Comparison of 10 Perfusion MRI Parameters in 97 Sub-6-Hour Stroke Patients Using Voxel-Based Receiver Operating Characteristics Analysis

Søren Christensen, MSc; Kim Mouridsen, PhD; Ona Wu, PhD; Niels Hjort, MD, PhD; Henrik Karstoft, PhD; Götz Thomalla, MD; Joachim Röther, MD; Jens Fiehler, MD, PhD; Thomas Kucinski, MD, PhD; Leif Østergaard, MD, PhD

Background and Purpose—Perfusion-weighted imaging can predict infarct growth in acute stroke and potentially be used to select patients with tissue at risk for reperfusion therapies. However, the lack of consensus and evidence on how to best create PWI maps that reflect tissue at risk challenges comparisons of results and acute decision-making in trials. Deconvolution using an arterial input function has been hypothesized to generate maps of a more quantitative nature and with better prognostic value than simpler summary measures such as time-to-peak or the first moment of the concentration time curve. We sought to compare 10 different perfusion parameters by their ability to predict tissue infarction in acute ischemic stroke.

Methods—In a retrospective analysis of 97 patients with acute stroke studied within 6 hours from symptom onset, we used receiver operating characteristics in a voxel-based analysis to compare 10 perfusion parameters: time-to-peak, first moment, cerebral blood volume and flow, and 6 variants of time to peak of the residue function and mean transit time maps. Subanalysis assessed the effect of reperfusion on outcome prediction.

Results—The most predictive maps were the summary measures first moment and time-to-peak. First moment was significantly more predictive than time to peak of the residue function and local arterial input function-based methods ($P<0.05$), but not significantly better than conventional mean transit time maps.

Conclusion—Results indicated that if a single map type was to be used to predict infarction, first moment maps performed at least as well as deconvolved measures. Deconvolution decouples delay from tissue perfusion; we speculate this negatively impacts infarct prediction. (Stroke. 2009;40:2055-2061.)

Key Words: diffusion magnetic resonance imaging ■ magnetic resonance angiography ■ perfusion weighted MRI ■ stroke

Together with diffusion-weighted imaging (DWI), perfusion-weighted imaging (PWI) has shown promise as a means to select patients with potential benefit of thrombolytic treatment.1 It has been hypothesized that brain parenchyma with a so-called PWI–DWI mismatch largely corresponds to tissue at risk of infarction in the absence of timely reperfusion.2–6 The apparent simplicity of this model is challenged by observations that DWI lesions may be reversible and that part of the perfusion alteration is benign oligemia and not penumbral tissue.7,8

The challenge of delineating salvageable tissue in acute stroke has led to study of various perfusion metrics and their corresponding thresholds in terms of predicting subsequent ischemic damage.4,9 It is increasingly clear, however, that the specific choice of perfusion metric critically affects assessments of tissue at risk.10,11

Several studies have performed side-by-side comparisons of various PWI postprocessing methods using a variety of perfusion maps, normalization methods, and comparison schemes. These studies have yielded conflicting results, some favoring the use of semiquantitative deconvolved measures and some arguing that processing makes little to no difference or that PWI has little predictive value at all.3,10–14

In this study, we compared 10 different state-of-the-art perfusion metrics using voxel-based receiver operating characteristic (ROC) analysis with the aim to identify the parameter that gives the best prediction of subsequent tissue infarction. We hypothesized that deconvolution-based maps...
would outperform nondeconvolution-based maps in predicting tissue that infarcts.

**Methods**

**Patients**

MRI data were retrieved from a database of prospectively collected consecutive patients with acute ischemic stroke that were studied by MRI with PWI and DWI imaging within 6 hours of symptom onset and follow-up DWI and T2 imaging at least 5 days later. Patients with parenchymal hematomas (class 1 and 2) or receiving intra-arterial tissue plasminogen activator (tPA) were excluded.

All patients scanned within 3 hours of symptom onset were treated with intravenous thrombolysis according to the European Stroke Initiative Recommendations for Stroke Management. Within 3 to 6 hours, thrombolysis was performed as an individual decision based on MRI findings. Informed consent was obtained from all subjects or next of kin.

A subset of these patients had a subacute MRI on Day 1 allowing for assessment of tissue reperfusion based on MR angiography and PWI using the modified Thrombolysis in Myocardial Infarction (TIMI) criteria for perfusion and vessel status. ROC analysis was performed by measuring the average sensitivity and specificity for final infarction across patients as a function of successively increasing thresholds.

