Atlas-Based Topographical Scoring for Magnetic Resonance Imaging of Acute Stroke

Robert K. Kosior, BSc; M. Louis Lauzon, PhD; Nikolai Steffenhagen, MD; Jayme C. Kosior, PhD; Andrew Demchuk, MD, PhD; Richard Frayne, PhD

Background and Purpose—The Alberta Stroke Program Early CT Score (ASPECTS), a 10-point scale, is a clinical tool for assessment of early ischemic changes after stroke based on the location and extent of a visible stroke lesion. It has been extended for use with MR diffusion-weighted imaging. The purpose of this work was to automate a MR topographical score (MR-TS) using a digital atlas to develop an objective tool for large-scale analyses and possibly reduce interrater variability and slice orientation differences.

Methods—We assessed 30 patients with acute ischemic stroke with a diffusion lesion who provided informed consent. Patients were imaged by CT and MRI within 24 hours of symptom onset. An MR-TS digital atlas was generated using the ASPECTS scoring sheet and anatomic MR data sets. Automated MR topographical scores (auto-MR-TS) were obtained based on the overlap of lesions on apparent diffusion coefficient maps with MR-TS atlas regions. Auto-MR-TS scores were then compared with scores derived manually (man-MR-TS) and with conventional CT ASPECTS scores.

Results—Of the 30 patients, 29 were assessed with auto-MR-TS. Auto-MR-TS was significantly lower than CT ASPECTS (P<0.001), but with a median difference of only 1 point. There was no significant difference between the auto-MR-TS and the man-MR-TS with a median difference of 0 points; 86% of patient scores differed by ≤1 point.

Conclusion—Auto-MR-TS provides a measure of stroke severity in an automated fashion and facilitates more objective, sensitive, and potentially more complex ASPECTS-based scoring. (Stroke. 2010;41:455-460.)

Key Words: acute stroke ■ brain infarction ■ brain ischemia ■ magnetic resonance ■ scales

Ischemic stroke results from reduced blood flow to brain tissue and can lead to infarction that is detectable by CT or by MR diffusion-weighted imaging (DWI).1 Neurological scoring methods such as the National Institutes of Health Stroke Scale score2 are used to quantify stroke severity to aid in the prediction of patient outcome and assessment of suitability for thrombolytic treatment with tissue plasminogen activator. Efforts have been made to improve prediction from the National Institutes of Health Stroke Scale score by combining it with imaging information such as with the Three-item Scale for the Prediction of Stroke Recovery, which is a composite of National Institutes of Health Stroke Scale, time from onset, and DWI infarct volume.3 The additional metric of DWI volume, however, is only a single feature from the breadth of information that could potentially be provided by MR stroke imaging.

The Alberta Stroke Program Early CT Score (ASPECTS) score is a topographical scoring system, that is, location-based.4 It is a 10-point scoring scale from which point deductions are made based on regional occupancy of an identifiable lesion on CT images.4 A small stroke lesion in a critical location can be far more serious than a large stroke encompassing a less critical region, which is what ASPECTS attempts to account for. The ASPECTS approach can be extended to DWI, because excellent intermodality agreement with CT has been demonstrated.5 Compared with CT, DWI has a higher contrast-to-noise ratio in acute infarction,6 which may facilitate computer automation and development of a topographical scoring method. Computer automation would provide an objective tool for large-scale analyses of data without pre-existing MR topographical score (MR-TS) data. Furthermore, computer automation may reduce interrater variability and slice orientation differences, which have been cited as sources of ASPECTS variability.7

The purpose of this work is to first extend the topographical ASPECTS method to MR DWI images, hereafter called MR-TS. Second, a computer automated (auto-MR-TS) implementation was developed and compared with manual scoring of the MR and CT images (man-MR-TS and CT ASPECTS). In a number of studies, ASPECTS has been

Received September 7, 2009; final revision received November 30, 2009; accepted December 3, 2009.

From Biomedical Engineering (R.K.K., R.F.), Electrical and Computer Engineering (J.C.K., R.F.), Radiology (M.L.L., R.F.), and Clinical Neurosciences (N.S., A.D., R.F.), University of Calgary, Calgary, Alberta, Canada; Hotchkiss Brain Institute (R.K.K., J.C.K., A.D., R.F.), University of Calgary, Calgary, Alberta, Canada; and Seaman Family MR Research Centre (R.K.K., M.L.L., N.S., J.C.K., A.D., R.F.), and Calgary Stroke Program (N.S., A.D.), Foothills Medical Centre, Alberta Health Services, Calgary, Alberta, Canada.

Correspondence to Richard Frayne, PhD, Seaman Family MR Research Centre, Foothills Medical Centre, 1403-29th Street NW, Calgary, Alberta, Canada T2N 2T9. E-mail rfrayne@ucalgary.ca

© 2010 American Heart Association, Inc.

Stroke is available at http://stroke.ahajournals.org

DOI: 10.1161/STROKEAHA.109.567289
validated and proven to show clear benefit7–9 and thus provides a strong clinical basis for our topographical scoring methods.

**Methods**

**Magnetic Resonance Imaging**

A retrospective study was performed using 30 patients with acute ischemic stroke with an acute DWI lesion. Informed consent was obtained. Patients were imaged by CT followed by MRI. All neuroimaging occurred within 24 hours of symptom onset. All MRI patient data were acquired on a 3-T scanner (Signa VH/i; General Electric Healthcare, Milwaukee, Wis) with a quadrature head coil. DWI images were acquired using a single-shot spin-echo echoplanar technique ($b=1000\, \text{s/mm}^2$, TR/TE/flip = 7000 ms to 9000 ms/73.1 to 93 ms/90°, 192×115 or 144×144 acquisition matrix, 32 cm×19.2 cm or 24 cm×24 cm field of view, and 19 slices, 5-mm thick with a 2-mm gap or 27 contiguous 5-mm slices) to evaluate the infarct. All patients received ASPECTS scores based on their noncontrast-enhanced CT scans.

