Cranial Computed Tomography Interpretation in Acute Stroke

Physician Accuracy in Determining Eligibility for Thrombolytic Therapy

David L. Schriger, MD, MPH; Mary Kalafut, MD, MS; Sidney Starkman, MD; Michelle Krueger, MD; Jeffrey L. Saver, MD

Context.—Intracranial hemorrhage must be excluded prior to administration of thrombolytic agents in acute stroke.

Objective.—To evaluate physician accuracy in cranial computed tomography scan interpretation for determining eligibility for thrombolytic therapy in acute stroke.

Design.—Administration of randomly selected, randomly ordered series of 15 computed tomography scans from a pool of 54 scans that demonstrated intracerebral hemorrhage, acute infarction, intracerebral calcifications (impostor for hemorrhage), old cerebral infarction (impostor for acute infarction), and normal findings.

Participants.—A convenience sample of 38 emergency physicians, 29 neurologists, and 36 general radiologists.

Main Outcome Measures.—Physician determination of eligibility for thrombolytic therapy based on computed tomography scan interpretation.

Results.—Average correct score by all physicians on all computed tomography scans was 77% (95% confidence interval, 74%-80%). Of 569 computed tomography readings by emergency physicians, 67% were correct; of 435 readings by neurologists, 83% were correct; and of 540 readings by radiologists, 83% were correct. Overall sensitivity for detecting hemorrhage was 82% (95% confidence interval, 78%-85%); 17% of emergency physicians, 40% of neurologists, and 52% of radiologists achieved 100% sensitivity for identification of hemorrhage.

Conclusion.—Physicians in this study did not uniformly achieve a level of sensitivity for identification of intracerebral hemorrhage sufficient to permit safe selection of candidates for thrombolytic therapy.

JAMA. 1998;279:1293-1297

From the Emergency Medicine Center (Drs Schriger, Starkman, and Krueger) and the Department of Neurology (Drs Kalafut, Starkman, and Saver), School of Medicine, and the Stroke Center (Drs Kalafut, Starkman, and Saver), University of California at Los Angeles.

Dr Starkman has received contracts and grants as an investigator in acute stroke studies sponsored by Genentech Inc, South San Francisco, Calif, and Boehringer-Ingelheim Pharmaceuticals Inc, Ridgefield, Conn. He also is a member of the Acute Ischemic Stroke Advisory Board and the Stroke Speakers' Bureau, and he has participated in stroke conferences sponsored by Genentech Inc and Boehringer-Ingelheim Pharmaceuticals Inc.

Dr Saver has participated as an investigator in research trials funded by Genentech Inc and Boehringer-Ingelheim Pharmaceuticals Inc, and has received honoraria from Genentech Inc for educational speeches. He also has served on a scientific advisory board for Boehringer-Ingelheim Pharmaceuticals Inc, Roche Laboratories Inc, Nutley, NJ, has a substantial interest in Genentech Inc. These companies have supported the activities of the University of California at Los Angeles Stroke Center with unrestricted educational grants.

Corresponding author: David L. Schriger, MD, MPH, 924 Westwood Blvd, Suite 300, Los Angeles, CA 90024-2924 (e-mail: schriger@ucla.edu).

RECOMBINANT tissue plasminogen activator has been approved for select patients within 3 hours of onset of acute ischemic stroke. Since thrombolytic therapy may produce lethal bleeding in patients with intracranial hemorrhage, the presence of intracranial blood on the initial computed tomography (CT) scan has been an exclusion criterion in the 5 trials of thrombolytic therapy for stroke and in the recommendations of expert panels. Thus, before thrombolytic therapy may be given in stroke, a physician highly skilled in identifying intracranial hemorrhage must interpret the CT scan.

For editorial comment see p 1307.

Early signs of major cerebral infarction (sulcal effacement, mass effect, and edema) also are associated with an increased risk for intracerebral hemorrhage in patients who receive thrombolytic therapy, and some guidelines recommend avoiding thrombolytic therapy when these findings are present.

We conducted this study to determine how well emergency physicians, neurologists, and general radiologists identified cranial CT scans that have evidence of intracranial hemorrhage. We also assessed these physicians' accuracy in interpreting other CT findings commonly seen in patients with acute ischemic stroke to examine whether physicians could distinguish CT scans that demonstrate subtle imaging abnormalities representing po-
tential contraindications to thrombolytics from those that do not preclude safe administration of these agents.