**MRI Protocols**

MRI studies were performed on a 1.5-T clinical whole-body scanner (Magnetom Symphony/Sonata Siemens) using a standard head coil. Part of this patient material has been used in previous studies in which the MRI protocol has been described in detail. MRI and CT studies.

**PWI Postprocessing**

We produced 2 summary/composite map types: first moment (FM) and TTP, both including the bolus arrival time determined by gamma variate fitting. Five deconvolution-based maps using a contralateral arterial input function (AIF): mean transmit time (MTT) and time to peak of the residue function (Tmax) by both standard (sMTT, sTmax) and oscillation index regularized block circulant (oMTT, oTmax, and oCBF [cerebral blood flow]) singular value decomposition (SVD).

Two MTT maps were generated by implementation of local AIF methods that are hypothesized to provide superior estimates: local MTT method 1 based on a published method and local MTT method 2 based on an algorithm approximating local arterial territories. Finally, a cerebral blood volume (CBV) map was created.

**ROC Analysis**

ROC analysis is a well-established and comprehensive method to evaluate the efficacy of a continuous parameter in predicting a binary outcome and can be viewed as an analysis testing the predictive performance of a PWI map across all possible thresholds.

We used ROC analysis to measure the predictive value of a given PWI map relative to the final lesion on a pixel-by-pixel basis and quantified this by a means of the area under the ROC curve, the AUC. The AUC is the probability that infarcting and noninfarcting voxels are correctly ranked relative to each other (eg, the probability that an infarcting voxel has a higher MTT value than a noninfarcting voxel). Consequently, the AUC varies from 0.5 for a map with no predictive value to 1 for a map in which a single threshold value can separate infarct from salvaged tissue completely. ROC analysis has previously been applied to voxel-based analysis in other acute stroke MRI and CT studies.

The ROC analysis was performed by measuring the average sensitivity and specificity for final infarction across patients as a function of successively increasing thresholds for each map type. Subanalysis included stratification by reperfusion status. A repeat analysis was made using normalized perfusion maps, in which all temporal parameters were normalized by subtraction of the mean contralateral value and CBF and CBV were expressed as a ratio to the contralateral mean value.

**Regions of Interest used in the ROC Analysis**

Figure 1 illustrates the regions of interest (ROIs) used in the ROC analysis. Follow-up lesions were manually delineated on follow-up T2 (b=0 image) and DWI images and the union of the 2 masks used as the follow-up lesion ROI. All ROIs were transformed and resampled to the acute PWI slice locations after coregistration of acute DWI and follow-up images to acute PWI images using MINC software developed at the Montreal Neurological Institute. All coregistrations were visually inspected for proper alignment.
Follow-up ROIs were mirrored to the contralateral hemisphere and used as contralateral, homologous tissue ROIs in the analysis. DWI and PWI brain masks were created by manually applying a threshold to the DWI and raw baseline PWI average image so as to exclude noncerebral tissue. A single brain mask was then created using the intersection of the 2. To restrict ROC analysis to well-defined tissue regions, the analysis was limited to the ipsilateral hemisphere and the homologous contralateral tissue ROI described here. The ipsilateral mask was further restricted by manual removal of the extremes of the temporal poles as well as the cerebellum (due to distortions in these regions and venous signals from the transverse sinus mimicking long tissue transit times). For normalization purposes, an ROI was defined covering the whole contralateral hemisphere of the brain mask.

Statistical Analysis

Confidence intervals on the ROC curves and on the AUC estimates were generated by bootstrapping. The remainder of the tests is specified in the text along with the probability value.

Results

Baseline, Treatment, Reperfusion, and Infarct Growth Characteristics

We retrieved 106 patients, imaged between January 2000 and December 2004, fulfilling the acute and follow-up MRI time criteria described. Of these patients, 9 were excluded due to pre-existing ischemic lesions on T2-weighted imaging (2 patients), PWI of insufficient quality (one patient), and parenchymal hematomas class 1 and 2 (6 patients).

Baseline characteristics were as follows [median (interquartile range)]: age 62 years (51–71 years), acute National Institutes of Health Stroke Scale 13 (8–16), time from the beginning of stroke symptoms to scan 2 hours 35 minutes (1 hour 54 minutes to 3 hours 11 minutes), median time to follow-up imaging 7.0 days (6.7 to 7.9 days), and 34% were female. Seventy-nine patients (81%) received intravenous tPA.

Day 1 imaging, and thereby reperfusion assessment, was available in 70 patients (72%) of which 59 received tPA and 11 conservative treatment. By the reperfusion dichotomy defined previously, 37 patients did not reperfuse.