A 3-dimensional MR-TS digital atlas was generated by manually tracing regions on T1 anatomic data sets (MNI, www.bic.mni.mcgill.ca/brainweb) at a resolution of $1\, \text{mm}^3$. A published reference of anatomic regions and blood supply territories was used in the region-drawing process.10

**Manual MR-TS**

Like with ASPECTS, man-MR-TS scores are calculated based on the occupancy of the lesion within the scoring regions.4 There are 12 predefined regions per hemisphere for a total of 24 regions (Figure 1), but lesion occupancy in the anterior or posterior blood supply territories does not warrant deductions with ASPECTS, leaving 10 eligible regions. Points are deducted from 10 where visible occupancy of a scored region within the lesion warrants a point deduction. A stroke fellow (N.S.) with ASPECTS training was permitted to use both the apparent diffusion coefficient (ADC) maps and the raw DWI data for man-MR-TS, because both sets of images would be used in clinical practice.

**Auto-MR-TS Calculations**

Auto-MR-TS was performed regionally using ADC maps and the MR-TS atlas based on lesion-region overlay in registered space (Figure 2). This methodology is similar to processing methods applied by other groups for CT11 and MR12 data. The infarct region on the ADC map was defined using a computer-assisted volumetric methodology using region-growing methods that has been previously validated.13 This segmentation method allows for tortuous, or geometrically irregular, regions to be selected. ADC lesions were segmented by (1) selecting a lesion with the mouse cursor; (2) adjusting an intensity threshold; and (3) manually adding or removing voxels to the computer-selected region if necessary (in cases where leakage occurred). As a stopping criterion for (2), an upper threshold was set to 80% of the mean of a region of interest placed in contralateral normal tissue on a slice with clear ADC deficit. Nonlinear registration was performed with SPM2 (Wellcome Department of Imaging Neuroscience, www.fil.ion.ucl.ac.uk/spm, 2004) to register the ADC map and lesion into the MNI brain atlas space. The T2-weighted baseline image with no diffusion weighting served as the source image to obtain the registration parameters that were applied to the ADC map and lesion.
An MR-TS score was calculated from the overlap of the lesion within region territories by computing the intersection between the various regions and lesion data. Sufficient occupancy was required to warrant a deduction based on user-defined occupancy thresholds. These thresholds were intended to control for cases where occupancy was negligible by visual inspection or subject to small preprocessing errors. The region threshold was defined as a percentage of occupancy of the lesion volume that must be attained by a region for consideration, that is, a threshold of 5% dictates that at least 5% of the lesion volume must intersect with a given region to be considered for deduction. The lesion threshold, on the other hand, is defined as a percentage of occupancy of the region volume that must be attained by a lesion for consideration. An absolute threshold was also set based on the total number of voxels in the lesion and region intersection. The region threshold was set to 5%, the lesion threshold was set to 10%, and the absolute threshold was set to 60 voxels (equivalent to 0.12 ml). Only one of the thresholds had to be satisfied to warrant a point deduction. Nonparametric Friedman and Wilcoxon signed rank tests were used to compare auto-MR-TS, man-MR-TS, and ASPECTS (with \(\alpha=0.05\) chosen as the significance level).

Visualization of the (1) ASPECTS regions; (2) overlaid segmented stroke lesion; and (3) 3-dimensional rendering of the lesion was performed using MatLab (7.4.0.287; MathWorks, 2007). For qualitative validation of the MR-TS regions and the segmented stroke, a visualization scheme was developed with overlay of the stroke lesion and ASPECTS regions onto a roadmap image such as the registered ADC map.

Results

One patient was excluded from stroke scoring due to the presence of MR image artifact. Specifically, ghosting was observed in the phase-encoded direction of the diffusion images, which effectively obscured the extent of the lesion on the ADC map so that proper lesion segmentation for auto-MR-TS could not be performed. Demographics and topographical scores are shown in Table 1 for the remaining 29 patients. These patients exhibited either no ghosting or only minor ghosting where the ghosting artifact was deemed to not interfere with the lesion segmentation. All scans were performed within 24 hours as per protocol, except for one patient (Patient 16) whose MR scan was performed 27 hours after stroke onset. The median delay time between CT and MR scans was 4.4 hours.

The median ASPECTS score for the 29 patients was 10 compared with a median of 8 from both auto-MR-TS and man-MR-TS (Table 1). These results indicate that the studied group represents patients with relatively mild strokes, because a threshold of 7 has been previously suggested to indicate high stroke severity and associated with poor patient outcome.4

There was a significant difference among the 3 scoring methods as summarized in Table 2 (Friedman test, \(P=0.0016\), by scoring technique: auto-MR-TS, man-MR-TS, and ASPECTS). Auto-MR-TS and man-MR-TS were both significantly lower than ASPECTS (\(P<0.0001\)), although the median difference was only 1 point between both MR-TS-derived scores and ASPECTS (Table 2). For the 4 patients with large discrepancies (difference \(\geq 3\) points), there were clear ADC lesions with changes not seen on CT (ie, ASPECTS=10).

For auto-MR-TS versus ASPECTS, 17 (59%) scores differed by \(\leq 1\) point, whereas 6 (21%) scores differed by \(\geq 3\) points. Five of the 6 patients (83%) with large discrepancies exhibited clear ADC lesions when no changes had been seen on CT (ie, ASPECTS=10). Thus, the score obtained from auto-MR-TS was consistently more severe than that obtained from ASPECTS. In only one patient was the score from auto-MR-TS less severe; in this case, it was higher by 1 point and the same as that determined by man-MR-TS (Patient 3).

