4-9-2005

Development and validation of the stroke action test

Susan Billings-Gagliardi
University of Massachusetts Medical School

Kathleen M. Mazor
University of Massachusetts Medical School

Stroke

JOURNAL OF THE AMERICAN HEART ASSOCIATION

American Stroke Association

A Division of American Heart Association

Development and Validation of the Stroke Action Test

Susan Billings-Gagliardi and Kathleen M. Mazor

Stroke 2005;36;1035-1039; originally published online Apr 7, 2005;
DOI: 10.1161/01.STR.0000162716.82295.ac
Stroke is published by the American Heart Association. 7272 Greenville Avenue, Dallas, TX 72514
 Print ISSN: 0039-2499. Online ISSN: 1524-4628

Development and Validation of the Stroke Action Test

Susan Billings-Gagliardi, PhD; Kathleen M. Mazor, EdD

Background and Purpose—Accurately assessing the public's readiness to respond to stroke is important. Most published measures are based on recall or recognition of stroke symptoms, or knowledge of the best action for stroke when the diagnosis is provided. The purpose of this study was to develop and evaluate a new written instrument whose items require the respondent to associate individual symptoms with the most appropriate action.

Methods—The Stroke Action Test (STAT) contains 21 items that name or describe stroke symptoms from all 5 groups of warning signs and 7 items that are nonstroke symptoms. For each item, the respondent selects 1 of 4 options: call 911, call doctor, wait 1 hour, or wait 1 day. The instrument validation sample included 249 subjects from community-based organizations. Score reliability and validity were analyzed using multiple data and information sources.

Results—The mean overall STAT score (all 28 items) for the lay people was 36.8%. On average, they chose call 911 for 34.1% of the stroke symptoms. They chose call doctor for 39.4% of the stroke symptoms, wait 1 hour for 20.1%, and wait 1 day for 6.0%. Score reliability is good (α=0.83). Evidence confirming score validity is presented based on analysis of item content and response patterns, and examination of the relationships between test scores and key variables related to stroke knowledge.

Conclusions—STAT directly assesses a critical aspect of practical stroke knowledge that has been largely overlooked and provides scores with good reliability and validity. **(*Stroke*. 2005;36:1035-1039.)**

Key Words: reproducibility of results ■ stroke ■ stroke assessment ■ survey instrument

To receive effective stroke therapy, lay people need to quickly access emergency medical services when early symptoms occur.[1,2] Accurately assessing this critical component of stroke knowledge is important for epidemiological research and for evaluating the effects of public education programs. The ultimate and most meaningful "test" is the behavior of patients and witnesses during actual strokes.[3,4] The challenge for researchers is to develop more practical assessments of this behavior, and to provide evidence of score validity.[5–7] In this context, validity evidence consists of information and data supporting the argument that assessment scores predict the actions that lay people would take if a stroke occurs.

Most recent assessments of stroke warning sign knowledge in English-speaking countries use *recall tasks*, in which respondents named as many warning signs of stroke as they could,[8–15] *recognition tasks*, in which respondents selected all stroke warning signs from short lists,[10,16–18] or responses to a general question such as, "What would you do first if you or a family member were having a stroke?" using an open-ended[9–12,15,16] or multiple-choice[18] format. However, the content of these assessments is not directly related to the task facing a person in an actual stroke situation. During a stroke, symptoms are experienced or observed. Identification of a symptom as a stroke warning sign is of limited value unless identification leads to urgent action. A person may know that a 911 call is the best response to stroke when the diagnosis is provided, but may not know to call 911 when a stroke symptom presents.[11] All of these factors confound interpretation of data that the assessments produce.

The purpose of this study was to develop and evaluate the Stroke Action Test (STAT), a new written instrument whose items require the respondent to associate a symptom with the most appropriate action. To investigate the validity of STAT scores, evidence was collected concerning item content (eg, do items approximate the task being measured, is the most important content emphasized, do examinees understand what is being asked, is the range of possible responses appropriate), reliability of scores, and relationship of scores to previously well-studied variables such as education, stroke experience, or reports of actual stroke situations.