Final infarct volumes were larger among patients without reperfusion (65 mL versus 32 mL, $P < 0.005$, Wilcoxon rank sum test). The same was evident for relative infarct expansion (final volume/acute DWI volume) with a growth ratio of 2.1 in patients without reperfusion versus 1.2 in patients with reperfusion ($P = 0.002$, Wilcoxon rank sum test).

Occlusion types were assessed in 94 patients: middle cerebral artery (MCA) branch occlusion (31%), MCA trifurcation (13%), MCA trunk (20%), internal carotid artery/MCA (17%), carotid T occlusion (14%), MCA/anterior cerebral artery (3%), internal carotid artery (1%), and finally no visible occlusion in 1%.

ROC Analysis

Figure 2 shows the results from the ROC analysis for each map type and the effect of normalization to the contralateral hemisphere. Normalizing to the contralateral whole hemisphere performed best in 10 of 10 map types ($P < 0.002$, sign test) and the remaining results are limited to the normalized maps.

Figure 3 shows the global AUCs for each map type in all patients, nonreperfusers and reperfusers. AUCs were significantly higher in nonreperfusers than in reperfusers for all map types as seen by nonoverlapping 95% CIs.28

Focusing on nonreperfusers, FM had the highest AUC (0.82) followed by TTP (0.81). However, FM CIs show overlap with both oMTT and sMTT but not with Tmax, the local AIF methods, CBF, and CBV, which indicates statistical significance of these differences with $P < 0.05$.28

All deconvolved parameter maps showed overlapping CIs between them, whereas CBV performed significantly worse than any other metric with an AUC of 0.61. The Table lists AUC results from all subgroups.

Based on this analysis, we considered only nonreperfusing patients using normalization and selected 3 map types that were representative of the highest (FM), midrange (MTT by circular SVD), and lower AUC values (Tmax by circular SVD). CBV was not considered due to the low performance.

Figure 4 shows the ROC curves for these maps with the sensitivities and specificities at the so-called “optimal operating point” where the compromise between sensitivity and specificity are balanced (equal weights were attributed to specificity and sensitivity).29
A FM prolongation of 3.5 seconds predicted follow-up infarction with a sensitivity of 75% and a specificity of 78%.

In comparison, MTT that was the best performing deconvolution approach had an optimal operating point at 1.78 seconds MTT prolongation with a sensitivity of 72% and a specificity of 75%. As shown by the 95% CIs, there was substantial overlap between the FM and MTT performance. Tmax did, however, perform significantly worse than FM as seen by nonoverlapping CIs.28

Discussion

The highest predictive performance was found in nondeconvolved measures, FM, and TTP followed by MTT and Tmax maps using a whole brain AIF. The local AIF hypothesized to be superior did not outperform the conventional methods.

These findings were surprising and contradicted our working hypothesis.

The effects of these differences on clinical mismatch classifications are appreciable. A specificity of 75% means that 25% of healthy tissue is misclassified as penumbral. Thus, low specificities lead to overestimates of tissue at risk. If, for instance, a sensitivity of 80% is required, the graph shows that these methods would yield specificities of (FM, MTT, Tmax) 72%, 64%, and 54%, respectively.

The Effect of Normalization

The performance of all PWI maps in terms of predicting outcome improved on normalization to contralateral hemispheric values. It is well known that normalization is required for TTP and FM maps to account for injection profile and arrival time differences, but these results show that deconvolved measures, despite the deconvolution procedure being a normalization technique, also benefit from residual normalization. Normalization to contralateral periventricular white matter (results not shown) performed comparable but slightly inferior to whole hemisphere normalization.

Dependence of Predictive Performance on Subsequent Reperfusion

The predictive performance was shown to be better (higher AUCs) among nonreperfusers than in reperfusers (Figure 3). This corroborates the notion that reperfusion and the variability in the timing of this event cause heterogeneity in the perfusion thresholds that best predict infarction.30,31 This heterogeneity will be reflected in lower AUCs in a cohort in which these thresholds differ in accordance with the results of this study.

This finding stresses the importance of establishing thresholds for infarction in a patient population without reperfusion when guiding treatment in subsequent patients.