There was no significant difference between auto-MR-TS and man-MR-TS (\(P=0.12\)) with a median difference of 0 points, and 25 (86%) scores differing by \(\leq 1\) point. Overall, consistent with the man-MR-TS versus ASPECTS comparison, auto-MR-TS yielded a score that was 1 point lower than from ASPECTS (median of 7 points versus 8 points). One score differed by 3 (Patient 13; Figure 3). In this patient, auto-MR-TS determined occupancy in the same regions as man-MR-TS as well as 3 additional adjacent regions (Figure 3). Figure 4 illustrates a case (Patient 28) in which man-MR-TS, auto-MR-TS, and ASPECTS yielded different scores (6, 7, and 8, respectively). Different deductions were made between the MR-TS methods because the lesion fell near the boundaries of a number of adjacent ASPECTS regions. The final scores, however, were in general agreement across methods (maximum difference of 2 points). The high contrast-to-noise ratio in the ADC maps (Figures 3 and 4) facilitated lesion segmentation.

The overlay of auto-MR-TS scoring regions and stroke lesion onto a roadmap provided for qualitative validation of the regions and the lesion as depicted in Figures 3 and 4. The overlay of the lesion allowed for qualitative verification that the scores were being computed correctly and that registration to atlas space was performed satisfactorily. This is seen as a useful adjunct step in verifying the automation process that can be applied to man-MR-TS as well.

Discussion

This study demonstrated an automated topographical scoring system for MR DWI. The differences between both MR-TS methods and ASPECTS are most likely attributable to the inherent differences in pathophysiologically derived image contrast between CT and MR in infarct assessment. DWI and ADC are more sensitive to the detection of early ischemic lesions suggesting that auto-MR-TS could simply be quantitative the lesion occupancy more sensitively. The high ADC lesion contrast allows for easier lesion segmentation that makes automation of MR-TS easier than for conventional ASPECTS. Furthermore, man-MR-TS and especially auto-MR-TS require negligible training compared with manual CT ASPECTS.16 Acquisition delay between CT and MR presents a confounding factor (median delay of 4.4 hours), although our results are in concordance with a previous comparison of manual MR and CT ASPECTS in which a median difference of 1 point was also observed despite there being a mean delay of only 1.7 hours for that study.5

The median ASPECTS scores were relatively high (8 for both MR-TS methods and 10 for CT ASPECTS), indicating that the studied group represents patients with relatively mild strokes, in which a score of \(\leq 7\) has been previously suggested to indicate high stroke severity associated with poor patient outcome independent of thrombolytic therapy (Table 2).4 It is important to note that MR stroke imaging and any associated topographical scoring will find greatest clinical importance if...
they add value to patients whose strokes are mild or moderate because most severe strokes can be more easily identified directly through clinical evaluation, CT scanning, and/or x-ray angiography.

The National Institutes of Health Stroke Scale score has seen widespread use and has been suggested to provide better prediction of patient outcome compared with lesion volume information acquired from MR stroke imaging.\textsuperscript{17} Topograph-

### Table 1. Summary of Demographics and Topographical Scores for Assessed Patients (17 [66\%] female, 12 [33\%] male, one patient excluded due to ghosting artifact)

<table>
<thead>
<tr>
<th>Patient No. (Gender)</th>
<th>Age, Years</th>
<th>Delay to CT</th>
<th>Delay to MR</th>
<th>Interscan Delay</th>
<th>Topographical Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ASPECTS</td>
</tr>
<tr>
<td>1 (F)</td>
<td>71</td>
<td>3.1</td>
<td>6.2</td>
<td>3.1</td>
<td>10</td>
</tr>
<tr>
<td>2 (F)</td>
<td>69</td>
<td>5.0</td>
<td>17.3</td>
<td>12.3</td>
<td>10</td>
</tr>
<tr>
<td>3 (F)</td>
<td>54</td>
<td>2.9</td>
<td>12.3</td>
<td>9.4</td>
<td>5</td>
</tr>
<tr>
<td>4 (F)</td>
<td>47</td>
<td>5.7</td>
<td>6.6</td>
<td>0.9</td>
<td>9</td>
</tr>
<tr>
<td>5 (M)</td>
<td>77</td>
<td>0.6</td>
<td>12.8</td>
<td>12.2</td>
<td>10</td>
</tr>
<tr>
<td>6 (F)</td>
<td>61</td>
<td>1.7</td>
<td>2.2</td>
<td>0.5</td>
<td>10</td>
</tr>
<tr>
<td>7 (M)</td>
<td>79</td>
<td>1.7</td>
<td>18.1</td>
<td>16.4</td>
<td>10</td>
</tr>
<tr>
<td>8 (M)</td>
<td>69</td>
<td>1.5</td>
<td>4.4</td>
<td>2.9</td>
<td>10</td>
</tr>
<tr>
<td>9 (F)</td>
<td>55</td>
<td>2.4</td>
<td>5.8</td>
<td>3.5</td>
<td>8</td>
</tr>
<tr>
<td>10 (M)</td>
<td>77</td>
<td>1.3</td>
<td>2.8</td>
<td>1.6</td>
<td>10</td>
</tr>
<tr>
<td>11 (M)</td>
<td>77</td>
<td>2.1</td>
<td>6.4</td>
<td>4.3</td>
<td>9</td>
</tr>
<tr>
<td>12 (F)</td>
<td>84</td>
<td>6.7</td>
<td>23.1</td>
<td>16.4</td>
<td>10</td>
</tr>
<tr>
<td>13 (F)</td>
<td>66</td>
<td>1.4</td>
<td>14.0</td>
<td>12.6</td>
<td>10</td>
</tr>
<tr>
<td>14 (F)</td>
<td>80</td>
<td>3.6</td>
<td>11.5</td>
<td>7.9</td>
<td>10</td>
</tr>
<tr>
<td>15 (F)</td>
<td>52</td>
<td>2.2</td>
<td>12.1</td>
<td>9.9</td>
<td>10</td>
</tr>
<tr>
<td>16 (M)</td>
<td>49</td>
<td>4.6</td>
<td>27.0*</td>
<td>22.4</td>
<td>5</td>
</tr>
<tr>
<td>17 (M)</td>
<td>74</td>
<td>9.6</td>
<td>11.7</td>
<td>2.1</td>
<td>9</td>
</tr>
<tr>
<td>18 (F)</td>
<td>85</td>
<td>4.4</td>
<td>14.7</td>
<td>10.3</td>
<td>10</td>
</tr>
<tr>
<td>19 (M)</td>
<td>64</td>
<td>8.0</td>
<td>22.4</td>
<td>14.4</td>
<td>10</td>
</tr>
<tr>
<td>20 (M)</td>
<td>52</td>
<td>7.9</td>
<td>9.2</td>
<td>1.3</td>
<td>10</td>
</tr>
<tr>
<td>21 (F)</td>
<td>43</td>
<td>1.6</td>
<td>4.3</td>
<td>2.7</td>
<td>8</td>
</tr>
<tr>
<td>22 (F)</td>
<td>76</td>
<td>3.4</td>
<td>4.8</td>
<td>1.4</td>
<td>10</td>
</tr>
<tr>
<td>23 (F)</td>
<td>77</td>
<td>3.5</td>
<td>15.0</td>
<td>11.5</td>
<td>9</td>
</tr>
<tr>
<td>24 (M)</td>
<td>70</td>
<td>10.7</td>
<td>15.1</td>
<td>4.4</td>
<td>10</td>
</tr>
<tr>
<td>25 (F)</td>
<td>63</td>
<td>2.7</td>
<td>13.8</td>
<td>11.1</td>
<td>10</td>
</tr>
<tr>
<td>26 (M)</td>
<td>69</td>
<td>1.0</td>
<td>18.1</td>
<td>17.1</td>
<td>8</td>
</tr>
<tr>
<td>27 (F)</td>
<td>72</td>
<td>1.3</td>
<td>4.9</td>
<td>3.6</td>
<td>8</td>
</tr>
<tr>
<td>28 (M)</td>
<td>76</td>
<td>1.2</td>
<td>3.7</td>
<td>2.6</td>
<td>8</td>
</tr>
<tr>
<td>29 (M)</td>
<td>27</td>
<td>2.2</td>
<td>5.2</td>
<td>3.0</td>
<td>10</td>
</tr>
</tbody>
</table>