Materials and Methods

STAT Content and Scoring

Content of the STAT is based on the consensus statement about stroke warning signs that major US organizations have agreed to use in their public education materials on stroke.[19–21] Using this statement as the framework, items were written by a neuroscientist and a psychometrician, and reviewed by a neurologist. The draft instrument was refined through an iterative process that included think-aloud interviews with lay people to test comprehension and wording,

Received December 21, 2004; final revision received January 20, 2005; accepted January 24, 2005.

From the Departments of Cell Biology and Neurology (S.B.-G.), University of Massachusetts Medical School, Worcester, Mass; Meyers Primary Care Institute (K.M.M.), Fallon Foundation, Worcester, Mass.

Correspondence to Susan Billings-Gagliardi, Department of Cell Biology (S7–147), University of Massachusetts Medical School, 55 Lake Avenue North, Worcester, MA 01655. E-mail susan.gagliardi@umassmed.edu

Stroke is available at http://www.strokeaha.org DOI: 10.1161/01.STR.0000162716.82295.ac

TABLE 1. Sample Test Items

If this happened to you or an adult friend/relative, what would you do?*
Sudden weakness of the face, especially on one side. (medical language, stroke symptom)
He was trying to eat lunch, but pieces of his sandwich kept falling out of the right side of his mouth. That hadn't ever happened before. (lay language, stroke symptom)
His finger joints were sore, and then a finger locked-up so he couldn't open his hand. (lay language, nonurgent medical condition)

*Response options: call 911 immediately, call doctor's office immediately, wait 1 hour then decide, wait 1 day then decide

score validation by clinicians, and pilot testing of the written form. The version of STAT evaluated in this study contains 28 items that name or describe a symptom. For each item, the task of the respondent is to answer the question, "If this happened to you or an adult friend/relative, what would you do?" by selecting 1 of 4 response options: (1) call 911 immediately; (2) call doctor's office immediately; (3) wait 1 hour and then decide; or (4) wait 1 day and then decide (Table 1). For scoring purposes, each correct response receives 1 point; incorrect responses receive 0 points. The total score is reported as percent of correct responses. Tests are either hand-scored or computer-scored after the answer sheet is scanned.

STAT items include 21 stroke symptoms representing all 5 groups of warning signs, as well as 7 nonstroke symptoms. Eleven items involve sudden unilateral numbness or weakness of the face, arm or leg, or trouble speaking or understanding. Two items contain a common stroke syndrome (eg, sudden right-side weakness of the face and arm, together with trouble speaking). The 7 nonstroke symptoms represent both urgent and nonurgent medical conditions. The overall SMOG Readability Index of the items is grade 7.[22] The complete STAT instrument can be accessed online at http://www.umassmed.edu/entities/cellbio/stat.cfm.

Additional Study Features Providing Data on Score Validity

To allow examination of the effect of item language on STAT scores, symptoms were presented in terse medical language (for the stroke symptoms, quotations from the consensus statement on stroke warning signs), or in lay language, based on descriptions given by lay people who had personally experienced or observed that symptom. For direct comparisons, 8 stroke symptoms were presented in *both* medical and lay language. To facilitate investigation of the STAT question format itself, 28 supplemental questions presented the same content but required only that the examinee decide whether a symptom was a stroke warning sign. To allow investigation of whether knowledge of the appropriate response to stroke (diagnosis provided) is equivalent to knowledge of the appropriate response to stroke symptoms, examinees were also asked to respond to the statement, "The first thing to do if you think you are having a stroke is to call 911," by selecting 1 of 5 options, ranging from strongly agree to strongly disagree.

