Color-Coded Fluid-Attenuated Inversion Recovery Images Improve Inter-Rater Reliability of Fluid-Attenuated Inversion Recovery Signal Changes Within Acute Diffusion-Weighted Image Lesions

Bum Joon Kim, MD; Yong-Hwan Kim, MS; Yeon-Jung Kim, MD; Sung Ho Ahn, MD; Deok Hee Lee, MD; Sun U. Kwon, MD; Sang Joon Kim, MD; Jong S. Kim, MD; Dong-Wha Kang, MD

Background and Purpose—Diffusion-weighted image fluid-attenuated inversion recovery (FLAIR) mismatch has been considered to represent ischemic lesion age. However, the inter-rater agreement of diffusion-weighted image FLAIR mismatch is low. We hypothesized that color-coded images would increase its inter-rater agreement.

Methods—Patients with ischemic stroke <24 hours of a clear onset were retrospectively studied. FLAIR signal change was rated as negative, subtle, or obvious on conventional and color-coded FLAIR images based on visual inspection. Inter-rater agreement was evaluated using κ and percent agreement. The predictive value of diffusion-weighted image FLAIR mismatch for identification of patients <4.5 hours of symptom onset was evaluated.

Results—One hundred and thirteen patients were enrolled. The inter-rater agreement of FLAIR signal change improved from 69.9% (κ=0.538) with conventional images to 85.8% (κ=0.754) with color-coded images (P=0.004). Discrepancy rated patients on conventional, but not on color-coded images, had a higher prevalence of cardioembolic stroke (P=0.02) and cortical infarction (P=0.04). The positive predictive value for patients <4.5 hours of onset was 85.3% and 71.9% with conventional and 95.7% and 82.1% with color-coded images, by each rater.

Conclusions—Color-coded FLAIR images increased the inter-rater agreement of diffusion-weighted image FLAIR recovery mismatch and may ultimately help identify unknown-onset stroke patients appropriate for thrombolysis. (Stroke. 2014;45:2801-2804.)

Key Words: magnetic resonance imaging ■ stroke

Mismatch of lesion visibility between diffusion-weighted image and fluid-attenuated inversion recovery image (DWI-FLAIR mismatch) is a useful surrogate marker for estimating the lesion age in ischemic stroke.1–3 Reperfusion therapy based on multimodal imaging, including the information on DWI-FLAIR mismatch, is considered feasible and safe in patients with unclear-onset stroke.4

However, FLAIR signal intensity within an acute DWI lesion is heterogeneous, and subtle changes limited to the partial region were often regarded as negative for FLAIR signal change.5 Because the lesion conspicuity is subjective, inter-rater agreement of FLAIR signal change was low, especially in subtle changes.6 Quantitative analysis of FLAIR signal intensity ratio failed to increase the specificity for predicting ischemic lesions within the therapeutic window7 or the inter-rater agreement on FLAIR signal change.8

Color coding may increase the visual distinguishability, thus enabling clinicians to identify small changes in signal intensity. We hypothesized that color-coded FLAIR images would enhance the inter-rater agreement of DWI-FLAIR mismatch.

Methods

Subjects
The records of patients with ischemic stroke admitted to Asan Medical Center <24 hours of clear onset between January and July 2013 were reviewed. Patients were excluded if (1) DWI lesion was absent, (2) images were obtained from other centers, or (3) imaging data quality was insufficient. Clinical data were obtained from a prospectively collected registry. The Institutional Review Board of the Asan Medical Center approved this study. Informed consent was not obtained because of the retrospective design.

Image Preparation
Patients received a 1.5-T MRI scan (Magnetom Avanto; Siemens). The DWI slice with the largest lesion and the slice of corresponding FLAIR image were prepared. The FLAIR image slice was extracted...
as a DICOM file, converted to a NIfTI file and was presented on MRicroN (http://www.nitrc.org/projects/mricron). The actc look-up table, which consists of 256 color codes, was used to assign the voxel color. The auto-contrast setting determined the intensity ranges (1–99 percentiles) of color codes. Finally, 3 corresponding slices of different sequences were prepared for each patient (Figure 1).