Summary statistic

- Mean (SD): 66 (14) 3.6 (2.7) 11.2 (11.7) 7.6 (6.1)
- Median (IQR): 69 (22) 2.7 (3.0) 6.6 (8.8) 4.4 (9.5) 10 (1) 8 (2) 8 (2)

\*Patient scanned 3 hours outside of the 24-hour acute time window.

F indicates female; M, male; IQR, interquartile range.

### Table 2. Summary of Comparisons Among Topographical Scoring Methods*

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Median of Differences</th>
<th>P Value‡</th>
<th>Breakdown of Differences in Scoring Comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>-3</td>
</tr>
<tr>
<td>Man-MR-TS versus ASPECTS</td>
<td>-1</td>
<td>0.0001</td>
<td>4 (14%)</td>
</tr>
<tr>
<td>Auto-MR-TS versus ASPECTS</td>
<td>-1</td>
<td>&lt;0.0001</td>
<td>6 (21%)</td>
</tr>
<tr>
<td>Auto-MR-TS vs man-MR-TS</td>
<td>0</td>
<td>0.124</td>
<td>1 (3.4%)</td>
</tr>
</tbody>
</table>

*Initial Friedman test showed significant difference among the 3 methods (\textit{F}=55.1, \textit{P}=0.0016).

‡Significant at \(\alpha=0.05\) significance level (Wilcoxon signed-rank test).
ical scoring may improve prediction of patient outcome, because it has been found, for instance, that functional outcome (dependence and morbidity) using ASPECTS scoring has proven to have high sensitivity and very high specificity (0.78 and 0.96, respectively). Furthermore, it has been shown that the tandem use of National Institutes of Health Stroke Scale and ASPECTS yielded improved prediction of outcome, better than either scoring technique alone, for patients treated with intravenous tissue plasminogen activator. The ASPECTS method has been extensively validated, has been used to manage acute therapy, and has also been used to compare different stroke imaging techniques. For these reasons, we used the CT-based ASPECTS to develop our atlas-based MR-TS approach.

One limitation of our automated approach is that segmentation of regions demands some operator input. However, this input is minimized by the use of predefined ADC thresholds. Another limitation is that artifact can impinge on proper lesion segmentation. We excluded one patient due to aliasing artifact, although auto-MR-TS may be robust in the vast majority of cases (93% technical success rate in our study). Manual scoring, however, may still be necessary in some cases to distinguish between features arising from artifact or degraded signal information and features arising from stroke.

In our judgment, these cases represent only a small portion of examinations and do not mitigate the main impetus for auto-MR-TS. It is important to consider that even if the auto-MR-TS method is not suitable, perhaps due to image artifact, our auto-MR-TS visualization scheme depicted in Figures 3 and 4 can still be useful for guiding a final subjective score. The overlays can assist in arriving at a MR-TS score by indicating the general boundaries of the scoring regions, even when the automatic score is compromised by artifact.

We used region and lesion thresholds for occupancy calculations to control for false-positive deductions. We kept these thresholds as low as possible, at the same time still controlling for cases in which some overcalls seemed present, that is, cases in which occupancy was negligible. In some data sets for example, a large lesion encroached slightly into regions that subjectively appeared not warrant a deduction. The thresholds also help control for false deductions based on registration and segmentations errors. Although the auto-MR-TS approach is inherently more objective than ASPECTS, it still allows for subjective tuning of threshold parameters both for the lesion segmentation and the occupancy calculations.

