“Clinical-CT” Mismatch and the Response to Systemic Thrombolytic Therapy in Acute Ischemic Stroke

David M. Kent, MD, MS; Michael D. Hill, MD, MS; Robin Ruthazer, MPH; Shelagh Cou tts, MBChB; Andrew M. Demchuk, MD; Imanuel Dzialowski, MD; Olaf Wunderlich, MD; Rudiger von Kummer, MD

Background and Purpose—Mismatch between clinical deficits and imaging lesions in acute stroke has been proposed as a method of identifying patients who have hypoperfused but still have viable brain, and may be especially apt to respond to reperfusion therapy. We explored this hypothesis using a combined database including 4 major clinical trials of intravenous (IV) thrombolytic therapy.

Methods—To determine what the radiological correlates of a “matched” functional deficit are, we calculated the relationship between the ASPECT score of the 24-hour (follow-up) CT scan and the 24-hour National Institutes of Health Stroke Scale (NIHSS) score on the subsample with ASPECT scores performed at this time (n = 820). Based on this empirical relationship, we computed the absolute difference between the observed baseline ASPECT score and the “expected” score (ie, matched) based on baseline NIHSS for all patients (n = 2131). We tested whether patients with better than expected baseline ASPECTS were more likely to benefit from IV recombinant tissue plasminogen activation (rtPA).

Results—At 24 hours, there was a strong, linear, negative correlation between NIHSS and ASPECTS ($r^2 = 0.33$, $P < 0.0001$); on average, an increase of 10 points on NIHSS corresponded to a decrease of $\approx 3$ points on ASPECTS. At baseline, the average degree of mismatch between the observed and “expected” ASPECTS was 2.1 points (interquartile range, 1.0 to 3.4). However, multiple analyses failed to reveal a consistent relationship between the degree of clinical-CT mismatch at baseline and a patient’s likelihood of benefiting from IV rtPA.

Conclusion—Clinical-CT mismatch using ASPECT scoring does not reliably identify patients more or less likely to benefit from IV rtPA. (Stroke. 2005;36:000-000.)

Key Words: cerebrovascular accident □ computed tomography □ emergency treatment □ neuroimaging □ stroke, acute □ stroke, ischemic □ thrombolytic therapy

Intravenous thrombolytic therapy has been found to improve 90-day functional outcome when administered within 3 hours of symptom onset in the National Institute for Neurological Disorders and Stroke (NINDS) Trial, and a recent combined analysis of all major intravenous thrombolytic therapy trials suggests that its benefits may extend beyond this time window. However, its use is associated with a 4.8% risk of major intracranial hemorrhage, which has a mortality rate of 60% at 3 months. Thus, the risks and benefits of the therapy may be finely balanced, especially at later time intervals. There is therefore great interest in methods of patient selection, such that those who are unlikely to benefit from therapy would not be unnecessarily exposed to these risks, and those likely to benefit are not excluded from therapy simply because they present beyond a rigid time window. In particular, much attention has focused on imaging techniques. Some have speculated that the presence of a large degree of mismatch between the perfusion-weighted imaging lesion and the smaller diffusion-weighted imaging (DWI) lesion on magnetic resonance imaging may identify these patients especially apt to benefit from thrombolytic therapy. Further, noting that symptom severity in acute stroke is more highly correlated with perfusion-weighted imaging volume than DWI volume, it has also been suggested that so-called clinical DWI mismatch (an incongruously low baseline DWI lesion volume compared with the baseline National Institutes of Health Stroke Scale [NIHSS] score) might similarly identify “high-benefit” patients. However, conventional CT scanning remains the modality most typically used for acute stroke patients. Thus, we...
hypothesized that mismatch between a patient’s clinical stroke severity (an indicator of the area of compromised perfusion) and the CT scan lesion seen at presentation (an indicator of the irreversibly infarcted area) might similarly identify such patients.