METHODS
Development of Scan Library

We reviewed interpretations of all cranial CT scans performed at a university teaching hospital from December 1994 to January 1996 to identify scans exhibiting hemispheric parenchymal hemorrhage or early infarction. Scans with calcifications (used as impostors for hemorrhage), scans with old infarction (used as impostors for acute infarction), and normal scans without calcification also were identified. Scans with other abnormalities (including subarachnoid and extracerebral hemorrhage), scans with multiple findings, and scans that could not be definitively placed in 1 of these categories were excluded. Only scans for which there was unanimous diagnostic agreement among the authors (and a consulting neuroradiologist) were included. Using a consensus process, we classified each scan that demonstrated hemorrhage as easy or difficult to interpret and each scan that demonstrated acute infarction as easy, intermediate, or difficult to interpret based on the subtlety of the findings.

Subject Recruitment

Thirty-eight emergency physicians were recruited at the Scientific Assembly of the California Chapter of the American College of Emergency Physicians in May 1996. Seventeen community-based neurologists who periodically attend at a university hospital neurology clinic were tested in private sessions and 12 additional neurologists were tested while attending the University of California at Los Angeles (UCLA) Stroke Center symposium, held in October 1996. Seven radiologists were contacted through the physician directors of local community hospital radiology departments and tested at their hospitals. Twenty-nine other radiologists were tested at the 49th Annual Midwinter Radiological/Oncology Conference of the Los Angeles Radiological Society, held in January 1997. Board-certified and board-eligible radiologists who did not have additional training in neuroradiology were tested.

At the meetings, a booth was set up in the exhibits area of the conference with a sign reading “Test Your Skills at CT.” Subjects included physicians who spontaneously approached the booth and those who, when approached, agreed to participate in the study. At conference test sites, subjects were offered a T-shirt and a sign reading “Test Your Skills at CT.” Subjects included physicians who spontaneously approached the booth and those who, when approached, agreed to participate in the study. At conference test sites, subjects were offered a T-shirt as an incentive to participate and were informed that the physician in each specialty with the highest score would receive a textbook of his or her choice. Physicians who gave verbal informed consent to participate completed a single-sheet questionnaire regarding their age, years of clinical experience, residency training, board certification, and typical involvement in the reading of cranial CT scans. After completing the questionnaire, each subject was shown the CT scans on a view box in an individual session with unlimited time to interpret each scan. The study was approved by the UCLA Institutional Review Board.

Scan Presentation

Physician-subjects were asked to assume that each scan was of a patient who arrived at the hospital within the first few hours after the onset of an acute hemispheric neurologic deficit (eg, aphasia, hemiparesis). As each CT scan was presented, the physician-subject was told which side of the patient’s body was affected. Subjects were asked to accept that each patient was eligible for thrombolytic therapy provided the CT scan had no contraindications. For each patient, the physician was asked, “Based solely on scan findings, could thrombolytics be administered to this patient?” Answer choices were (1) yes; (2) no, because of hemorrhage; or (3) no, because of signs of acute infarction. A list of contraindications (hemorrhage, early hypodensity, mass effect, and shift) and a list of findings that did not preclude the administration of thrombolytics (calcification, atrophy, and old infarction) were provided to remind subjects of the criteria for this study. Subjects were informed that 20% to 60% of the scans would have no contraindication to thrombolytic therapy.

Each subject was presented with 5 initial scans: 2 difficult hemorrhages, 1 intermediate acute infarction, 1 impostor, and 1 normal (Figure). Subjects who responded correctly to all 5 scans were placed in an advanced track and were then presented with 10 scans: 3 difficult hemorrhages, 3 difficult acute infarctions, 1 intermediate acute infarction, 2 impostors, and 1 normal. Subjects who responded incorrectly to 1 or more of the first 5 scans were placed in the standard track and presented with 3 difficult hemorrhages, 1 easy hemorrhage, 1 intermediate acute infarction, 2 easy acute infarctions, 1 impostor, and 2 normal scans. The 2-track strategy was designed to ensure that subjects were given scans that were appropriate to their skill levels and would maximize the discriminative capacity of the test.