Auto-MR-TS may be extended to include augmented versions of ASPECTS such as a version that includes post-
rior circulation flow called posterior circulation ASPECTS (pc-ASPECTS).21 Note that with basic ASPECTS scoring, only the subcortical and middle cerebral artery blood supply territories are considered for scoring, whereas lesion occupancy in the anterior or posterior blood supply territory regions are disregarded in the score. Another important application would incorporate perfusion imaging to estimate a penumbral volume and an auto-MR-WS mismatch grade, extending on a study that showed how perfusion-diffusion MR ASPECTS mismatch scoring is effective for stroke assessment.9 Further augmentations may now be considered based on the advantages afforded by auto-MR-WS automation steps. The basic ASPECTS methodology had to balance reproducibility (related to ease of use by a manual scorer) and sensitivity (related to how finely the regions were allocated). Auto-MR-WS overcomes problems associated with subjectivity and ASPECTS training, which obviated a more complex scoring scheme.

In conclusion, auto-MR-WS using digital brain atlas mapping facilitates larger-scale studies and provides for an objective, reproducible, whole brain approach to quantitative stroke severity scoring.

Sources of Funding
This work was supported by the Alberta Heritage Foundation for Medical Research (AHFMR), the Canada Foundation for Innovation (CFI), and the Canadian Institutes of Health Research (CIHR). R.K.K. and J.C.K. are supported by AHFMR, Natural Sciences and Engineering Research Council of Canada, and Informatics Cir of Research Excellence Graduate Scholarship awards. A.M.D. is an AHFMR Scholar. R.F. is a Canada Research Chair and an AHFMR Senior Scholar. A.D. has received a research grant from Novo Nordisk Canada.

Disclosures
A.D. has received financial support from Sanofi-Aventis and Boehringer Ingelheim through the Speakers’ Bureau and from Boehringer Ingelheim as a consultant. There are no other conflicts to report.

References
Atlas-Based Topographical Scoring for Magnetic Resonance Imaging of Acute Stroke
Robert K. Kosior, M. Louis Lauzon, Nikolai Steffenhagen, Jayme C. Kosior, Andrew Demchuk
and Richard Frayne

Stroke. 2010;41:455-460; originally published online January 21, 2010;
doi: 10.1161/STROKEAHA.109.567289
Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2010 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/41/3/455

Data Supplement (unedited) at:
http://stroke.ahajournals.org/content/suppl/2016/03/31/STROKEAHA.109.567289.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/
急性卒中磁共振成像的图像定位评分
Atlas-Based Topographical Scoring for Magnetic Resonance Imaging of Acute Stroke

Robert K. Kosior, BSc; M. Louis Lauzon, PhD; Nikolai Steffenhagen, MD; Jayme C. Kosior, PhD; Andrew Demchuk, MD, PhD; Richard Frayne, PhD

背景和目的: Alberta卒中早期CT评分(ASPECTS)——满分为10分的量表，是根据卒中病灶的位置和大小评估卒中后缺血性改变的临床工具,已被扩展到与磁共振(MR)弥散加权成像(DWI)一起使用。本研究的目的是用数字化图谱创建一个客观工具实现磁共振定位评分(MR-TS)自动化,用以大样本分析,并可能减少不同评估者间的差异及分层差异。


结果: 30例患者中,29例参与了自动MR-TS。自动MR-TS明显低于CT的ASPECTS(P<0.001),但是中位数仅相差1分。自动MR-TS与人工MR-TS没有显著差异,中位数相差0分。86%的患者的评分相差≤1分。

结论: 自动MR-TS提供了一种自动评估卒中严重程度的评分方法,使得建立在ASPECTS基础上的评分更加客观、敏感,也可能更复杂。

关键词: 急性卒中,脑梗死,脑缺血,磁共振,量表

(Stroke. 2010;41:455-460. 糜建华 译 李焰生 校)

脑组织血流量减少导致的缺血性卒中可产生CT或MRI的弥散加权成像(DWI)上可见的梗死[1]。国立卫生研究院卒中量表评分(NIHSS)评分[2]等神经科的评分方法可以量化卒中的严重程度,预测患者的预后,评估是否适合使用组织型纤溶酶原激活剂进行溶栓治疗。很多学者尝试着通过结合影像学信息来改善NIHSS评分的预见性,例如预测卒中恢复的三项评分,包括NIHSS评分、发病时间、DWI上梗死体积[3]。然而,DWI梗死体积仅仅是MRI成像所提供信息的一小部分。

Alberta卒中早期CT评分(ASPECTS)是一个基于病变部位的定位评分系统[4]。满分为10分,分数扣除是基于CT上确切病灶的局部占位[4]。重要部位小的梗死灶可能比其他部位的大片梗死后果更严重,这点是ASPECTS评分考虑到的。ASPECTS方法也可以用于DWI,因为其与CT模式的高度一致性已经被证实[4]。与CT相比, DWI 在急性脑梗死中声噪比更高[4],使得计算机自动化更容易,从而创造出一种地形评分方法。计算机自动化提供了一个不

