Moving Beyond a Single Perfusion Threshold to Define Penumbra
A Novel Probabilistic Mismatch Definition

Yoshinari Nagakane, MD, PhD*; Soren Christensen, PhD*; Toshiyasu Ogata, MD, PhD; Leonid Churilov, PhD; Henry Ma, MBBS; Mark W. Parsons, PhD; Patricia M. Desmond, MD, FRACP; Christopher R. Levi, FRACP; Kenneth S. Butcher, MD, PhD; Stephen M. Davis, MD, FRACP; Geoffrey A. Donnan, MD, FRACP; for the EPITHET Investigators

Background and Purpose—The mismatch lesion volumes defined by perfusion-weighted imaging (PWI) exceeding diffusion-weighted imaging (DWI) have been used as a marker of ischemic penumbral tissue. Defining the perfusion lesion by thresholding has shown promise as a practical tool; several positron emission tomography studies have indicated a more probabilistic relationship between perfusion and infarction. Here, we used a randomized controlled trial dataset of tissue-type plasminogen activator 3 to 6 hours after stroke to: (1) quantify the relationship between severity of hypoperfusion (measured by Tmax) and risk of infarction; (2) exploit this relationship to present a novel definition of mismatch based on infarct probabilities rather than dichotomies; and (3) examine the treatment response in the subgroup of patients with mismatch by the new definition.

Methods—Patients from the Echoplanar Imaging Thrombolytic Evaluation Trial (EPITHET) were included. Baseline PWI and 90-day T2-weighted imaging were coregistered. PWI lesion volumes were divided into 10 Tmax delay strata, and infarct risk was defined as the fraction of the tissue at a given Tmax strata that progressed to infarction by day 90.

Results—Sixty-two patients were studied. Infarct risk was an increasing function of Tmax for all subgroups, including the whole cohort. The probabilistic approach outperformed all Tmax thresholds, with exception of the Tmax ≥10 threshold, for which it was only favored by a trend.

Conclusions—Infarct risk and treatment effect increased with severity of perfusion abnormalities. This suggests that a severity-weighted mismatch definition may define penumbral tissue more accurately. (Stroke. 2012;43:00-00.)

Key Words: magnetic resonance imaging ■ mismatch ■ penumbra ■ perfusion-weighted imaging ■ tissue-type plasminogen activator

Ischemic penumbral tissue is at risk of infarction but has the potential to be salvaged by reperfusion and/or other strategies. With the advent of functional MRI techniques, the concept of mismatch defined by perfusion-weighted imaging (PWI) exceeding diffusion-weighted imaging (DWI) lesion volumes was postulated as a penumbral marker. Patients with PWI/DWI mismatch are likely to benefit from timely reperfusion and may be eligible for thrombolytic therapy beyond the current therapeutic time window. A recent meta-analysis showed that MRI-selected patients treated with thrombolysis beyond 3 hours had improved clinical outcomes with reperfusion/recanalization. In acute stroke, PWI is used to estimate the tissue regions that will infarct if hypoperfusion persists and reperfusion is not achieved. Positron emission tomography studies have focused on determining thresholds for the separation between benign oligemia and penumbra, the so-called penumbral thresholds. Based on these findings, several MR studies have aimed at identifying parameters and thresholds that correspond best with this penumbral threshold. In practice, the hypoperfused at-risk region is most commonly defined as a region above a certain threshold value of mean transit time or time to peak of the residue function (Tmax). It is clear that lower threshold values may include significant amounts of

Received November 3, 2011; accepted January 18, 2012.
From the National Stroke Research Institute (Y.N., T.O., L.C., H.M., G.A.D.), Florey Neuroscience Institutes, Austin Health, University of Melbourne, Australia; Department of Neurology (S.C., S.M.D.), Royal Melbourne Hospital, University of Melbourne, Australia; Department of Mathematics and Statistics (L.C.), University of Melbourne, Australia; Department of Neurology (P.M.D., C.R.L.), Hunter Medical Research Institute, John Hunter Hospital, University of Newcastle, Australia; Department of Neurology (K.S.B.), University of Alberta, Edmonton, Alberta, Canada; Department of Radiology (M.W.P.), Royal Melbourne Hospital, University of Melbourne, Australia.

*Y.N. and S.C. contributed equally to this work.

Correspondence to Geoffrey A. Donnan, Florey Neuroscience Institutes, University of Melbourne, Level 2 Alan Gilbert Building, 161 Barry Street, Carlton South Victoria 3053, Australia. E-mail gdonnan@unimelb.edu.au

© 2012 American Heart Association, Inc.