Instrument Validation Sample

Two hundred forty-nine examinees were recruited from community-based organizations in Central Massachusetts. Criteria for inclusion were ages 25 to 75, self-assessed ability to read English, and lack of professional medical training. Examinees took the STAT in a witnessed small-group setting. Time to complete the test was observed, but no time limit was set. Examinees also provided demographic information, rated their own overall health, and reported their level of experience with stroke. For this study, experience with stroke was coded as "yes" or "no," in which "yes" was defined as any personal interactions with individuals who had strokes. Participants received a $7 stipend. To allow investigation of the effects of stroke training on scores, a group of first-year medical student (MS1) volunteers took STAT before (n=93) and after (n=72) 10 hours of instruction in stroke prevention and recognition.[23] The study received exempt status from the Institutional Review Board at the University of Massachusetts Medical School.

TABLE 2. Characteristics of Lay People (n=249)

Characteristic	Lay People Mean (SD) or %
Age, y	43.9 (12.0) (range 25–75)
Gender	
Male	50.8
Female	49.2
Race	
White (non-Hispanic)	45.2
Hispanic or Latino	27.4
Black or African American	19.4
Asian	2.0
American Indian or Alaska Native	1.6
Other or unreported	4.4
Education (highest level achieved)	
Less than high school graduation	18.5
High school or GED	27.8
Some college or 2-y degree	27.8
College graduate	25.8
Overall health (self-assessed)	
Excellent to very good	49.6
Good	35.0
Fair to poor	15.4
Personal experiences with stroke	
Yes	49.6
No	50.4

Analyses

Descriptive statistics were used to summarize the characteristics of subjects, mean item and test scores, and frequencies with which they chose each response option. Reliability of test scores was evaluated using Cronbach α. To gather additional validity evidence, scores of subgroups of examinees and of MS1 before and after stroke instruction were compared (*t* tests and paired *t* tests). The possible effect of item language on scores was evaluated by calculating the proportion of lay people correctly answering equivalent items that named a symptom in medical language or described the symptom in lay language (see Table 1). Results were compared using McNemar χ^2 tests.

Results

Demographic characteristics, self-reported health, and personal stroke experiences of the 249 lay people in the instrument validation sample are summarized in Table 2. The majority of examinees completed the 28-item STAT in ≈5 minutes; almost no one took >10 minutes. The mean overall STAT score (based on all 28 items) for these lay people was 36.8%. The mean score on the 21 items containing stroke symptoms was 34.1%. This means that on average, participants in this study chose call 911 for 34.1% of the stroke

TABLE 3. Comparison of Mean STAT Scores by Examinee Subgroups

Group	No.	STAT Score, % Mean (SD)	Significance
Lay people	249	36.8 (19.2)	$t=-2.6$
MS1 before training	93	42.2 (16.2)	$P=0.011$
Lay people, no stroke experiences	118	31.0 (18.6)	$t=-4.3$
Lay people with stroke experiences	120	41.4 (18.4)	$P<0.001$
Lay people, not college graduates	184	34.8 (18.8)	$t=-2.7$
Lay people, college graduates	64	42.1 (19.4)	$P=0.008$
MS1 before stroke instruction	72*	42.1 (17.0)	$t=-13.3$
MS1 after stroke instruction	72*	73.6 (14.5)	$P<0.001$

*Paired data.

TABLE 4. Effect of Language on STAT Scores for Equivalent Stroke Symptoms: Comparison of Lay People's Descriptions and Medical Statements

Stroke Symptom	% Correct Lay People's Descriptions	% Correct Medical Statements	Significance, P
Weakness of 1 side of face	47	39	NS, 0.013
Weakness of 1 arm	44 ↑	32	<0.001
Trouble speaking	47 ↑	36	0.002
Trouble understanding	37 ↑	23	<0.001
Trouble seeing in 1 eye	12	14	NS, 0.461
Loss of coordination	43 ↑	24	<0.001
Weakness of the face and arm on 1 side *and* trouble speaking	74 ↑	61	<0.001
Trouble seeing *and* dizziness	43 ↑	26	<0.001

For each symptom, the language producing higher scores ($P<0.01$) is designated by ↑.

symptoms. They chose call doctor for 39.4% of the stroke symptoms, wait 1 hour for 20.1%, and wait 1 day for 6.0%.