需要MR定位评分数据就能分析大样本数据的客观工具,而且可以降低不同评估者间的差异及分层的差异,这些都是ASPECTS评分差异性的来源[7]。


方法
且存在 DWI 新鲜病灶的患者。获得知情同意后，患者先后进行 CT 及 MRI 检查。所有神经影像学检查在发病 24 小时之内进行。所有患者 MRI 数据都是在 3-T 扫描仪 (Signa VH/i; General Electric Healthcare, Milwaukee, Wis) 上通过头颅正交线圈获得的。DWI 图像则是通过单摄自旋回波平面扫描技术 ($b=1000 \text{ s/mm}^2$，脉冲时间 / 回波时间 / 自旋角度 = 7000-9000 ms/73.1-93 ms/90°，192 × 115 或 24 cm × 24 cm 视野，19 层，5 mm 厚，层间隔 2 mm 或 5 mm 厚连续不间断 27 层) 来评估梗死灶。所有患者通过非增强的 CT 扫描得出 ASPECTS。在 1mm×1mm×2mm 分辨率的 T1 解剖数据集 (MNI, www.bic.mni.mcgill.ca/brainweb) 上通过头颅正交线圈获得的。DWI 图像则是通过单摄自旋回波平面扫描技术 ($b=1000 \text{ s/mm}^2$，脉冲时间 / 回波时间 / 自旋角度 = 7000-9000 ms/73.1-93 ms/90°，192 × 115 或 144 × 144 采集矩阵，32 cm × 19.2 cm 或者 24 cm × 24 cm 视野，19 层，5 mm 厚，层间隔 2 mm 或 5 mm 厚连续不间断 27 层) 来评估梗死灶。所有患者通过非增强的 CT 扫描得出 ASPECTS。

人工 MR-TS

与 ASPECTS 评分相似，人工 MR-TS 分数依据评分部位的病灶计算得出 [4]。每侧大脑半球有预定的 12 个部位，两侧总共有 24 个部位 (图 1)，但是病灶位于前循环还是后循环并不能保证 ASPECTS 评分分数的扣除，所以只剩下 10 个合理的部位。满分 10 分，评分部位存在病变就扣除 1 分。一名接受过 ASPECTS 培训的卒中研究员被允许同时使用表观弥散系数 (ADC) 图和原始 DWI 数据进行人工评分，因为这两套图像都将在临床中使用。

自动 MR-TS

(3) 必要时对动添加或去除计算机选择的部位的体素 (例如当遗漏发生时)。作为步骤 2 的终止标准，上
取设置为存在明确 ADC 缺陷的对侧正常组织感兴趣
区分类之间平均的 80%。用 SPM2 软件 (Wellcome
Department of Imaging Neuroscience, www.fil.ion.ucl.ac.uk/spm, 2004) 完成非线性登记，将 ADC 图和病
灶录入 MRI 脑图中。没有 DWI 的 T2 加权像作为原
始图像用来获得将应用于 ADC 图和病灶的登记参数。

MR-TS 分数是通过计算各个区域与病灶叠加部
分的数据得出的。根据使用者定义的病灶阈值，病
灶足够充分才能保证扣分。阈值的设定目的是控制
那些目测无法识别的病灶或微小的预处理错误。区
域阈值是指病灶体积必须达到一定范围，即病灶体
积中至少 5% 与指定区域相交才可以扣分。而病灶
阈值则是指区域大小的百分比必须由病灶来衡量。
绝对阈值取决于病灶及区域相交部分的体素值。区
域阈值设定为 5%，病灶阈值设定为 10%，绝对
阈值设定为 60 个体素值 (相当于 0.12 mL)。要求达
到任何一项阈值才保证扣分。非参数 Friedman 和
Wilcoxon 符号秩检验用来比较自动 MR-TS、人工
MR-TS 和 ASPECTS 评分 (显著性水平 α=0.05)。

通过 MatLab(7.4.0.287; MathWorks, 2007) 显示 AS-
PECTS 区域、重叠的横断面界定的卒中病灶和病灶
的三维分布。为了定义验证 MR-TS 和横断面界定
的卒中，建立了一种可视化方案，即把卒中病灶和
ASPECTS 区域列在路线图成像上，如已登记的 ADC 图。

结果
由于出现 MRI 上的伪影，我们从卒中评分中剔
除 1 例患者。众所周知，弥散成像相位编码方向上
的镜像影会导致 ADC 图的病灶的边界模糊，因此准
确的病灶的自动 MR-TS 就不能完成。其余 29 例患
者的人口学资料和图像分析资料见表 1。这 29 例患
者的 MRI 没有镜像影或影很小不影响病灶划分。
所有检查根据计划在 24 小时内完成，仅 1 例是在卒
中发病后 27 小时完成 MRI 检查。CT 和 MRI 检查
的时间延迟的中位数为 4.4 小时。

29 例患者中 ASPECTS 评分中位数是 10。而自
动 MR-TS 和人工 MR-TS 的中位数都是 8(表 1)。结
果表明研究组患者是相对轻的卒中患者，因为先前
的研究表明≤7 分预示着卒中较严重且预后差 16。

表 2 显示了 3 种评分方法的显著性差异 (Fried-
man 检验，P=0.0016)。虽然 MR-TS 和 ASPECTS 评
分之差的中位数只是 1 分，手工和自动 MR-TS 的评
分均显著低于 ASPECTS 评分 (P<0.001)(表 2)。对
于 4 名有较大差异的患者 (相差≥3 分)，在 ADC 上
有明确的病灶但 CT 上却看不到 (ASPECTS=10)。

对于自动 MR-TS 与 ASPECTS 评分，17 例 (59%)
相差≤1 分，6 例 (21%) ≥3 分。6 例中 5 例 (83%) 存
在明确的 ADC 病灶但 CT 却看不到 (ASPECTS=10)。
所以自动 MR-TS 较 ASPECTS 评分更严谨。只有 1
例患者的自动 MR-TS 没有 ASPECTS 评分严重，比
之高 1 分，与人工 MR-TS 相同 (3 号患者)。