Testing Protocols

Two hundred protocols, each containing a script of 25 scans, were created.
Each protocol specified that a subject must receive an initial series of 5 scans, with the scans to be included and their order of presentation determined using the “uniform” random numbers function of STATA 5.0 (Stata Corp, College Station, Tex). By a similar method, additional series of 10 scans each were prepared for the standard and advanced tracks.

**Statistical Methods**

We designed the experiment to provide stable estimates of the performance of each subject, the difficulty of each scan, and the sensitivity for detecting hemorrhage, without unduly burdening the volunteers. Using a conservative simulation, we determined that 30 subjects per specialty, each reading 15 scans, would meet these goals. Ninety-five percent confidence intervals (CIs) surrounding percent correct values were calculated using robust clustered logistic regression, which accounts for the fact that the scan readings may be associated with the skills of the reader and, therefore, are not completely independent.15

### RESULTS

All 29 neurologists, all 36 radiologists, and 74% of the 38 emergency physicians were board-certified (another 13% of the emergency physicians were senior residents in emergency medicine). Emergency physicians averaged 9 years in postresidency practice and 36 clinical hours per week, neurologists averaged 13 years in practice and 42 clinical hours per week, and radiologists averaged 15 years in practice and 42 clinical hours per week. Twenty-four percent of emergency physicians routinely read cranial CT scans. All neurologists reported reading CT scans; roughly half of them did so before seeing the radiologist’s report. All of the radiologists spent some time reading cranial CT scans (15% of clinical practice on average), although 22% reported that the interpretation of cranial CT and magnetic resonance imaging scans constituted less than 5% of their practice.

There were few violations in protocol. One emergency physician was inadvertently given only 14 scans. One radiologist was placed in the standard track despite achieving a perfect score on the first 5 scans. Also, during testing of the first group of radiologists it became clear that the lowest stratum of overall performance, easy hemorrhages, intermediate acute infarctions, normal scans, and difficult acute infarctions (40% correct) (Table 2). The correlation of percent correct with scan difficulty (easy, intermediate, or difficult) substantiated the validity of our classification schema.

Subjects in the advanced track (those who responded correctly to the first 5 scans) consistently scored higher on each type of scan than those in the standard track (Table 3). For all but those in the lowest stratum of overall performance, easy hemorrhages and easy acute infarctions were read with near perfect sensitivity. Scores for other types of scans improved with increasing overall skill.

**COMMENT**

The United States, recombinant tissue plasminogen activator is approved for use in acute ischemic stroke when administered within 180 minutes of symptom onset.1 During this time interval the patient must recognize the symptoms, get to a hospital, undergo evaluation, and have a CT scan performed and interpreted. There will be situations when the 3-hour limit is rapidly approaching, no neuroradiologist is available, and an emergency physician, neurologist, or general radiologist is the only physician available to interpret the CT scan. Our study asks whether physicians in these specialties are capable, without additional training, of interpreting the CT scan with sufficient sensitivity to de-
terminate if thrombolytics may be administered safely. Overall sensitivity for intracerebral hemorrhage, an absolute contraindication to thrombolytic therapy, was 82%. Two levels of sensitivity for hemorrhage may be expected of physicians who interpret cranial CT scans, one based on a single-case perspective, the other based on a population perspective. From the single-case perspective, the administration of a thrombolytic agent to a patient with an intracerebral hemorrhage may be lethal or have other catastrophic consequences. The potential consequences to the patient and the physician of administering thrombolytic therapy after failing to recognize hemorrhage on the CT scan mandate that sensitivity for intracerebral hemorrhage be extremely high, certainly over 95% and ideally higher than 99%.

A population-based perspective provides a more relaxed requirement for acceptable sensitivity. For instance, in the 624-patient National Institute of Neurological Disorders and Stroke study, there were 10 more survivors in the treatment group than in the control group at 3 months.4 Assuming that 15% of all acute hemispheric strokes are hemorrhagic,10 that patients with small hemorrhagic strokes have prognoses equal to those with ischemic stroke, and that all patients with hemorrhagic stroke who receive thrombolytics die, any sensitivity for hemorrhage higher than 75% will preserve a potential net death benefit in the treated group (unpublished data, D.L.S.). By similar reasoning, the sensitivity for hemorrhage may decrease to as low as 22% before negating improvements in neurologic outcome as measured with the Barthel Index.

