist assistant in this model suggests that remote administration of the NIHSS may be implemented in diverse settings, including patients in the emergency medical services system or prehospital arena. This may be important for tertiary care centers serving widely distributed, low-density rural populations. The use of training videotapes to demonstrate the performance of the NIHSS over a telemedicine link may enhance the reliability of "telestroke" assessments.[7]

To avoid interruption of the time-sensitive nature of acute stroke evaluation with an unproven modality (ie, telemedicine-administered neurological examination), we deliberately excluded patients in the acute phase. Recruitment of clinically stable patients also minimized examination variability attributable to clinical fluctuations in the acute setting. Because of the subacute nature of our test bed, the current data must be considered preliminary in determining their potential impact on actual clinical decision making. Our future research aims to reproduce the results of this study in acute and hyperacute stroke patients under real emergency room conditions. It is also important to note that while telemedicine-enabled neurological assessment may expedite stroke-related decision making, it cannot and should not be thought of as a substitute for the comprehensive clinical evaluation of the acute stroke patient, including thorough medical and cardiac evaluations.

Acknowledgments

The authors gratefully acknowledge the assistance of Carol Baer, RN, PhD, and Tommie-Sue Montgomery, PhD. Dr Shafqat is supported under a Partners Telemedicine Fellowship.

References

1. National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group. Tissue plasminogen activator for acute ischemic stroke. *N Engl J Med*. 1995;333:1581–1587.
2. Furlan AJ, Higashida R, Wechsler L, Schulz G, PROACT II Investigators. PROACT II: Recombinant prourokinase (r-ProUK) in acute cerebral thrombo-embolism. In: Program of the 24th American Heart Association International Conference on Stroke and Cerebral Circulation; February 4–6, 1999; Nashville, Tenn. Abstract 15.
3. Perednia DA, Allen A. Telemedicine technology and clinical applications. *JAMA*. 1995;273:483–488.

4. Goldstein LB, Bertels C, Davis JN. Interrater reliability of the NIH Stroke Scale. *Arch Neurol*. 1989;46:660–662.
5. Brott T, Adams HP, Olinger CP, Marler JR, Barsan WG, Biller J, Spilker J, Holleran R, Eberle R, Hertzberg V, Rorick M, Moomaw CJ, Walker M. Measurements of acute cerebral infarction: a clinical examination scale. *Stroke*. 1989;20:864–870.
6. Spilker J, Kongable G, Barch C, Braimah J, Brattina P, Daley S, Donnarumma R, Rapp K, Sailor S, NINDS rt-PA Stroke Study Group. Using the NIH Stroke Scale to assess stroke patients. *J Neurosci Nurs*. 1997;29:384–392.
7. Lyden P, Brott T, Tilley B, Welch KMA, Mascha EJ, Levine S, Haley EC, Grotta J, Marler J, NINDS rt-PA Stroke Study Group. Improved reliability of the NIH Stroke Scale using video training. *Stroke*. 1994;25: 2220–2226.
8. Wityk R, Pessin MS, Kaplan RF, Caplan LR. Serial assessment of acute stroke using the NIH Stroke Scale. *Stroke*. 1994;25:362–365.
9. Fleiss JL. *Statistical Methods for Rates and Proportions.* New York, NY: John Wiley & Sons Inc; 1981;212–236.
10. Grigsby J, Sanders JH. Telemedicine: where it is and where it's going. *Ann Intern Med*. 1998;129:123–127.
11. Levine SR, Gorman M. "Telestroke": the application of telemedicine for stroke. *Stroke*. 1999;30:464–469.
12. DeCorato DR, Kagetsu NJ, Ablow RC. Off-hours interpretation of radiologic images of patients admitted to the emergency department: efficacy of teleradiology. *AJR Am J Roentgenol*. 1995;165:1293–1296.
13. Taylor P. A survey of research in telemedicine, 2: telemedicine services. *J Telemed Telecare*. 1998;4:63–71.
14. Heinemann AW, Harvey RL, McGuire JR, Ingberman D, Lovell L, Semik P, Roth EJ. Measurement properties of the NIH Stroke Scale during acute rehabilitation. *Stroke*. 1997;28:1174–1180.
15. Lai S-M, Duncan P, Keighley J. Prediction of functional outcome after stroke: comparison of the Orpington Prognostic Scale and the NIH Stroke Scale. *Stroke*. 1998;29:1838–1842.
16. Clark WM, Warach SJ, Pettigrew LC, Gammans RE, Sabounjian LA, Citicoline Stroke Study Group. A randomized dose-response trial of citicoline in acute ischemic stroke patients. *Neurology*. 1997;49:671–678.