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

DOI: 10.1161/STROKEAHA.111.643932
by higher Tmax values. The Tmax maps were stratified into 9 2-second strata as well as a ≥20-second strata, as depicted in Figure 1. In the individual patient, the infarct risk for each Tmax strata was then defined as the fraction of the strata with infarction on follow-up, as determined from the coregistered day 90 infarct regions of interest. This analysis was restricted to a manually edited region of interest containing only tissue with Tmax ≥2 seconds and excluding obvious false-positive regions. Regions that were DWI-positive were included in the analysis.

We investigated the infarct risks in all patients as well as in subgroups defined by treatment versus placebo and by recanalization versus no recanalization. As per the EPITHET protocol, recanalization was defined as improvement from baseline to day 3 to 5 arterial obstruction by ≥2 points, based on an adaptation of the Thrombolysis In Myocardial Infarction grading on MRA (0=complete occlusion, 1=severe stenosis, 2=mild to moderate stenosis, and 3=normal arterial caliber).14

**Prediction of Final Lesion Volumes**

We used the median infarct risks from nonrecanalized patients as a severity-weighted predictive model for the relationship between Tmax severity and final infarct risk. In each patient, the predicted infarct volume was defined as the sum of each Tmax strata (Volume_{Tmax}) multiplied by the median infarct risks of that strata (Risk_{Tmax}):

\[
\text{Predicted infarct volume} = \text{Risk}_{Tmax 2-4} \cdot \text{Volume}_{Tmax 2-4} + \text{Risk}_{Tmax 6} \cdot \text{Volume}_{Tmax 6} + \cdots + \text{Risk}_{Tmax 18-20} \cdot \text{Volume}_{Tmax 18-20} + \text{Risk}_{Tmax >20} \cdot \text{Volume}_{Tmax >20}
\]

We used a jack-knifing technique to avoid using the same data for model estimation as for performance validation. The predicted infarct volume for each patient was estimated using a model trained on the 10 remaining patients. Cycling through the patients in this way, we estimated these coefficients of Risk_{Tmax} 11 times, each time applying the model to the patient left out of the coefficient estimation. The severity-weighted infarct volume predictions were compared with those obtained by Tmax thresholding (Tmax 2 to Tmax 14 seconds). Analysis was restricted to nonrecanalized patients.

**Severity-Weighted Mismatch Definition**

Finally, we applied the Tmax severity versus risk model to the entire EPITHET cohort of 101 patients and classified mismatch patients using this new definition. A severity-weighted mismatch was defined as predicted infarct volume/DWI >1.2 and absolute difference >10 mL. We examined whether infarct growth was attenuated by alteplase in patients with mismatch by the new definition.

**Statistical Analysis**

Statistical analyses were performed using Stata/IC 10 (StataCorp). The 2-sample Wilcoxon rank-sum test was used for continuous variables, and Fisher exact test was used for categorical variables between patients with alteplase and placebo and between recanalization and no recanalization. The relationship between Tmax severity and infarct risk was quantified by a multilevel random-effect model estimation as for performance validation. The predicted infarct volume for each patient was estimated using a model trained on the 10 remaining patients. Cycling through the patients in this way, we estimated these coefficients of Risk_{Tmax} 11 times, each time applying the model to the patient left out of the coefficient estimation. The severity-weighted infarct volume predictions were compared with those obtained by Tmax thresholding (Tmax 2 to Tmax 14 seconds). Analysis was restricted to nonrecanalized patients.

**Severity-Weighted Mismatch Definition**

Finally, we applied the Tmax severity versus risk model to the entire EPITHET cohort of 101 patients and classified mismatch patients using this new definition. A severity-weighted mismatch was defined as predicted infarct volume/DWI >1.2 and absolute difference >10 mL. We examined whether infarct growth was attenuated by alteplase in patients with mismatch by the new definition.

**Statistical Analysis**

Statistical analyses were performed using Stata/IC 10 (StataCorp). The 2-sample Wilcoxon rank-sum test was used for continuous variables, and Fisher exact test was used for categorical variables between patients with alteplase and placebo and between recanalization and no recanalization. The relationship between Tmax severity and infarct risk was quantified by a multilevel random-effect model estimation as for performance validation. The predicted infarct volume for each patient was estimated using a model trained on the 10 remaining patients. Cycling through the patients in this way, we estimated these coefficients of Risk_{Tmax} 11 times, each time applying the model to the patient left out of the coefficient estimation. The severity-weighted infarct volume predictions were compared with those obtained by Tmax thresholding (Tmax 2 to Tmax 14 seconds). Analysis was restricted to nonrecanalized patients.