Measurement of Inter-Rater Agreement
Inter-rater agreement was measured in 2 pairs of raters blinded to clinical data (pair A: well-trained stroke neurologists; pair B: neurologists in residency training; see the online-only Data Supplement). To avoid a learning effect, rating was done with color-coded image first and a week later with the conventional image. FLAIR signal change was categorized into obvious (a hyperintense lesion within the DWI lesion), subtle (hyperintensity identified only with reference to the DWI lesion), or negative (no hyperintensity at all within the DWI lesion). Ten examples were provided as a reference before rating. FLAIR images rated negative were regarded as positive for DWI-FLAIR mismatch.

Statistical Analysis
κ Value and percent agreement were used to quantify inter-rater agreement of FLAIR signal change and McNemar test to compare the difference. Patients with discrepant ratings by conventional images but concordant with color-coded images were regarded as patients with benefit from color-coded FLAIR image. Characteristics of these patients were compared with those of other patients using χ² and t tests, appropriately. The predictive value of DWI-FLAIR mismatch for identification of patients <4.5 hours of symptom onset was also evaluated.

Results
Among the 155 patients enrolled, 42 patients were excluded (no FLAIR image, n=23; negative DWI, n=11; insufficient image quality or clinical information, n=8). Thus, 113 patients were included for final analysis: mean age was 67.0±12.6 years and 61 (54.0%) were men. The initial National Institutes of Health Stroke Scale score was 4±5. Patients received MRI at 7.0±7.1 hours from symptom onset. Fifteen patients (13.3%) received tissue plasminogen activator before and during MRI.

In pair A, the inter-rater agreement of DWI-FLAIR mismatch was 0.538 with conventional images and 0.754 with color-coded images (Figure 2). The percent agreement was significantly different between conventional and color-coded images (69.9% versus 85.8%, respectively; P=0.004). The proportion of patients categorized as subtle change was higher with conventional than with color-coded images (38.1% versus 25.7%, respectively; P=0.01). The inter-rater agreement was increased by color-coded image, especially in patients with obvious change (Figure 2).

Of the 113 patients, 34 patients (30.1%) demonstrated discrepant ratings with conventional images. It was resolved in 28 of 34 patients (82.4%) with color-coded images. Those 28 patients demonstrated a high proportion of cardioembolic stroke (46.4% versus 21.2%; P=0.02) and cortical lesion location (42.9% versus 17.6%; P=0.04) compared with other patients (Table I in the online-only Data Supplement). Color-coded FLAIR image decreased the proportion of obvious and subtle FLAIR signal change, especially in patients beyond the time window of thrombolysis (Figure 3). Therefore, the positive predictive value of DWI-FLAIR mismatch predicting lesion <4.5 hours of stroke onset was 85.3% and 71.9% with conventional FLAIR images and 95.7% and 82.1% with color-coded FLAIR images, respectively, by each rater.

Discussion
Color-coded FLAIR image, a simple and an easily applicable method by clinicians, increased the inter-rater agreement of FLAIR signal change, especially in patients with cortical lesions. These focal cortical lesions were previously disregarded from conventional FLAIR images, but may increase hemorrhagic risk after thrombolysis. Consistent with our results of conventional FLAIR images, recent studies that included patients with similar stroke severity to the present study demonstrated a moderate inter-rater agreement (κ=0.57 and 0.48). Semiautomatic measurement of signal intensity ratios inside a circular region of interest did not significantly improve inter-rater agreement. This may be related with the fact that the pathological status of infarcted tissue is heterogeneous, which was well demonstrated by color-coded FLAIR image.

Additionally, color-coded FLAIR images decreased the proportion of patients graded as subtle change. Previously, the inter-rater agreement and the correlation between time from stroke onset and FLAIR signal change was low in patients graded as subtle FLAIR change. Therefore, reducing
patients rated as subtle change may help increase the reliability of DWI-FLAIR mismatch in the clinical field. Finally, excluding patients >4.5 hours of symptom onset is important to avoid unexpected side effects of reperfusion therapy in the ongoing clinical trials with patients with unknown-onset stroke. Therefore, the high specificity and positive predictive value of color-coded FLAIR images may have advantage in safety concerns of thrombolysis (Table II in the online-only Data Supplement).

Limitations stemming from the possibility of overestimation and estimation bias occurring from learning effect cannot be excluded. However, restricting the area for evaluation to the DWI lesion and the previously described various efforts to minimize the learning effect may help resolve these issues.

Figure 2. Inter-rater agreements according to fluid-attenuated inversion recovery (FLAIR) signal change. *Statistically significant difference ($P<0.05$).