To investigate how examinees were using the different response options, we counted the number of times an individual chose each of the 4 options; 24.9% (62 individuals) chose call 911 for ≤2 of the stroke symptoms. Only 4.8% (12 individuals) chose call 911 for ≥16 of the stroke symptoms. Only 10% (25 individuals) chose call doctor for ≥16 of the stroke symptoms, only 1.2% (3 individuals) chose wait 1 hour, and only 0.4% (1 individual) chose wait 1 day. The latter result indicates that examinees tended to distribute their responses across the 3 incorrect options rather than consistently selecting the same incorrect response.

The stroke symptoms for which the fewest of the examinees would call 911 were sudden dizziness, sudden trouble seeing in one eye, and a transient visual loss. The symptoms for which the highest percentage would call 911 were sudden face and arm weakness that presented together with trouble speaking.

The reliability of the 28-item test was good ($\alpha=0.83$). Investigations of score validity included comparing mean STAT scores across subgroups of examinees. As Table 3 shows, STAT differentiated individuals with different levels of education and stroke training or experience, as would be expected. MS1s without training scored higher than lay people; lay people who reported personal stroke experiences scored higher than lay people without them; and lay people who were college graduates scored higher than lay people who were not. In addition, scores of the same individuals (MS1s) were significantly increased by intense instruction about stroke.

Table 4 compares scores of examinees on items that presented stroke symptoms or combinations of symptoms in lay language and in medical language (from the consensus statement). For 6 of 8 items, scores were higher when symptoms were described in lay language ($P<0.01$). For 2 symptoms, the difference was not statistically significant.

We also investigated whether similar items requiring only recognition of stroke symptoms produced higher scores than STAT items, which required a decision about an action. The mean overall score on recognition items was 47.8%, whereas the mean score on STAT action items was 34.1% ($P<0.001$). Individual stroke items showed the same trend, because scores for 8 of 11 were higher in the recognition format ($P<0.01$). Most examinees demonstrated knowledge of the correct response to stroke when the diagnosis was provided, because 93.8% agreed with the statement, "If you think you are having a stroke, the first thing to do is call 911 or an ambulance."

Discussion

The STAT is a new written instrument designed to assess lay people's knowledge of the correct response to individual stroke symptoms. Most lay people complete the test in ≈5 to 7 minutes, and it can be easily scored. The reliability of STAT scores, as estimated by the Cronbach α coefficient, is good. Data and information presented in Results also provide evidence that STAT scores are likely to reflect knowledge of the correct action to take if stroke occurs. This validity evidence is as follows.

First, STAT contains 28 items. The symptoms most often reported in stroke databases[24,25] are most heavily represented, and all 5 major groups of stroke warning signs are sampled. Seven nonstroke symptoms are included to broaden the range of correct responses, because if only stroke items were included then there would be only a single correct answer: call 911.

Second, there is a logical relationship between the task the test-taker must perform and the underlying behavior that the test seeks to predict: specifically, calling 911 in the event of a stroke.[6] Most previous measures are based on recall or recognition of stroke symptoms, or knowledge of the best action for stroke (diagnosis provided). By contrast, STAT assesses whether the respondent can connect symptoms and appropriate action. The importance of this distinction is underscored by our findings that 94% of examinees agreed that calling 911 is the best response to stroke; yet, on average, only 34% selected call 911 in response to specific symptoms. A similar discrepancy was reported in an epidemiological study of 882 Australian lay people conducted by telephone interviews.[11]

Third, STAT items are presented in lay and medical language, a decision made early in the process of designing the test in an effort to assure that the majority of test-takers would understand what was being asked. That decision is now supported by data showing that scores were higher on items that described stroke symptoms in lay language than on items that presented the same symptoms in the medical language of the consensus statement on stroke warning signs. We had expected that medical terms and terse style might increase scores because it would make a symptom appear more ominous or urgent.[26] Because just the opposite occurred, we now hypothesize that some of the examinees did not recognize the symptom in medical language, or did not understand what was meant.