**Statistical Analysis**

Statistical analyses were performed using Stata/IC 10 (StataCorp). The 2-sample Wilcoxon rank-sum test was used for continuous variables, and Fisher exact test was used for categorical variables between patients with alteplase and placebo and between recanalization and no recanalization. The relationship between Tmax severity and infarct risk was quantified by a multilevel random-effect model estimation as for performance validation. The predicted infarct volume for each patient was estimated using a model trained on the 10 remaining patients. Cycling through the patients in this way, we estimated these coefficients of Risk_{Tmax} 11 times, each time applying the model to the patient left out of the coefficient estimation. The severity-weighted infarct volume predictions were compared with those obtained by Tmax thresholding (Tmax 2 to Tmax 14 seconds). Analysis was restricted to nonrecanalized patients.
Results

Out of 101 patients enrolled in EPITHET, 99 patients had baseline PWI and 92 had both baseline PWI and day 90 T2-weighted imaging (day 92 interquartile range, 90–95). Six patients with parenchymal hematoma and 4 patients with severely distorted EPI imaging were excluded from the assessment of Tmax stratified infarct risk. For the remaining 62 patients (29 received alteplase and 33 received placebo), baseline PWI and day 90 T2-weighted imaging were used for this analysis. Recanalization status could be assessed in 29 patients with adequate pretreatment and posttreatment MRA scans. Of these, recanalization was present in 18 patients. The Table shows baseline characteristics, and there was no significant difference between the alteplase and placebo groups (interaction Tmax by treatment, \( P > 0.05 \)).

<table>
<thead>
<tr>
<th>Table. Baseline Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
</tr>
<tr>
<td>Median age (y)</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Hypertension</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
</tr>
<tr>
<td>Hyperlipidemia</td>
</tr>
<tr>
<td>Atrial fibrillation</td>
</tr>
<tr>
<td>Current or past smoker</td>
</tr>
<tr>
<td>Median baseline DWI volume (mL)</td>
</tr>
</tbody>
</table>

The positive relationship between infarct risks and Tmax severity was maintained in subgroups of treatment (Figure 3A) and recanalization status (Figure 3B). Corresponding increases in risk differed significantly between alteplase and placebo subgroups (interaction Tmax by treatment, \( P < 0.001 \)) and between recanalization and no recanalization subgroups (interaction Tmax by recanalization, \( P < 0.001 \)). Patients without recanalization and those treated by placebo had high infarct risks, particularly in severe Tmax strata.

Prediction of Final Lesion Volumes

Figure 4 shows the Tmax versus severity relationships in the 11 patients with no recanalization. To assess the accuracy of infarct volume prediction, we calculated the differences (residuals) between baseline PWI volumes and 90-day infarct volumes for conventional Tmax thresholds in 11 patients with no recanalization (Figure 5A). Low Tmax thresholds, such as \( \leq 2 \), \( \geq 4 \), and \( \geq 6 \) seconds, overestimated the 90-day infarct volumes, whereas high thresholds, such as \( \geq 12 \) and \( \geq 14 \), underestimated the same volumes. Tmax thresholds of \( \geq 8 \) and \( \geq 10 \) seconds provided the most accurate estimation of the 90-day infarct volumes (median differences of 31 and \( \sim 35 \) mL, respectively). The Tmax \( \geq 10 \) estimates did not differ significantly from the severity-weighted estimate (\( P = 0.2 \)), but all other Tmax estimates showed significantly worse estimation of the final infarct with a maximum \( P = 0.0068 \). For comparison, we also analyzed the DWI performance, which was outperformed by the severity-weighted prediction (\( P = 0.032 \)).

Severity-Weighted Mismatch Definition

When mismatch was assessed in all EPITHET patients using the severity-weighted predicted infarct volumes generated by the infarct risks in the nonrecanlalized patients, the prevalence of the severity-weighted mismatch was 62% (61/99; Figure 5B). For the EPITHET primary outcome measure, the patients with severity-weighted mismatch showed trends to infarct growth attenuation with alteplase compared with placebo, such as geometric mean (\( P = 0.2681 \)), relative growth (\( P = 0.0716 \)), absolute growth (\( P = 0.0873 \)), and difference in cube-root lesion volume (\( P = 0.0563 \)), although it did not reach statistical significance, perhaps because of the lower number of patients. In comparison, the original EPITHET mismatch definition using the threshold approach with...
Tmax $\approx 2$ seconds that classified 86% (85/99) of patients with mismatch was also nonsignificant for all analytical methods of infarct growth, such as geometric mean ($P=0.239$), relative growth ($P=0.054$), absolute growth ($P=0.126$), and difference in cube-root lesion volume ($P=0.069$).