Figure 3. Fluid-attenuated inversion recovery (FLAIR) signal change according to time from stroke onset. *Table expressed by number of patients
Color coding of FLAIR images, an easily applicable technique, increased the inter-rater agreement of FLAIR signal change. It may be potentially beneficial in informing the decision regarding thrombolysis in patients with unknown-onset stroke.

Sources of Funding
This study was supported by grants from the Korea Health Technology R&D Project, Ministry for Health and Welfare, Korea (HI12C1847), and the Asan Institute for Life Sciences (No. 2013–365), Seoul, Korea.

Disclosures
None.

References
Color-Coded Fluid-Attenuated Inversion Recovery Images Improve Inter-Rater Reliability of Fluid-Attenuated Inversion Recovery Signal Changes Within Acute Diffusion-Weighted Image Lesions

Bum Joon Kim, Yong-Hwan Kim, Yeon-Jung Kim, Sung Ho Ahn, Deok Hee Lee, Sun U. Kwon, Sang Joon Kim, Jong S. Kim and Dong-Wha Kang

Stroke. 2014;45:2801-2804; originally published online July 31, 2014;
doi: 10.1161/STROKEAHA.114.006515

Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2014 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/45/9/2801

Data Supplement (unedited) at:
http://stroke.ahajournals.org/content/suppl/2014/07/31/STROKEAHA.114.006515.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/
ONLINE SUPPLEMENT

Color-coded FLAIR images improve inter-rater reliability of FLAIR signal changes within acute DWI lesions

Bum Joon Kim, MD¹, Yong-Hwan Kim, MS², Yeon Jeong Kim, MD¹, Sung Ho Ahn, MD¹, Deok Hee Lee, MD, PhD³, Sun U. Kwon, MD, PhD¹, Sang Joon Kim, MD, PhD³, Jong S. Kim, MD, PhD¹, Dong-Wha Kang, MD, PhD*¹,²

Departments of Neurology¹ and Radiology³, and Vision, Image and Learning Laboratory², Asan Institute for Life Sciences, Asan Medical Center, University of Ulsan College of Medicine, Seoul, South Korea

**Supplemental Table I** Characteristics of patients with and without benefit from color-coded FLAIR images

**Supplemental Table II** Predictive value of conventional and color-coded FLAIR images for the identification of patients within 4.5 h of symptom onset
Supplemental Methods

DWI was obtained using a single-shot spin-echo echo-planar imaging technique with the following parameters: repetition time, 3000 ms; echo time, 86 ms; matrix number, 192 × 192; voxel size, 0.651 × 0.651; 5 mm slice thickness with a 2 mm inter-slice gap. FLAIR images were acquired using a fast spin-echo sequence with the following parameters: repetition time, 9000 ms; echo time, 100 ms; inversion time, 2500 ms; matrix number, 256 × 190; voxel size, 0.8594 × 0.8594; 5 mm slice thickness with a 2 mm inter-slice gap.

Inter-rater agreement was measured in two pairs of raters. Pair A was two well-trained stroke neurologists with 5–10 years of experience in neurology. Pair B was two neurologists in residency training. The results for pair B is expressed in supplemental results. The inter-rater agreement (percent agreement) according to the degree of FLAIR signal change (obvious, subtle and negative) was presented as; (number of patients rated as the particular signal intensity concordantly by the two raters / number of patients rated as the particular signal intensity from more than one rater) and was compared between conventional and color-coded FLAIR images. A two-tailed p<0.05 was considered statistically significant. All statistical analyses were performed using SPSS for Windows (version 17.0; SPSS Inc.).

Supplemental Results

In pair B, the inter-rater agreement of DWI-FLAIR mismatch was 0.538 (0.413–0.662) with conventional FLAIR images and 0.675 (0.553–0.798) with color-coded FLAIR images. The percent agreement was significantly different between conventional and color-coded FLAIR images (69.9% vs. 83.2%, p=0.02). The proportion of patients categorized as subtle FLAIR signal change from more than one rater was higher with conventional FLAIR images than with color-coded FLAIR image (32.7% vs. 18.6%, p < 0.001).
### Supplemental Table I  Characteristics of patients with and without benefit from color-coded FLAIR images