Materials and Methods

Overview and Data

Using a combined database including the NINDS Recombinant Tissue Plasminogen Activator (rtPA) Trial,1 Alteplase Thrombolysis for Acute Noninterventional Therapy in Ischemic Stroke (ATLANTIS) A16 and B17 Trials, and the Second European-Australasian Acute Stroke Study (ECASS II),18 we examined whether the presence of mismatch between the CT scan findings observed at baseline and those expected based on stroke severity influenced the effect of systemic thrombolytic therapy on outcomes in patients randomized to rtPA versus placebo. The included trials are described in detail in their original reports and the important differences between the databases are reviewed in a recent combined analysis.1,2,16–18

ASPECTS Scoring

The extent of the CT scan lesion was determined by the Alberta Stroke Program Early CT Score (ASPECTS). ASPECTS is a weighted volumetric scale used to score the degree of ischemic change present on an acute stroke patient’s CT scan within the first 24 hours from symptom onset. For ASPECTS, “ischemic changes” refers to hypo-attenuating brain tissue because of ischemic edema and brain tissue swelling. Because it is weighted, it is operationally a combined lesion volume/lesion location score overcoming the known poor correlation between lesion volume and clinical outcome.20,21 The score applies to the middle cerebral artery territory only and ranges from 0 to 10, with 10 implying no evidence of ischemic change and 0 implying a complete MCA territory infarct.1 The content validity and reliability of this scale has been demonstrated.22–24

Analysis

To operationally define mismatch, we first sought to empirically define what constitutes an appropriately “matched” functional deficit on CT scan. To do so, we required a sample of patients who were unlikely to show significant mismatch. We used the subsample of patients with follow-up ASPECTS at 24 hours (n = 820), comprising all the ATLANTIS Trial patients as well as 200 randomly selected patients from ECASS II (evenly divided among treatment and control groups).

We used linear regression to derive an equation describing the relationship between 24-hour NIHSS score and 24-hour ASPECTS using nonparametric splines to explore for potential nonlinearities. Then, using this derived equation and each patient’s baseline NIHSS, we calculated an “expected” ASPECTS for each patient. This “expected” ASPECTS was taken to represent a “matched” radiologic lesion for the stroke severity on presentation. Our measure of the degree of clinical-CT mismatch was simply the absolute difference between each patient’s “expected” ASPECTS and the actual ASPECTS at baseline. We called this difference the “ASPECTS residual.”

We hypothesized that those with a higher (positive) residual (ie, a better than expected CT scan) would be more likely to benefit from rtPA than those with lower or negative ASPECTS residuals (those with CT scan lesions more appropriate to the clinical deficit). We tested this hypothesis using logistic regression by examining whether the interaction between ASPECTS residual and rtPA treatment was a significant predictor of outcome. For this analysis, the primary outcome was a normal or near-normal functional outcome (defined by a modified Rankin score [mRS] ≤1). In secondary analyses, we tested the same hypothesis using the outcome of no major disability (mRS ≤2). We also tested the same hypothesis using generalized estimating equations to assess the benefit of rtPA on the global outcome,25 combining normal/near-normal outcomes on the mRS, the NIHSS, and the Barthel Index.

Because the use of clinical-CT mismatch is most germane in patients with clinically moderate to severe stroke, we tested the hypothesis on this subgroup of patients using predefined clinical and radiological thresholds. In this analysis, we examined whether, among patients with severe clinical strokes (NIHSS 7–15), those without severe radiologic findings (ASPECTS >7) were more likely to benefit than those with severe findings (ASPECTS ≤7). As a secondary analysis, we repeated this analysis on patients with moderate to severe stroke (NIHSS ≥8).

Results

The ASPECT score of the 24-hour CT scan was negatively and linearly correlated to the 24-hour NIHSS from an NIHSS of 0 to 20, beyond which the average ASPECTS remains approximately constant regardless of severity. Linear regression with NIHSS truncated at 20 demonstrated that every 10-point increase in NIHSS was associated with a 3-point decrease in ASPECTS, as shown in Figure 1.

