Interobserver Variation of ASPECTS in Real Time

Shelagh B. Coutts, MBChB; Andrew M. Demchuk, MD; Philip A. Barber, MBChB; William Y. Hu, MD; Jessica E. Simon, MBChB; Alastair M. Buchan, MD; Michael D. Hill, MD; for the VISION Study Group

Background—The Alberta Stroke Program Early CT Score (ASPECTS) has been used to quantify early ischemic changes on computed tomography (CT) brain scans of acute stroke patients. We sought to assess the reliability of the score when performed in real time as compared with an expert rating performed at a later time point.

Methods—Two hundred fourteen patients presenting with acute ischemic stroke or transient ischemic attack were prospectively recruited if they had a brain CT scan performed within 12 hours of symptom onset. Each scan was read for ASPECTS prospectively by the treating physician and later by 1 expert reader. A weighted kappa statistic was used to determine the interobserver agreement.

Results—The median baseline National Institutes of Health Stroke Scale score was 5 (range: 0 to 32) and the median time to CT scan was 152 minutes (range: 22 to 769). The interobserver agreement between ASPECTS performed in real time and expert ASPECTS was substantial ($\kappa = 0.69$). The mean difference between real-time ASPECTS and expert ASPECTS was 0 (SD: 1.1).

Conclusions—ASPECTS is a reliable clinical scale for rating early ischemic changes on CT when performed in real time.

(Stroke. 2004;35:e103-e105.)

Key Words: cerebral infarction • computed tomography

Thrombolysis for stroke within 3 hours of symptom onset remains the only approved acute treatment for ischemic stroke. Appropriate patient selection for thrombolysis in acute stroke is important and continues to be refined. Computed tomographic (CT) imaging may be central to this selection by identifying ischemic changes in the hyperacute phase.

The Alberta Stroke Program Early CT Score (ASPECTS) was developed to offer a reproducible grading system to assess early ischemic changes on pretreatment CT studies in patients with acute ischemic stroke. It is a 10-point scale that grades the extent of ischemic change within the territory of the middle cerebral artery. With good training and experience, the early changes of acute ischemia in the middle cerebral artery can be detected reliably. However, previous studies of ASPECTS have been performed with consensus assessment or not prospectively. In this article, we describe the reliability of ASPECTS rated prospectively by the treating physician as compared with the ASPECTS performed at a later date by an expert rater.

Patients and Methods

Prospectively, 214 patients were recruited into this study. Inclusion criteria were stroke or transient ischemic attack (consisting of hemiparesis or aphasia lasting >5 minutes) that were scanned within 12 hours of when the patient was last seen as being well, were older than 18 years, and were functionally independent on the modified Rankin scale (score ≤2). Patient demographics were recorded at the time of admission to the emergency department. The protocol was approved by the local institutional ethics review board.

Imaging

Standard noncontrast CT was performed with a fourth-generation multi-slice CT scanner (GE Medical Systems) in the emergency room. The noncontrast CT scanning technique was as follows: 120 kV, 170 mA, 2-second scan time, and 5-mm slice thickness. Coverage was from skull base to vertex with contiguous axial slices parallel to the inferior orbitomeatal line. A window width of 75 to 80 Hounsfield units (HU) and window level of 30 to 40 HU were used to maximize tissue contrast. Physicians were able to alter the window width and leveling as appropriate to maximize the appearance of ischemic changes.

Image Analysis

The ASPECTS was recorded prospectively by the treating physician at the time of CT scan. The treating physician (stroke fellow or stroke neurologist experienced in rating ASPECTS) scored the ASPECTS without knowledge of any other imaging modalities, but with knowledge of clinical symptoms. At a later point, 1 of 4 expert readers (different individual than the treating physician, including 1 neuroradiologist, 2 stroke neurologists, and 1 stroke fellow) rated the baseline CT scan using ASPECTS. The expert reader was blind to all clinical information except symptom side. Each scan was read by the
Statistics