Grandin et al compared PWI map types in 66 patients scanned sub-6 hours with follow-up T2 fluid-attenuated inversion recovery imaging at 24 hours using ROC analysis. Reperfusion was not assessed. They concluded that deconvolution offered no benefits over summary parameters.3

Rivers et al addressed whether PWI/DWI mismatch could predict growth in 46 patients and found no apparent prognostic value of mismatch as defined in this study.14 Although this type of study answers a clinical relevant question in a particular setting, the dichotomization of both mismatch and

<table>
<thead>
<tr>
<th>Table. AUC Results From the ROC Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Map Type</td>
</tr>
<tr>
<td>-----------------------------</td>
</tr>
<tr>
<td>FM including arrival time</td>
</tr>
<tr>
<td>TTP including arrival time</td>
</tr>
<tr>
<td>Tmax by standard SVD</td>
</tr>
<tr>
<td>Tmax by circular SVD</td>
</tr>
<tr>
<td>CBF by circular SVD</td>
</tr>
<tr>
<td>Local AIF method 2</td>
</tr>
<tr>
<td>Local AIF method 1</td>
</tr>
<tr>
<td>CBFV</td>
</tr>
<tr>
<td>MTT by standard SVD</td>
</tr>
<tr>
<td>MTT by circular SVD</td>
</tr>
</tbody>
</table>

SVD indicates singular value decomposition.

Figure 4. Global ROC curves for patients without reperfusion using normalization to the contralateral hemisphere. OOP denotes optimal operating point (see text) and is marked “O” on the ROC curve.
Infarct growth, unknown reperfusion status, and a PWI volume defined as a manual outline of maximum lesion extent may cause loss of statistical power.

Kane et al found a large variation between map types and what type of outcome that was best predicted in a sub-24-hour population of 32 patients with no control for reperfusion and visual estimation of hyperperfusion in each case.11

Butcher et al compared mismatch definitions in 40 patients and found no significant difference between methods.10 This comparison investigated a subset of the thresholds used here and compared the methods by volume correlations with final outcome lesion volumes.

Schellinger et al compared 4 PWI maps in 20 patients using their volume–National Institutes of Health Stroke Scale correlation as a comparison measure and found no significant differences between PWI methods.12

Implications for Stroke Imaging: To Deconvolve or Not to Deconvolve?
Given the theoretical relation of deconvolved perfusion maps to underlying physiological parameters, the finding that such maps achieved lower AUC values than summary parameter maps may at first seem surprising. Summary maps are, however, influenced by several phenomena such as tracer arrival delay, arterial dispersion, and actual tissue transit time and can therefore be viewed as “composite” measures. Deconvolution, on the other hand, aims at extracting tissue MTT by decoupling delay and arterial dispersion effects. Therefore, if arrival delay and arterial dispersion contain predictive information, it would be expected that a single, deconvolved map of, say, MTT performs worse than a composite map. Indeed, studies have demonstrated considerable success of Tmax/delay in predicting outcome tissue at risk and response to reperfusion indicating a predictive role of arrival delay.23

Additional analysis (results not shown) supported the notion that arrival delay and tissue MTT in conjunction provide the best prediction. Fitted arrival delay by itself was a poor predictor (AUC <0.7) and the AUC for FM and TTP without arrival delay included dropped on average 0.05. However, technical issues regarding the precision of the fitting makes it difficult to reliably distinguish tissue dispersion and delay and we did not explore any causal relationships further as a result thereof.20

In theory, patients with chronic regional hemodynamic delay will have TTP and FM abnormalities despite absence of tissue at risk and deconvolution may be a better option in these patients, but this is still an unanswered question in practice.13

We speculate that deconvolution-based maps may display their advantage over threshold prediction approaches in multiparameter predictive models in which the use of well-defined, physiological parameters (apparent diffusion coefficient, CBF, CBV, MTT, Tmax, and so on) provide risk estimates and a means to quantify their individual importance in determining infarct risk.32 Composite parameters are inherently complex in terms of disentangling the roles of separate hemodynamic effects in a single patient.

Although sophisticated multiparameter predictive models of infarct progression are being developed, our results indicate that as a single map type to predict infarction in routine imaging, FM may be the best choice provided it is normalized to the contralateral hemisphere. The practical implication of this is that a simple and user-independent postprocessing scheme performs as good as deconvolution methods if just one PWI map type is used for decision-making. The normalization to the contralateral hemisphere is easy to implement in a fully automated fashion, allowing the visualization of above-threshold values indicating tissue at risk. It should be noted that the TTP and FM maps produced here were based on a thoroughly optimized fitting procedure, likely not similar to algorithms in the literature or commercial software packages.20 The MTT maps produced by local AIF methods did not outperform conventional whole-brain AIF approaches. This may be due to additional decoupling of hemodynamic effects with potential predictive value (arterial dispersion) or by partial volumed signals in small arteries.