<table>
<thead>
<tr>
<th></th>
<th>Benefit (n = 28)</th>
<th>No benefit (n = 85)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (years)</strong></td>
<td>66.9 ± 15.4</td>
<td>67.1 ± 11.6</td>
<td>0.93</td>
</tr>
<tr>
<td><strong>Male</strong></td>
<td>13 (46.4)</td>
<td>48 (56.5)</td>
<td>0.36</td>
</tr>
<tr>
<td><strong>Hypertension</strong></td>
<td>22 (78.6)</td>
<td>55 (74.7)</td>
<td>0.17</td>
</tr>
<tr>
<td><strong>Diabetes</strong></td>
<td>8 (28.6)</td>
<td>29 (34.1)</td>
<td>0.59</td>
</tr>
<tr>
<td><strong>Hyperlipidemia</strong></td>
<td>9 (32.1)</td>
<td>38 (44.7)</td>
<td>0.24</td>
</tr>
<tr>
<td><strong>Smoking</strong></td>
<td>4 (14.3)</td>
<td>15 (17.6)</td>
<td>0.68</td>
</tr>
<tr>
<td><strong>Atrial fibrillation</strong></td>
<td>10 (35.7)</td>
<td>16 (18.8)</td>
<td>0.07</td>
</tr>
<tr>
<td><strong>History of CAD</strong></td>
<td>4 (14.3)</td>
<td>5 (5.9)</td>
<td>0.15</td>
</tr>
<tr>
<td><strong>NIHSS score</strong></td>
<td>4.1 ± 3.6</td>
<td>4.0 ± 5.5</td>
<td>0.93</td>
</tr>
<tr>
<td><strong>Stroke subtypes</strong></td>
<td></td>
<td></td>
<td>0.02</td>
</tr>
<tr>
<td>Large artery atherosclerosis</td>
<td>6 (21.4)</td>
<td>35 (41.2)</td>
<td></td>
</tr>
<tr>
<td>Small vessel occlusion</td>
<td>8 (28.6)</td>
<td>19 (22.4)</td>
<td></td>
</tr>
<tr>
<td>Cardioembolism</td>
<td>13 (46.4)</td>
<td>18 (21.2)</td>
<td></td>
</tr>
<tr>
<td>Undetermined or other</td>
<td>1 (3.6)</td>
<td>13 (15.3)</td>
<td></td>
</tr>
<tr>
<td><strong>Symptom-to-MRI time (h)</strong></td>
<td>8.3 ± 7.5</td>
<td>6.7 ± 7.0</td>
<td>0.30</td>
</tr>
<tr>
<td><strong>Use of tissue plasminogen activator</strong></td>
<td></td>
<td></td>
<td>0.41</td>
</tr>
<tr>
<td>Cortical</td>
<td>12 (42.9)</td>
<td>15 (17.6)</td>
<td></td>
</tr>
<tr>
<td>Subcortical</td>
<td>12 (42.9)</td>
<td>43 (50.6)</td>
<td></td>
</tr>
<tr>
<td>Cortico-subcortical</td>
<td>4 (14.3)</td>
<td>23 (27.1)</td>
<td></td>
</tr>
<tr>
<td>Territorial</td>
<td>4 (4.7)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
</tbody>
</table>

Results are expressed as number (% column) or mean ± SD.

FLAIR: fluid-attenuated inversion recovery; CAD: coronary artery disease; NIHSS: National Institutes of Health Stroke Scale; MRI: magnetic resonance imaging.
**Supplemental Table II** Predictive value of conventional and color-coded FLAIR images for the identification of patients within 4.5 h of symptom onset

<table>
<thead>
<tr>
<th></th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conventional FLAIR (rater 1)</td>
<td>47.5%</td>
<td>90.4%</td>
<td>85.3%</td>
<td>59.5%</td>
</tr>
<tr>
<td>Conventional FLAIR (rater 2)</td>
<td>37.7%</td>
<td>82.7%</td>
<td>71.9%</td>
<td>53.1%</td>
</tr>
<tr>
<td>Color-coded FLAIR (rater 1)</td>
<td>36.1%</td>
<td>98.1%</td>
<td>95.7%</td>
<td>56.7%</td>
</tr>
<tr>
<td>Color-coded FLAIR (rater 2)</td>
<td>37.7%</td>
<td>90.4%</td>
<td>82.1%</td>
<td>55.3%</td>
</tr>
</tbody>
</table>

FLAIR: fluid-attenuated inversion recovery; PPV: positive predictive value; NPV: negative predictive value.

Raters 1 and 2 were raters in pair A of the main manuscript.