自动与人工 MR-TS 评分中没有显著性差异 (P=0.12)。
中位数差值为零，25 例患者 (86%) 相差≤1 分。总体上，
与人工 MR-TS 和 ASPECTS 评分的一致性相当，自
动 MR-TS 较 ASPECTS 评分低 1 分 (评分的中位数
分别为 7 分和 8 分)。只有 1 例患者相差 3 分 (13 号
患者，表 3)。这例患者自动 MR-TS 和人工 MR-TS
d 的评分在 3 个相邻区域内相同 (图 3)。图 4 显
示了 1 例患者的手工、自动 MR-TS 及 ASPECTS 评
分的分数各不相同 (分别是 6、7、8 分)。MR-TS 之
间不同的扣分方法是因为病灶位于 ASPECTS 评分
区的多个相邻边缘上。但是使用不同方法其最终的
分数大致上是一致的 (最多相差 2 分)。ADC 图 (图
3、4) 的高噪声有助于病灶界定。

如图 3 和图 4 中描述，自动 MR-TS 评分区域与
卒中病灶在路线图上的叠加提供了评分区域和病灶的
定性验证。这种病灶的叠加使得验证评分计算正确以
及在数字图上的记录被很好地完成。这是验证自动化
过程可用于人工 MR-TS 有效的一个重要的附加步骤。

讨论
本研究展示了自动化的磁共振 DWI 的定位评分
系统。MR-TS 与 ASPECTS 评分的不同主要来自于
梗死评估过程中内在的病理生理衍生的 CT 和 MRI
期缺血性病灶更加敏感，表明自动 MR-TS 可以更
加敏感地定位病灶大小。ADC 病灶的高对比性使
得病灶横断性界定更容易，从而使自动 MR-TS 较
传统的 ASPECTS 评分更简单。而且与手工的 CT
的 ASPECTS 评分相比，人工 MR-TS，尤其是自动
MR-TS 基本上不需要培训[15]。CT 和 MRI 间的检查
间隔时间是一个混杂因素 (中位间隔时间为 4.4 小
时)，虽然我们的结果与先前关于手动 MR 和 CT 的
ASPECTS 评分的比较一致 (中位数相差 1 分)，但
是后者平均延迟 1.7 小时[16]。

ASPECTS 评分中位数相对比较高 (自动与人工
MR-TS 均为 8 分，ASPECTS 评分为 10 分)，表明
研究组患者卒中相对较轻的患者，先前认为 AS-
表 1 受评估患者的人口统计和图像分析资料汇总 (17 名 [66%] 女性，12 名 [33%] 男性，1 名患者因为伪影被剔除)

<table>
<thead>
<tr>
<th>患者编号</th>
<th>性别</th>
<th>年龄</th>
<th>CT 检查时间</th>
<th>行磁共振检查时间</th>
<th>两者相隔时间</th>
<th>影像分析评分</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (F)</td>
<td>71</td>
<td>3.1</td>
<td>6.2</td>
<td>3.1</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>2 (F)</td>
<td>69</td>
<td>5.0</td>
<td>17.3</td>
<td>12.3</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>3 (F)</td>
<td>54</td>
<td>2.9</td>
<td>12.3</td>
<td>9.4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>4 (F)</td>
<td>47</td>
<td>5.7</td>
<td>6.6</td>
<td>0.9</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>5 (M)</td>
<td>77</td>
<td>0.6</td>
<td>12.8</td>
<td>12.2</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>6 (F)</td>
<td>61</td>
<td>1.7</td>
<td>2.2</td>
<td>0.5</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>7 (M)</td>
<td>79</td>
<td>1.7</td>
<td>18.1</td>
<td>16.4</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>8 (F)</td>
<td>69</td>
<td>1.5</td>
<td>4.4</td>
<td>2.9</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>9 (F)</td>
<td>55</td>
<td>2.4</td>
<td>5.8</td>
<td>3.5</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>10 (M)</td>
<td>77</td>
<td>1.3</td>
<td>2.8</td>
<td>1.6</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>11 (M)</td>
<td>77</td>
<td>2.1</td>
<td>6.4</td>
<td>4.3</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>12 (F)</td>
<td>84</td>
<td>6.7</td>
<td>23.1</td>
<td>16.4</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>13 (F)</td>
<td>66</td>
<td>1.4</td>
<td>14.0</td>
<td>12.6</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>14 (F)</td>
<td>80</td>
<td>3.6</td>
<td>11.5</td>
<td>7.9</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>15 (F)</td>
<td>52</td>
<td>2.2</td>
<td>12.1</td>
<td>9.9</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>16 (M)</td>
<td>49</td>
<td>4.6</td>
<td>27.0*</td>
<td>22.4</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>17 (M)</td>
<td>74</td>
<td>9.6</td>
<td>11.7</td>
<td>2.1</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>18 (F)</td>
<td>64</td>
<td>4.4</td>
<td>14.7</td>
<td>10.3</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>19 (M)</td>
<td>65</td>
<td>8.0</td>
<td>22.4</td>
<td>14.4</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>20 (M)</td>
<td>52</td>
<td>7.9</td>
<td>9.2</td>
<td>1.3</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>21 (F)</td>
<td>43</td>
<td>1.6</td>
<td>4.3</td>
<td>2.7</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>22 (F)</td>
<td>76</td>
<td>3.4</td>
<td>4.8</td>
<td>1.4</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>23 (F)</td>
<td>77</td>
<td>3.5</td>
<td>15.0</td>
<td>11.5</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>24 (M)</td>
<td>70</td>
<td>10.7</td>
<td>15.1</td>
<td>4.4</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>25 (F)</td>
<td>63</td>
<td>2.7</td>
<td>13.8</td>
<td>11.1</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>26 (M)</td>
<td>69</td>
<td>1.0</td>
<td>18.1</td>
<td>17.1</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>27 (F)</td>
<td>72</td>
<td>1.3</td>
<td>4.9</td>
<td>3.6</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>28 (M)</td>
<td>76</td>
<td>1.2</td>
<td>3.7</td>
<td>2.6</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>29 (M)</td>
<td>27</td>
<td>2.2</td>
<td>5.2</td>
<td>3.0</td>
<td>10</td>
<td>9</td>
</tr>
</tbody>
</table>

