Reliability of MR Perfusion-Weighted and Diffusion-Weighted Imaging Mismatch Measurement Methods

M. Luby and S. Warach

AJNR Am J Neuroradiol 2007, 28 (9) 1674-1678
doi: https://doi.org/10.3174/ajnr.A0643
http://www.ajnr.org/content/28/9/1674
Reliability of MR Perfusion-Weighted and Diffusion-Weighted Imaging Mismatch Measurement Methods

**BACKGROUND AND PURPOSE:** We investigated 2 methods of measuring MR imaging perfusion-diffusion mismatch to determine whether reliability is improved by direct measurement on a single, blended map.

**MATERIALS AND METHODS:** Image software was used for measurement of lesion volumes from diffusion-weighted images (DWI) and mean transit time (MTT) calculated from perfusion-weighted (PWI) images on 64 patients with acute stroke. For the first method, the DWI and MTT lesions were measured separately. For the second method, the mismatch volume was measured directly on the blended images created from the registered DWI and MTT images.

**RESULTS:** Test-retest agreement was 100% and 97% for the separate and blended methods using mismatch cutoffs of 20% or more versus less than 20%. There were no significant differences in the mismatch statistics between the methods.

**CONCLUSIONS:** Mismatch volumes by a single reader can provide highly reliable and consistent results even when separately measuring DWI and MTT lesions. Propagation of measurement error was not demonstrated, and the methods were statistically comparable.

The MR perfusion-weighted imaging (PWI)-diffusion-weighted imaging (DWI) mismatch concept has become increasingly important for the inclusion of patients with stroke in clinical trials and for thrombolysis therapy. The Desmoteplase in Acute Ischemic Stroke and Dose Escalation of Desmoteplase for Acute Ischemic Stroke studies demonstrated that patients with a PWI-DWI mismatch who were treated with desmoteplase experienced a higher rate of reperfusion and better clinical outcome. An existing mismatch indicates the presence of hypoperfused but possibly still salvageable tissue, that is, “tissue at risk.” However, the standard measurement of mismatch used across these MR imaging studies, the difference in volume between perfusion and diffusion lesions, may contain compounded error introduced when measuring these MR imaging sequences separately. The purpose of this study was to investigate the reliability of 2 methods of measuring PWI-DWI mismatch volume to determine whether the reliability of estimating mismatch volume is improved by direct measurement on a single, blended PWI-DWI map.

**Methods and Materials**

**Patients**

This study is part of a prospective, natural history study of MR imaging in a consecutive series of tPA-treated patients at the National Institutes of Neurologic Disorders and Stroke (NINDS) and Suburban Hospital (Bethesda, Md). The institutional review boards at NINDS and Suburban Hospital approved the study. From February 2000 through January 2005, 147 patients were treated with standard intravenous tPA, of those, 81 patients had an MR imaging before tPA treatment and were the subject of the mismatch analysis. Patients were eligible for this analysis if they received an MR imaging scan (“acute” scan) followed by standard intravenous tPA within 3 hours from stroke onset. Only image time points and sequences with identifiable lesions were included to yield the current study sample of 64 patients. Six of the 81 patients were excluded because of unavailable acute MR imaging sequences. Eleven of the 81 patients were excluded for not having detectable lesions on DWI and PWI.

**Imaging Sequences**

MR imaging sequences were performed on a 1.5T clinical scanner (TwinSpeed; General Electric, Waukesha, Wis).

**DWI**

In this study, the DWI spin-echo-planar sequence included 20 contiguous axial oblique sections, with $b = 0$ and $b = 1000 \text{ s/mm}^2$, isotropically weighted in 3 gradient directions (R, S, and P), with 1 signal intensity average, using TR/TE at 6000 ms/72 ms, acquisition matrix of 128 $\times$ 128, 7-mm section thickness, 0-mm section gap, and 24-cm FOV.

**PWI**

In this study, the PWI gradient-echo-planar sequence included 20 contiguous axial oblique sections, with single-dose gadolinium injection of 0.1 mmol/kg via power injector, using 25 phase measurements (2 seconds per phase measurement), TR/TE at 2000 ms/45 ms, acquisition matrix of 64 $\times$ 64, 7-mm section thickness, 0-mm section gap, and 24 cm FOV.

**Mean Transit Time**

Mean transit time (MTT) maps were calculated from PWI using time concentration curves (TCC) in this study as the first moment of the...
were used to address the key question of whether one set of measurements is sufficiently representative or if 2 sets of measurements are required for providing the most accurate results. The 95% confidence limits are proposed as the repeatability coefficients of one type of measurement for another, that is, one set of measurements is sufficient rather than requiring 2 sets in this particular study.8

Results
Mismatch was measurable in 64 patients at the acute time point. Only measurements where a lesion was seen (ie, non-0 volume) on the DWI or MTT sequences for at least one read are reported in the tables and all of the figures.

Method 1 (Separate Lesion Volume Measurement)
Calculation: Mismatch (MTT − DWI). For Method 1, the calculated mismatch volume (MTT − DWI) statistics are contained in Table 1. All of the Spearman correlations were significant at the 0.01 level (2-tailed) as presented in Table 1.

Fig 2A displays the Bland-Altman plot for the Mismatch Method 1 for the acute mismatch volumes. The upper and lower limits, as shown thick black solid lines, were calculated for each plot to represent ±2 SD from the mean. A total of 97% of data points were within these boundary limits for the mismatch data.

For Method 1, the calculated percentage of mismatch (mismatch/MTT) statistics is contained in Table 2. All of the Spearman correlations were significant at the 0.01 level (2-tailed) as presented in Table 2.