We weigh heavily the single-case perspective and believe that, on average, the sample of physicians in our study did not have the skills needed to recognize hemorrhage on CT scans and determine which patients may safely receive thrombolytic therapy. One implication of this finding is that physicians in these specialties should not assume that the standard of care dictates that they should all be able to make these decisions independently.

Sensitivity for identification of early signs of major infarction was variable. Easy acute infarctions were identified by most physicians, but even the best performers failed to identify a third of the difficult acute infarctions. The clinical importance of this finding in patients with stroke within and beyond 3 hours of symptom onset remains to be determined.30,37,14 Poor specificity would deprive eligible patients the opportunity to benefit from thrombolytic therapy. While this is less catastrophic than administering thrombolytic agents to patients with contraindications, our study demonstrates that many physicians will have trouble differentiating hemorrhage from calcification and acute infarction from old infarction.

It is possible that our study results could represent a biased estimate of national average performance. Physicians in our study may not be representative of the national population of physicians in each specialty, leading to sampling bias. Furthermore, physicians taking a test may achieve results different from those they achieve in actual practice. Motivation in the simulated testing situation could be higher (eg, years of conditioning to perform well, desire to win the textbook) or lower (eg, no patient’s life is at stake) than in clinical practice. We cannot predict the direction of this bias.

The inability of participants to view CT scans on a computer and measure image density to differentiate calcium from blood may have adversely affected specificity on normal scans with calcification from blood may have adversely affected specificity on normal scans with calcification.
ing what would happen if the physician were the only person available to promptly read the scan. This approach most likely decreased specificity (those physicians who expressed uncertainty about a scan often scored it as infarction or hemorrhage because they did not wish to miss contraindications), but should not have affected sensitivity.

The few breaks in protocol also could affect our estimates, but any decrease in performance resulting from the 2 ambiguous scans was likely balanced by the 2 subjects who were mistakenly placed in the standard track (where they likely scored higher than they would have in the advanced track). Given the magnitude of our results, and the fact that many physicians were able to attain a perfect sensitivity for hemorrhage, we do not believe that these testing biases are large enough to alter our findings substantially.

Our convenience sampling method is the main threat to the external validity of the study. While board certification is an imperfect proxy for skill at interpreting cranial CT scans, all neurologists and radiologists were board certified, and the percentage of board-certified emergency physicians in the sample exceeded the national percentage of full-time emergency physicians who are board certified. While we cannot prove that our sample was representative of the national population of physicians in these specialties, we observed that physicians who acknowledged that they were uncomfortable reading CT scans often declined participation, suggesting that our study most likely did not underestimate physician performance. In addition, overall physician performance correlated with performance on each scan type, which suggests that the examination had construct validity.

In conclusion, it appears that while some members of each of these physician groups are capable of identifying hemorrhage with perfect or near-perfect sensitivity, the majority of those tested are not. Board certification in emergency medicine, neurology, or general radiology is an inadequate marker for such competence. Physicians involved in the care of patients with acute ischemic stroke should ensure that the interpretation of the CT scan reliably identifies intracranial hemorrhage when present. This may be accomplished by providing physicians with enhanced training in the interpretation of cranial CT scans or by implementing teleradiography or other systems that facilitate immediate scan interpretation by qualified readers.

The National Stroke Association (NSA) allowed us to use their exhibit booth for execution of this study, and underwrote the cost of the T-shirts and medical books. The NSA supported these activities with funds provided by an educational grant from Knoll Pharmaceutical Company, Whippany, NJ.

We offer our sincere thanks to the physicians who participated in this study. We also thank William H. Rogers, PhD, for help with the modeling of sample size considerations; Michael Zucker, MD, and Emily Nicklin for help finding and duplicating scans; Pablo Villalobananc, MD, neuroradiologist, for aid in scan interpretations; Danica Barron for data entry and data checking; and Jules Leshar McGhee for help with recruitment and enrollment at the study sites.

References

1. Nightingale SL. t-PA approved for acute ischemic stroke [From the Food and Drug Administration]. JAMA. 1996;276:443.