平均数 (标准差) 66 (14)  3.6 (2.7)  11.2 (11.7)  7.6 (6.1)
中位数 (四分位数) 69 (22)  2.7 (3.0)  6.6 (9.8)  4.4 (9.5)  10 (1)  8 (2)  8 (2)

* 超过 24 小时急性时间窗 3 小时行相关检查的患者。M 男性 F 女性

PECTS 评分≤7 分代表卒中中较为严重且无论是否溶栓治疗 (表 2)。值得注意的是，MR 卒中成像和任何相关的定位评分对轻中度卒中都将具有重要的临床意义，因为严重的卒中可以简单地直接通过临床评估、CT 扫描或 X 线血管造影来识别。

NIHSS 评分已经被广泛应用，而且较通过 MR 影像信息提供的梗死体积对预后的预测更加准确[17]。定位评分可能改善对预后的判断，因为已经发现利用 ASPECTS 评分判断功能预后 (生活独立和发病率) 的敏感性和特异性更高 (分别是 0.78 和 0.96)。而且已被表明合并使用 NIHSS 评分和 ASPECTS 评分对判断使用静脉组织型纤维蛋白溶酶原激活剂治疗患者的预后更准确，比单独使用任何一种评分更准确[18]。ASPECTS 方法已被广泛地证实并用于急性治疗和不同卒中影像技术的比较 [5,7-9,19,20]。正因为这些理由，我们利用 ASPECTS 方法来发展 MR-TS 方法。

表 2 形状评分方法间的比较 *

<table>
<thead>
<tr>
<th>比较类型</th>
<th>中位数差异</th>
<th>P 值†</th>
</tr>
</thead>
<tbody>
<tr>
<td>人工 MR-TS vs ASPECTS</td>
<td>-1</td>
<td>0.0001</td>
</tr>
<tr>
<td>自动 MR-TS vs ASPECTS</td>
<td>-1</td>
<td>0.0001</td>
</tr>
<tr>
<td>自动 MR-TS vs 人工 MR-TS</td>
<td>0</td>
<td>0.124</td>
</tr>
</tbody>
</table>

各评分差异

<table>
<thead>
<tr>
<th>差值</th>
<th>≤-3</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1</td>
<td>4 (14%)</td>
<td>9 (31%)</td>
<td>9 (31%)</td>
<td>5 (17%)</td>
<td>1 (3.4%)</td>
<td>1 (3.4%)</td>
</tr>
<tr>
<td>-2</td>
<td>6 (21%)</td>
<td>6 (21%)</td>
<td>13 (45%)</td>
<td>3 (10%)</td>
<td>1 (3.4%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>0</td>
<td>1 (3.4%)</td>
<td>2 (7%)</td>
<td>8 (28%)</td>
<td>13 (45%)</td>
<td>4 (14%)</td>
<td>1 (3.4%)</td>
</tr>
</tbody>
</table>

* 初始 F 检验表明 3 种评分方法间存在显著性差异 (F=55.1, P=0.0016)
† 显著性水平 α=0.05 (Wilcoxon 符号秩检验)
本自动化方法的一个局限是病灶横断界定需要操作者输入, 但采用预定义的 ADC 阈值输入可以达到最小化。另一个局限是人为因素可以影响正确的病灶横断界定。虽然自动 MR-TS 可以应用于绝大多数的病例 (此项研究中成功率为 93%), 但我们还是因为人为混杂而剔除了 1 例患者。人工评分在一些病例中仍然是必须的, 以便于区分人为因素、信号衰减或是卒中。我们认为这些病例仅仅代表了极小部分的患者,并不减弱自动 MR-TS 的重要意义。重要的是, 即使可能由于伪影而不适合自动 MR-TS, 我们的 MR-TS 可视化方案 (如图 3 和图 4 中描述) 仍对引导最终的主观评分有益。通过提示评分区域的大致边界, 当自动评分被伪影干扰时, 叠加仍可以帮助得出 MR 定位评分。

我们通过区域和病灶阈值来计算占位以控制假阳性扣分。我们保持阈值尽可能低, 同时控制可能存在的多扣分的情况, 也就是病灶可以忽略的病例。例如在一些数据集中, 大病灶轻度侵入, 而主观上好像还不值得扣分。阈值也可以帮助控制由记录或横断界定错误导致的错误扣分。虽然自动化 MR-TS 比 ASPECTS 评分更客观, 但它仍允许为了病灶横断界定和病灶体积计算对病灶的阈值参数进行主观调整。

自动 MR-TS 可被扩展为 ASPECTS 改进版, 例如包括后循环血流的 ASPECTS 评分被称之为后循环 ASPECTS (pc-ASPECTS) [21]。值得注意的是在应用基础 ASPECTS 评分时, 只有皮层下和大脑中动脉供血区在评分范围内, 而前动脉和后动脉供血区在评分过程中则被忽略。另外一项重要应用是整合灌注成像来估计半暗带体积和自动 MR 定位评分不匹配区域的梯度, 一项研究表明灌注 - 弥散 MR 的 ASPECTS 不匹配评分对卒中评估有效 [9]。进一步的扩展可能来自于自动 MR-TS 的自动化优势。基础的 ASPECTS 方法需要平衡可重复性 (与手工记录员的易操作性有关) 和敏感性 (与区域划分精细程度有关) 两者间的关系。自动化 MR 定位评分克服了主观性和 ASPECTS 需要培训的问题, 不会成为一项更加复杂的评分方案。

总之, 自动 MR 定位评分利用脑数字地形图推进了大样本研究, 并且提供了一个客观的、可重复的、定量地针对全脑的卒中严重程度评分方法。

参考文献 (略)