ASPECTS residuals were then calculated on each patient as shown in Figure 2. The average ASPECTS residual was 2.1 (interquartile range [IQR], 1.0 to 3.4). This positive residual indicates baseline CT scans are, on average, somewhat better than would be clinically expected based on the relationship between the 24-hour stroke severity and CT scans. This is consistent with aggregate clinical improvement and/or radiologic worsening over the first 24 hours and is demonstrated by the relative density of patients in the “northeast” quadrant of Figure 2 compared with Figure 1.

In logistic regression modeling, the magnitude of the ASPECTS residual was found to be predictive of outcome; those with a higher residual were less likely to have a normal
or near normal outcome (odds ratio [OR]=0.91 for each 1-point increase in the residual [95% CI, 0.87 to 0.94; \(P\leq0.0001\)) because, on average, the ASPECTS residual increases with higher baseline NIHSS scores (as seen in Figure 2). Transforming the ASPECTS residual to a quadratic considerably improved prediction of outcome, increasing the area under the receiver-operator curve for the model from 0.58 to 0.62 (\(P\leq0.0001\)).

**Primary Analyses**

In an unadjusted logistic regression model, the ASPECTS residual did not appear to modify the effect of rtPA at all (probability value for the interaction of ASPECTS residual with rtPA treatment \(P=0.72\)). Adjusting for baseline NIHSS, age, glucose, symptom onset to treatment time [OTT], and an interaction term of NIHSS by age, and OTT by rtPA treatment did not appreciably change this result (\(P=0.73\)).

This indicates that patients with better-than-expected CT scans were no more likely to benefit from rtPA than others. Similarly, when we analyzed the subgroup of patients with severe strokes (NIHSS >15), we found that those with severe radiological findings (ASPECTS >7) were just as likely to benefit from rtPA treatment as those without such findings (ASPECTS >7) (Table 1). The adjusted probability value for the interaction between CT scan severity and treatment effect was 0.99.

**Secondary Analyses**

To further explore the possibility that the degree of clinical-CT mismatch as reflected in the ASPECTS residual influenced the treatment effect of rtPA, multiple secondary analyses tested whether ASPECTS residual modified the treatment effect of rtPA when it was expressed in quadratic form, when the global outcome (incorporating mRS, Barthel, and NIHSS) was used, or when the outcome mRS \(\leq2\) was used. None of these analyses revealed any significant effect of mismatch on the likelihood of benefit with rtPA. These results were also not changed when a term to control for any possible interaction between rtPA treatment and NIHSS was forced in to the model to control for the correlation between higher NIHSS and higher ASPECTS residual. The range of \(P\) values for the ASPECTS residual by rtPA treatment interaction term for these 10 secondary analyses were 0.32 to 0.96. Moreover, there was no differential treatment effect when outcome in patients with moderate to severe strokes (NIHSS \(\geq8\)) with ASPECTS \(\leq7\) were compared with those with ASPECTS >7.

To search further for a subgroup of patients with a high degree of clinical-CT mismatch that might be more likely than others to benefit, we categorized patients into groups based on the magnitude of their ASPECTS residuals. These results are shown in Table 2. Regardless of what threshold

---

**Table 1. Probability of a Normal or Near-Normal Outcome in Patients With Severe Strokes (NIHSS>15) and “Matched” vs “Mismatched” CT-Scans**

<table>
<thead>
<tr>
<th>Matched (ASPECTS (\leq7))</th>
<th>Mismatched (ASPECTS &gt;7)</th>
<th>Interaction P Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Placebo rtPA Placebo rtPA Placebo rtPA</td>
<td>Placebo rtPA Placebo rtPA Placebo rtPA</td>
<td>Placebo rtPA Placebo rtPA</td>
</tr>
<tr>
<td>mRS Unadjusted (Adjusted) Global Outcome Unadjusted (Adjusted)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8% (12/151) 16% (26/166) 11% (20/187) 19% (31/166)</td>
<td>0.81 (0.99) 0.99 (0.99)</td>
<td></td>
</tr>
<tr>
<td>(P=0.04) (P=0.04)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*These p-values test the significance of the difference in the treatment-effect of rtPA in the Matched vs the Mismatched patients, using both the mRS and the Global Outcome.
was used to separate patients with a small degree of mismatch from those with a large degree of mismatch, both the low ASPECT residual subgroup and the high ASPECT residual subgroup had a trend toward benefit with rtPA. Only when the ASPECT residual was dichotomized at 4 points was there a statistically significant difference in the degree of benefit with rtPA therapy between the 2 groups. Using the mRS as the outcome scale, the 18% of patients with a high ASPECT residual (\( \geq 4 \)) had a relative risk for a good outcome with treatment of 1.70 compared with 1.14 for those with a lower ASPECT residual (<4) (probability value for the interaction ASPECT residual and treatment = 0.04). However, this result was no longer significant when the global outcome scale was used \((P=0.08)\).