Descriptive statistics were used to evaluate the study population. Interrater reliability between experts was not measured but has previously been shown to be equivalent. Agreement between the real-time and expert ASPECTS rating was assessed using weighted kappa ($\kappa_w$) scores. The weighting was designed to heavily penalize any difference $>1$ ASPECTS point. The $\kappa_w$ values were interpreted as: slight agreement, 0.00 to 0.20; fair agreement, 0.21 to 0.40; moderate agreement, 0.41 to 0.60; substantial agreement, 0.61 to 0.80; or almost perfect agreement, 0.81 to 1.00.8

The distribution of ASPECTS scores (real-time and expert) was skewed. The difference between the real-time and expert ASPECTS scores was normally distributed. The unit of analysis was the difference between the real-time and expert ASPECTS. Analysis of variance was used to assess the difference between categories of real-time ASPECTS values and the difference between the real-time and expert ASPECTS. Hypothesis testing between groups was adjusted using the Bonferroni method. Linear regression was used to plot this relationship.

Results

There were 214 patients included in this study; 88 were female. The median baseline National Institutes of Health Stroke Scale score (NIHSS) was 5 (range: 0 to 32). Median age was 72.5 (27 to 91). Fifty-four patients had a transient ischemic attack by the current definition9 (NIHSS was 5 (range: 0 to 32). Median time from symptom onset to CT scan was 152 minutes (range: 22 to 769 minutes).

Interobserver agreement between real-time and expert ASPECTS was substantial; $\kappa_w=0.69$ (95% confidence interval [CI]: 0.59 to 0.79). The mean difference was 0 (SD: 1.1). There was no difference in the reliability if transient ischemic attack patients were excluded ($\kappa_w=0.66$; $n=160$), or if only patients scanned at $<6$ hours are considered ($\kappa_w=0.68$; $n=183$), or if only stroke patients scanned at $<6$ hours are considered ($\kappa_w=0.65$; $n=140$).

Trichotomizing the ASPECTS scale into $<3$ (unfavorable), 3 to 7 (neutral), and 8 to 10 (favorable) had no impact on the mean differences between real-time and expert ASPECTS, except when the real-time ASPECTS was $<3$. In this situation, the expert reader was likely to call the scan $\approx 2$ ASPECTS points greater than the treating physician ($P=0.007$ for 3 to 7; $P<0.001$ for 8 to 10). There was a trend toward the treating physician undercalling the ischemic change in the 8-to-10 group, ie, the treating physician gives a clinically more favorable score than does the expert. However, this did not reach statistical significance ($P=0.064$).

This relationship was not affected by age, blood glucose, or the baseline NIHSS score.

Figure 1 gives a graphic representation of the correlation between the real-time ASPECTS and the difference between the real-time ASPECTS and the expert ASPECTS. The slope of the line suggests that, at lower real-time ASPECTS ratings, the treating physician is likely to overcall the presence of ischemic change.

At higher ASPECTS (>7) scores (favorable scan appearance), the real-time observer tends to undercall ischemic change. The effect size is quite small and is likely to be clinically insignificant. At lower ASPECTS (<3) scores (unfavorable scan appearance), the real-time observer tends to overcall the presence of ischemic change.

Conclusions

The reliability of rating ischemic change on the acute stroke CT in routine clinical practice compared with expert rating has not been previously evaluated to our knowledge. We have found that ASPECTS is reliable between real-time and expert ratings. This is important because the true significance of a clinical scale is its ability to be used in routine clinical practice. However, it must be emphasized that this clinical scale is only useful for the middle cerebral artery territory strokes and not the posterior cerebral artery or anterior cerebral artery.

Figure 1. Graphic representation of the real-time ASPECTS compared with the real-time ASPECTS and expert ASPECTS. The slope of the line suggests that at lower real-time ASPECTS ratings, the treating physician is likely to overcall the presence of ischemic change.