CBV and CBF were poor predictors of infarction. This is likely owed, at least in part, to the bimodal distribution of CBV and CBF values due to large gray and white matter differences. This makes ROC analysis a suboptimal assessment tool for these parameters and their merit cannot be well quantified using ROC analysis. The analysis does indicate that a thresholding approach to CBF and CBV maps is unlikely to yield accurate predictions.

Limitations to Our Study
We considered if bias in this retrospective data set would favor our findings. As for the infarct growth dynamics of the population, the relative growth was 2.1 (interquartile range, 1.6 to 3.7) in patients without reperfusion comparable to 2.05 (interquartile range, 1.0 to 3.5) in EPITHET indicating comparable growth dynamics with a randomized study, although this must be viewed as a rough indication because EPITHET recruited in the 3- to 6-hour window only.

Most of the patients were treated with tPA. The single most important event in an evolving acute stroke is timely reperfusion, and we argue it is more reliable to consider reperfusion rather than tPA treatment in relation to prediction of tissue fate. Reperfusion alters tissue fate and predictions; tPA is only a catalyst of this process and may or may not provoke reperfusion. As such, we do not believe the high number of tPA-treated patients introduces bias in the comparisons of the PWI methods.

Methodological sources of error include coregistration error caused by edema at Day 7 follow-up. Previous work has shown that edema is on the decline at this time point peaking at Day 3.33 There is no established optimal time point for follow-up imaging when using coregistration-based analysis and some structural changes, whether swelling or later shrinkage, will inevitably challenge an exact spatial correspondence.

In many of the patients with large MCA infarctions and visible edema, it was our impression from visual inspection of the acute-to-follow-up coregistrations using synchronized markers that when swelling was visible, it often took place by displacement into the ventricles with only minor displace-
ments on infarct borders toward nonaffected tissue. The brain masks defined on the acute imaging exclude the ventricle-displaced regions from analysis. We estimate that misregistration affects a limited number of voxels.

We did not attempt to quantify the usefulness of PWI in defining treatment response or select a particular mismatch definition in this comparative study. The true value of PWI is assessed in the hands of an experienced clinician and not by summary AUC values. The aim of this study was to objectively compare PWI postprocessing methods in a cohort of acute stroke patients using an objective ROC methodology. Disregarding the CBV and CBF maps, our visual impression of the dissections was that of qualitative similarity between them. However, the measured specificities and sensitivities indicate that the differences are large enough to cause at least some patients to change mismatch/no-mismatch classification as used in clinical trials, indicating that the actual choice of map will have an effect on patient selection.

This study provides evidence that deconvolution techniques may not be the best option for mismatch classification based on prediction of tissue at risk when using a threshold approach to estimate tissue at risk using PWI. Because the end goal of perfusion imaging is identification of a responder subgroup, future work will need to elaborate on the agreement between methods and ultimately the treatment/perfusion response in the mismatch group as identified by the different map types.

In conclusion, the composite parameters TTP and FM performed superior to local AIF and at least as well as deconvolution-based methods in terms of prediction of tissue outcome. FM provided the highest AUC values and had a sensitivity/specificity of 75%/78% for final infarction. The best performing deconvolution parameter was MTT with a sensitivity/specificity of 72%/75%.

Deconvolution decouples hemodynamic information into several parametric maps that are likely less predictive than composite maps on an individual basis but hypothesized to be more informative using predictive models. These results need validation in independent data sets with follow-up imaging and assessment of reperfusion status.

Acknowledgments
We thank the Danish National Research Foundation, the Danish Medical Research Council, Schering AG, and GlaxoSmithKline for grant support.

Disclosures
None.

References
20. Ostergaard L, Weisskoff RM, Chesler DA, Gyldensted C, Rosen BR. High resolution measurement of cerebral blood flow using intravascular
Comparison of 10 Perfusion MRI Parameters in 97 Sub-6-Hour Stroke Patients Using Voxel-Based Receiver Operating Characteristics Analysis

Søren Christensen, Kim Mouridsen, Ona Wu, Niels Hjort, Henrik Karstoft, Götz Thomalla, Joachim Röther, Jens Fiehler, Thomas Kucinski and Leif Østergaard

*Stroke*. 2009;40:2055-2061; originally published online April 9, 2009;
doi: 10.1161/STROKEAHA.108.546069

*Stroke* is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2009 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/40/6/2055