Fig 2C displays the Bland-Altman plot for the Mismatch Method 1 for the acute mismatch percentages. The upper and lower limits, shown as thick black solid lines, were calculated for each plot to represent ±2 SD from the mean. A total of 98% of data points were within these boundary limits for the mismatch percentage data.

Method 2 (Blended Difference Volume Measurement)
For Method 2, the blended mismatch volume statistics are contained in Table 1. All of the Spearman correlations were significant at the 0.01 level (2-tailed) as presented in Table 1.

Fig 2B displays the Bland-Altman plot for the Mismatch Method 2 for the acute mismatch volumes. A total of 94% of acute data points were within the boundary limits for the mismatch data.

For Method 2, the calculated percentage of mismatch (mismatch/MTT) statistics is contained in Table 2. All of the Spearman correlations were significant at the 0.01 level (2-tailed) as presented in Table 2.

Fig 2D displays the Bland-Altman plot for the Mismatch Method 2 for the acute mismatch percentages. The upper and lower limits, shown as thick black solid lines, were calculated for each plot to represent ±2 SD from the mean. A total of 94% of data points were within these boundary limits for the mismatch percentage data.

Discussion
The presence of mismatch is increasingly important in the identification of target patients for thrombolysis and enrollment into clinical trials. However, the reliability of mismatch volume and percentage of measurements has not been extensively investigated within an acute stroke onset
window. Because physiologic change, measurement technique, and associated error may potentially change or obscure the amount and presence of mismatch, it is important to understand the available measurement techniques, as well as sources and magnitude of measurement error. One limitation of the current study is that all of the mismatch data were acquired 3 hours from stroke and from patients who subsequently received intravenous tPA. This is the main explanation for the large number of patients, 85%–89%, demonstrating at least a 20% mismatch, based on quantitative methods in this study, and probably confirmed before the time of thrombolysis. Therefore, indirectly this study confirms the qualitative interpretation for at least 20% mismatch performed as part of the clinical evaluation for thrombolysis consideration. This is consistent with other studies using the 20% criteria. However this study does not provide a quantitative method rapid enough to replace the current qualitative assessment used for clinical trial enrollment.

In the present study, technical variables that could affect lesion volume measurement, such as the MR imaging scanner type, pulse sequence parameters, image processing, and anal-

<table>
<thead>
<tr>
<th>Mismatch Volume</th>
<th>Read 1 Average Acute (n = 64)</th>
<th>Read 2 Average Acute (n = 64)</th>
<th>Spearman Correlation Coefficient</th>
<th>Absolute Volume Difference, Mean ± SD, Median</th>
<th>% Deviation, Mean ± SD, Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method 1</td>
<td>90.45</td>
<td>84.65</td>
<td>0.961</td>
<td>13.83 ± 23.85, 4.85</td>
<td>24.68 ± 90.35, 7.67</td>
</tr>
<tr>
<td>Method 2</td>
<td>83.15</td>
<td>82.16</td>
<td>0.975</td>
<td>13.23 ± 13.76, 7.57</td>
<td>37.22 ± 14.49, 14.49</td>
</tr>
</tbody>
</table>
ysis software, were held constant for each method. Thus, we have characterized the reliability independent of other sources of measurement error.

The amount of mismatch seen with Methods 1 and 2 is comparable with other studies using similar stroke populations but slightly smaller compared with studies with larger stroke onset windows or because of different perfusion processing software. The quantitative results presented in this study are significantly more reliable compared with other qualitative and manually derived quantitative methods. There are some technical limitations with the blended maps presented in this study. In some instances, as shown in Fig 1F, the DWI components from the blended method are smaller, and in Fig 1H, the MTT components are larger; however, the mismatch volumes produced were not significantly different from those of the standard method.

Based on the results from the Bland-Altman analyses for both mismatch volume and percentage, a single quantitative read of mismatch is sufficient for mismatch detection. The repeatability seen for mismatch was less than 6 mL and less than 7% for Method 1 at the acute time points. There were 2 outliers readily identified by the Bland-Altman plots in the mismatch reads. The repeatability seen for mismatch was less than 1 mL and less than 2% for Method 2 at the acute time points. There were 4 outliers readily identified by the Bland-Altman plots in the mismatch reads.

Mismatch volumes by a single reader can provide highly reliable and consistent results even when separately measuring DWI and MTT lesion volumes. There is a potential trend for increased sensitivity and decreased variability in the detection of mismatch with the blended difference measurement method. Although expected to be a source of significant variability in the separate volume measurement, propagation of
measurement error was not demonstrated as a factor, and the 2 methods were statistically comparable. This study validates the current approach for centralized reading of mismatch in stroke trials, which uses the standard method, separately measuring the DWI and MTT volumes.

Acknowledgments

We thank the National Institutes of Health stroke team and patients at Suburban Hospital for their valuable assistance and cooperation.

References


Table 2: Mismatch percentage reliability and comparison

<table>
<thead>
<tr>
<th>Mismatch Volume</th>
<th>Read 1 Average % (Mismatch Volume/MTT Volume) Acute (n = 64)</th>
<th>Read 2 Average % (Mismatch Volume/MTT Volume) Acute (n = 64)</th>
<th>Spearman Correlation Coefficient (Mismatch Volume/MTT Volume)</th>
<th>Cases with 20% or More Mismatch (Mismatch Volume/MTT Volume), Read 1</th>
<th>Cases with 20% or More Mismatch (Mismatch Volume/MTT Volume), Read 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method 1</td>
<td>59.33</td>
<td>51.92</td>
<td>0.966</td>
<td>52</td>
<td>52</td>
</tr>
<tr>
<td>Method 2</td>
<td>71.11</td>
<td>73.71</td>
<td>0.645</td>
<td>55</td>
<td>57</td>
</tr>
</tbody>
</table>

Note:—MTT indicates mean transit time.