Fourth, scores show positive relationships with other variables reported in the stroke literature. For example, examinees reporting personal experiences with stroke scored higher than those with none.[27] In addition, examinees scored highest on an item about sudden face and arm paralysis together with problems talking; these symptoms are among the most frequently reported by callers activating the emergency medical system in acute stroke.[28] Finally, as the MS1 data show, STAT scores are significantly improved by stroke training.

The purpose of this study was to develop and evaluate the STAT instrument, not to conduct an epidemiological investigation. Our study sample contains individuals with characteristics that are typical of the population to whom such a test would likely be administered in the future, including a range of ethnicities, ages, education, and experience. However, they are not intended to be a random sample of the US population at large. That fact acknowledged, it is disturbing that overall this sample of lay people would call 911 for only 34.1% of stroke symptoms. They chose an inappropriate response, call your doctor's office, more often than they chose call 911. Looking at the responses of each individual in the sample provides a different and more alarming perspective. Only 4.8% (12 people) chose call 911 for $\geq$16 of the 21 stroke symptoms. Often in testing, we consider 75% or more correct to be a "passing score." If we apply that criterion here, then only $\approx$5% of our sample "passed."

An important limitation of STAT, or any similar instrument, is that it can replicate only a fraction of what would actually be experienced by a patient or witness in an actual stroke situation. We need to continue working on ways to portray symptoms more realistically, perhaps with the use of multimedia technology, because this should further increase the predictive value of test scores. In the future, it would also be important to evaluate the instrument for use with groups of older or less healthy individuals, and to develop and validate a culturally sensitive Spanish language version.

In summary, STAT directly assesses a critical aspect of practical stroke knowledge that has been largely overlooked and provides scores with good reliability and validity. Our findings also focus attention on the importance of directing public education to the critical link between individual stroke symptoms and calling 911, and of using lay and medical language in describing stroke symptoms to the public. If STAT were given to a larger, population-based sample, responses to individual items could help identify which stroke symptoms lay people are least likely to respond to and permit more targeted public education efforts. Additionally, use of a standardized instrument, such as STAT, would greatly facilitate cross-study comparisons.

Acknowledgments

The development and evaluation of STAT was supported by a grant from the Meyers Primary Care Institute. We thank Nancy M. Fontneau, MD for reviewing test items and Joann Baril for technical assistance.