Figure 2. Overcalling in real time. A, B, and C, The representative cuts from a CT scan performed at 3 hours and 8 minutes after symptom onset on an acute stroke patient with right hemiparesis and aphasia (NIHSS 16). The treating physician rated the scan as having an ASPECTS of 4 with regions insula, M2, M3, M4, M5, and M6 rated as abnormal. Later expert review of the scan rated the scan as having an ASPECTS of 7 with regions insula, M2, and M3 rated as abnormal (MCA dot sign seen on cut A). D, E, and F, Diffusion-weighted magnetic resonance performed 4 hours and 40 minutes after symptom onset. An evolving infarct is seen in ASPECTS regions insula, M2, M3, and M5 (total score 6).
to overcall the ischemic change by nearly 2 points. The reasons for this probably reflect a combination of factors, including the human visual perception system’s tendency to overestimate boundaries.\textsuperscript{10} Human factors such as a desire not to administer thrombolysis may also spur the real-time observer to naturally err toward lower scores. There was no relationship between NIHSS and the differences seen between the real-time and expert ASPECTS. This suggests that the real-time physician does not overcall ischemic change because of bias introduced by knowledge of stroke severity. At the high end of the scale, the physician may be keen to offer thrombolysis and underinterpret the ischemic change. A final reason for disparity in scoring between the real-time and expert rating is that the treating physician often does not have ideal conditions for the rating of brain CT scan.

A weighted kappa was used for the statistics because artificially dichotomizing the results into $>7$ and $\leq 7$ would overly penalize a scan that an expert called a 7 and that the treating physician called a 6. A score of $\pm 1$ point on the ASPECTS scale has been shown to the expected margin of error,\textsuperscript{7} is felt not to be clinically significant, and would not deter thrombolytic treatment for any given patient. Hence, the kappa score was weighted to penalize any difference more than $\pm 1$ point. Our population included patients with transient ischemic attack and patients who were imaged between 6 and 12 hours after symptom onset. However, removing these patients from the analysis made no difference to the results.

The results in our study are limited to stroke neurologists and neuroradiologists. We did not assess the ability of emergency physicians or trainees to rate ASPECTS. This is an important limitation and future work is needed in this area.

In conclusion, we have found that ASPECTS performed in real time is a reliable method of quantifying the early ischemic changes of the middle cerebral artery in acute stroke.

Acknowledgments
This project received funding support from the Alberta Foundation for Health Research, Alberta Heritage Foundation for Medical Research, Heart and Stroke Foundation of Alberta, NWT and Nunavut, and Canadian Institutes of Health Research. Dr Barber was supported by Heart and Stroke Foundation of Canada (HSFC),

Canadian Institutes for Health Research, and Alberta Heritage. Dr Buchan was supported by the Heart and Stroke Foundation of Canada (HSFC), Canadian Institutes for Health Research, and The Canadian Stroke Network. Dr Coutts was supported by Heart and Stroke Foundation of Canada (HSFC) Fellowship and Alberta Heritage Foundation for Medical Research Fellowship. Dr Demchuk was supported by the Alberta Heritage Foundation for Medical Research and the Canadian Institutes for Health Research. Dr Hill was supported by Heart and Stroke Foundation of Alberta/NWT and Nunavut, and the Canadian Institutes for Health Research. Dr Simon was supported by an Alberta Heritage Foundation for Medical Research Fellowship. We acknowledge the following people within the VISION study group for their help with this study: Dr James Kennedy, Dr Alexis Gagnon, Dr Vanessa Palumbo, Dr Jayanta Roy, Dr Timothy Watson, Karyn Fischer, Linda Anderson-Armitage, Andrea Cole-Haskayne, Carol Kenney, Karla Ryckborst, Lisa Sinclair, Nancy Newcommon, and Marie McClelland.

References
Interobserver Variation of ASPECTS in Real Time
Shelagh B. Coutts, Andrew M. Demchuk, Philip A. Barber, William Y. Hu, Jessica E. Simon,
Alastair M. Buchan and Michael D. Hill
for the VISION Study Group

Stroke. 2004;35:e103-e105; originally published online April 8, 2004;
doi: 10.1161/01.STR.0000127082.19473.45
Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2004 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/35/5/e103