**Discussion**

In this study, among >2000 randomized patients, we failed to detect any consistent difference in response to rtPA based on the degree of mismatch between baseline stroke severity by NIHSS and baseline CT scan findings as measured by ASPECT scores. Specifically, patients with CT scans that were more normal than expected did not derive more benefit from rtPA compared with patients with CT scans showing ischemic changes more consistent with their clinical deficits (or worse). Because it is generally believed that the clinical deficit in the hyperacute phase reflects the size of both the infarcted and penumbral zones\(^{13,14}\) and that early CT scan findings reflect only the infarcted tissue,\(^{15}\) this finding is somewhat surprising because the effectiveness of thrombolytic therapy presumably depends on the presence of viable penumbral tissue (theoretically represented by the mismatch) that can be salvaged when blood flow is restored.

However, a closer examination of the assumptions on which the “clinical-CT mismatch hypothesis” depend yields clues as to why such a treatment effect interaction might be difficult to detect.

First, the hyperacute clinical deficit may not be related just to the zone of infarction and penumbra. For example, it may also reflect “stunned” regions or diaschisis, which ultimately recover regardless of reperfusion. Conversely, it is also possible that at least some penumbral tissue may retain some function during the hyperacute phase, causing the degree of clinical-CT mismatch to underestimate the penumbra.

Second, “ischemic changes” as defined by the ASPECTS may not well-reflect irreversibly infarcted tissue. Although areas involved at baseline are usually involved on follow-up,\(^{15}\) reversal of early ischemic changes has been observed under well-controlled conditions.\(^{21}\) This is especially the case because ASPECTS takes into account both hypo-attenuating and nonhypo-attenuating (ie, swelling) brain tissue changes. The nonhypo-attenuating changes (sulcal effacement, ventricular compression, cortical thickening) may represent compensatory arterial dilatation that are reversible by reperfusion.\(^{15}\) Conversely, it is also the case that irreversibly infarcted regions may at times be difficult to detect on baseline scans because of artifact, volume averaging, or adjacent encephalomalacia from previous infarction.

Third, the relationship between functional and anatomic deficit may not be well-specified. The central hypothesis assumes that the ASPECTS residuals shown in Figure 2 closely reflect the amount of salvageable penumbral tissue. However, other factors clearly contribute to the degree of scatter, which is seen even at 24 hours (Figure 1) when the zone of salvageable penumbra is presumably quite limited. Although highly correlated statistically, a patient’s NIHSS is not determined by ASPECTS alone but by the specific anatomic sites involved (including the side of the lesion\(^{26,27}\)), and by interindividual variation.

Finally, the reliability of the treatment effect itself also affects the likelihood of detecting a treatment effect interaction. Systemic thrombolytic therapy causes early reperfusion in \( \approx 50\% \) of patients, depending on the site of occlusion,\(^{28,29}\) and patients who are not treated will on occasion spontaneously recanalize. This considerably diminishes the effective sample size in which to uncover a treatment–effect interaction.