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. Kothari R, Jauch E, Broderick J, Brott T, Sauerbeck L, Khoury J, Liu T. Acute stroke: delays to presentation and emergency department evaluation. *Ann Emerg Med*. 1999;33:3–8.
3. Williams LS, Bruno A, Rouch D, Marriott DJ. Stroke patients' knowledge of stroke: influence on time of presentation. *Stroke*. 1997;28:912–915.
4. Wein TH, Staub L, Felberg R, Hickenbottom SL, Chan W, Grotta JC, Demchuk AM, Groff J, Bartholomew LK, Morgenstern LB. Activation of emergency medical services for acute stroke in a nonurban population: the T.L.L. Temple Foundation stroke project. *Stroke*. 2000;31:1925–1928.
5. Am Educational Research Association, Am Psychological Association, National Council on Measurement in Education. *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association; 1999.
6. Downing SM. Validity: on the meaningful interpretation of assessment data. *Med Educ*. 2003;37:830–837.
7. Kane MT. Current concerns in validity theory. *J Educ Measure*. 2001;38:319–342.
8. Pancioli AM, Broderick J, Kothari R, Brott T, Tuchfarber A, Miller R, Khoury J, Jauch E. Public perception of stroke warning signs and knowledge of potential risk factors. *JAMA*. 1998;279:1288–1292.
9. Becker KJ, Fruin MS, Gooding TD, Tirschwell DL, Love PJ, Mankowski TM. Community-based education improves stroke knowledge. *Cerebrovasc Dis*. 2001;11:34–43.
10. Rowe AK, Frankel MR, Sanders KA. Stroke awareness among Georgia adults: epidemiology and considerations regarding measurement. *South Med J*. 2001;94:613–618.
11. Yoon SS, Heller RF, Levi C, Wiggers J, Fitzgerald PE. Knowledge of stroke risk factors, warning symptoms, and treatment among an Australian urban population. *Stroke*. 2001;32:1926–1930.
12. Reeves MJ, Hogan JG, Rafferty AP. Knowledge of stroke risk factors and warning signs among Michigan adults. *Neurology*. 2002;59:1547–1552.
13. Schneider AT, Pancioli AM, Khoury JC, Rademacher E, Tuchfarber A, Miller R, Woo D, Kissela B, Broderick JP. Trends in community knowledge of the warning signs and risk factors for stroke. *JAMA*. 2003;289:343–346.
14. Silver FL, Rubini F, Black D, Hodgson CS. Advertising strategies to increase public knowledge of the warning signs of stroke. *Stroke*. 2003;34:1965–1969.
15. Carroll C, Hobart J, Fox C, Teare L, Gibson J. Stroke in Devon: knowledge was good, but action was poor. *J Neurol Neurosurg Psychiatry*. 2004;75:567–571.
16. Parahoo K, Thompson K, Cooper M, Stringer M, Ennis E, McCollam P. Stroke: awareness of the signs, symptoms and risk factors—a population-based survey. *Cerebrovasc Dis*. 2003;16:134–140.
17. DeLemos CD, Atkinson RP, Croopnick SL, Wentworth DA, Akins PT. How effective are "community" stroke screening programs at improving stroke knowledge and prevention practices? Results of a 3-month follow-up study. *Stroke*. 2003;34:e247–e249.
18. Greenlund KJ, Neff LJ, Zheng Z-J, Keenan NL, Giles WH, Ayala CA, Croft JB, Mensah GA. Low public recognition of major stroke symptoms. *Am J Prev Med*. 2003;25:315–319.

19. Learn to recognize a stroke. American Stroke Association Web site. Available at: http://www.strokeassociation.org/presenter.jhtml?identifier=1020. Accessed September 15, 2004.
20. Common symptoms and treatments. National Stroke Association Web site. Available at http://199.239.30.192/NationalStroke/SymptomsAndTreatments/default.htm. Accessed September 15, 2004.
21. NINDS stroke information page. Available at: http://www.ninds.nih.gov/health_and_medical/disorders/stroke.htm. Accessed September 15, 2004.
22. McLaughlin GH. SMOG grading—a new readability formula. *J Reading*. 1969;12:639–646.
23. Billings-Gagliardi S, Fontneau NM, Wolf MK, Barrett SV, Hademenos G, Mazor KM. Educating the next generation of physicians about stroke: incorporating stroke prevention into the medical school curriculum. *Stroke*. 2001;32:2854–2858.
24. Foulkes MA, Wolf PA, Price TR, Mohr JP, Hier DB. The stroke data bank: design, methods, and baseline characteristics. *Stroke*. 1988;19:547–554.
25. Bogousslavsky J, Van Melle G, Regli F. The Lausanne stroke registry: analysis of 1,000 consecutive patients with first stroke. *Stroke*. 1988;19:1083–1092.
26. Norman GR, Arfai B, Gupta A, Brooks LR, Eva KW. The privileged status of prestigious terminology: impact of "medicalese" on clinical judgments. *Acad Med*. 2003;10:S82–84.
27. Weltermann BM, Homann J, Rogalewski A, Brach S, Voss S, Ringelstein EB. Stroke knowledge among stroke support group members. *Stroke*. 2000;31:1230–1233.
28. Handschu R, Poppe R, Rauβ J, Neundörfer B, Erbguth F. Emergency calls in acute stroke. *Stroke*. 2003;34:1005–1009.