Taken together, even if the reasoning that forms the basis of the mismatch hypothesis is sound, these caveats would cumulatively and considerably weaken the link between clinical-CT mismatch and the likelihood of benefit from rtPA. The trends in Table 2 suggest that there may be some subtle effects in the expected direction, particularly for patients with large clinical deficits and normal or near-normal CT scans. Because the variance is high, much larger numbers or substantially more homogenous patient populations may be needed to tease out the presence of an interaction, if present. For example, the demonstration of an interaction effect with

### Table 2. 90-day Outcomes in Patients With High and Low ASPECTS-Residuals (ie dichotomized outcomes)

<table>
<thead>
<tr>
<th>ASPECTS-Residual Threshold</th>
<th>Percent of Patients With Mismatch, ( \geq ) Threshold (no.)</th>
<th>Matched (ASPECTS-residual &lt;threshold)</th>
<th>Mismatched (ASPECTS-residual ( \geq ) threshold)</th>
<th>Interaction P Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Placebo rTPA</td>
<td>Placebo rTPA</td>
<td>Placebo rTPA mRS</td>
<td>Global Outcome</td>
</tr>
<tr>
<td>2</td>
<td>54% (1111)</td>
<td>41%</td>
<td>46%</td>
<td>28%</td>
</tr>
<tr>
<td>3</td>
<td>35% (715)</td>
<td>42%</td>
<td>46%</td>
<td>21%</td>
</tr>
<tr>
<td>4</td>
<td>18% (366)</td>
<td>39%</td>
<td>43%</td>
<td>13%</td>
</tr>
<tr>
<td>5</td>
<td>10% (205)</td>
<td>37%</td>
<td>42%</td>
<td>10%</td>
</tr>
<tr>
<td>6</td>
<td>4% (85)</td>
<td>35%</td>
<td>41%</td>
<td>10%</td>
</tr>
</tbody>
</table>

*These \( P \) values test the significance of the difference in the treatment-effect of rTPA in the Matched vs the Mismatched patients, using both the mRS and the Global Outcome.
CT ASPECTS \(> 7\) in the PROACT-II study of intra-arterial prourokinase for MCA used a very pure population of stroke patients with angiographically proven occlusion and a high recanalization rate with therapy.\(^2\) However, we can conclude from our study that this type of mismatch is too unreliable to use clinically in individual patients, unless somehow embedded within decision support that takes into account a broad array of patient characteristics, each with a subtle effect on a patient's risk–benefit trade-off.

One important limitation of our study is that whereas patients with a wide range of ASPECT scores and ASPECTS residuals were included, \(<4\%\) of patients in this study had large-volume hypodensities on CT scan (ie, involvement of more than one-third of the MCA territory). Therefore, we lacked sufficient statistical power to test whether such scans themselves predict a low likelihood of benefit from rtPA therapy.

Many of the assumptions that underlie the clinical-CT mismatch hypothesis also underlie other imaging modalities promising for patient selection. Our findings offer caution that methods supported by sound physiological reasoning will not necessarily yield tests that are clinically useful in discriminating between patients likely or unlikely to respond to rtPA. Demonstrating a true treatment–effect interaction with patients randomized to rtPA and placebo may prove more difficult than preliminary "proof-of-concept" studies suggest.

Acknowledgments

D.M.K. is a Pfizer Scholar in Clinical Epidemiology and is supported by a Career Development Award from National Institute for Neurological Disorders and Stroke (K23 NS44929-01). M.D.H. was funded by the Canadian Institutes for Health Research and the Heart & Stroke Foundation of Alberta, Northwest Territories, and Nunavut. The reading of the ECASS II CT-scans was supported by a grant to D.M.K. from the Tufts-New England Medical Center Research Fund. S.C. was supported by Heart and Stroke Foundation Canada. The authors thank Drs David Thaler and Judith Hinchey for reviewing their paper for clarity.

References

"Clinical-CT" Mismatch and the Response to Systemic Thrombolytic Therapy in Acute Ischemic Stroke

David M. Kent, Michael D. Hill, Robin Ruthazer, Shelagh Coutts, Andrew M. Demchuk, Imanuel Dzialowski, Olaf Wunderlich and Rudiger von Kummer

Stroke. published online July 7, 2005;

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/early/2005/07/07/01.STR.0000173397.31469.4b.citation