**Discussion**

The present study showed that the risk of infarction increased with the severity of perfusion deficits instead of showing the expected threshold, although across patients Tmax 8 to 10 seconds did exhibit a sharp slope (Figure 3B). The source of this probabilistic relationship appeared to be caused by both intrapatient and interpatient variability in the Tmax–infarct risk relationship. This suggests that no single PWI threshold valid across patients can perfectly distinguish between salvageable tissue and areas destined to infarct (Figure 4). Hence, the mismatch between single PWI thresholds and DWI lesion volumes may provide an oversimplified penum-
A model to predict final infarction must be based on patients with no reperfusion. The clinically relevant question is what will happen to the tissue if there is no reperfusion, and this needs to be estimated in patients with no reperfusion. It has to be acknowledged that many patients have reperfusion spontaneously and will never have infarction to an extent predicted by the model. It is likely that further research can help identify this subset of patients by means of occlusion site or other reperfusion determining factors and, in this way, can improve the model for these patients. In the present study, we based the predictive model on patients without recanalization as a proxy for no reperfusion because the best reperfusion definition for this type of study is unknown. Interestingly, the infarct risks did not reach 100%, even for the highest T_max range in the nonrecanalized group. We speculate that factors such as collateral supply in some cases inflate T_max and do not cause infarction, which is consistent with recent findings that T_max reflects a combination of hemodynamic effects. Visual inspection indicated that regions close to sinuses and edges were often associated with high T_max values and absence of infarction. Lacking automatic ways to distinguish artifactual areas, and given that they often are of modest volume, we did not correct for these artifacts; however, future efforts to detect these artifactual regions are needed. Patients who died before the day 90 follow-up imaging represent a possible data bias. Another potential bias is tissue shrinkage by day 90, causing underestimates of the volume of infarction and thereby underestimation of risk for a given T_max value. Future studies may be able to address this with earlier imaging follow-up, limiting the influence of atrophy.

The probabilistic approach to infarct prediction used in this study is not novel and positron emission tomography studies have demonstrated a probabilistic, rather than threshold-based, relationship between acute perfusion and...
tissue infarction. Several methods have been proposed that integrate information from several modalities and use a variety of transformation techniques to estimate the infarct probability. However, these techniques still have not found widespread use despite their theoretical appeal. The method presented here only uses DWI and PWI to classify a mismatch and is simple to implement. It does require exclusion of false-positive areas, in our case performed by a manual exclusion of false-positives, which is relatively quick to perform (3–5 minutes). The actual calculation of the predicted volume is then a few algebraic operations of microseconds in duration. To the best of our knowledge, a severity-weighted mismatch definition has not been presented before. In addition, a map of infarct risk can be generated by substituting each Tmax strata with the associated probability to produce a visual topographical representation of infarct risk in the individual. In its current form, the method does not require coregistration to calculate a volumetric mismatch; however, as fully automated registration is becoming increasingly available, it is likely that registering the acute DWI to the PWI could provide more accurate mismatch estimates that account for DWI regions already reperfused. With coregistered data, it would also be possible to attribute higher infarct risks to tissue already DWI-positive, irrespective of the concurrent Tmax value.

There are some limitations to this study. First, the analyzed cohort was small and we had a limited number of patients without recanalization. The estimated probability of infarction in each Tmax strata could be improved by a larger cohort with acute PWI and follow-up images, and it would enhance the predicted infarct volume by the severity-weighted approach. Second, regarding all perfusion parameters, Tmax is sensitive to several hemodynamic effects and this is bound to introduce some variability in the risk estimates. Third, our estimates of risk will be affected by fluctuating perfusion levels. Ideally, this analysis could be improved by limiting it to nonreperfusing tissue as determined by coregistration of acute and follow-up perfusion imaging. Unfortunately, day 3 perfusion reperfusion assessment in these data makes this impossible for patients with edema, which peaks at day 3 to 5 and is largest in patients without reperfusion, which is exactly the group of interest. Earlier reperfusion assessments, at the end of infarct evolution, but before severe edema development, may partially address this issue. Limiting the at-risk region to nonreperfused tissue, albeit limited by coregistration errors, did not significantly improve prediction of final infarct volume over those based on Tmax ≥8-second estimates. Fourth, the mismatch calculation in this study assumes that DWI is a reasonable representation of a core matter that remains controversial, but this topic previously has been systematically examined.
in this data set and has been found to have negligible impact on mismatch classifications.\textsuperscript{24}

Despite these limitations, this study demonstrated that the probability of infarction is strongly tied to the severity of PWI deficits. A penumbral definition based on $T_{\text{max}} \geq 2$ is now considered inadequate because of a penumbral definition and our results corroborating this (showing that $T_{\text{max}}$ 2 seconds to $T_{\text{max}}$ 4 seconds have a median risk of only 8%); also, in the absence of more severe areas of hypoperfusion, a penumbral definition of 2 seconds would grossly overestimate the area of final infarction.\textsuperscript{25}

In summary, we have shown that infarct risks increased with the severity of perfusion deficits and that in a small population of 11 patients a probabilistic method for infarct volume prediction outperforms all $T_{\text{max}}$ thresholds statistically, with exception of the $T_{\text{max}} \geq 10$ threshold, for which it was only favored by a trend. Thus, although we do not provide formal evidence that the proposed method performs better than using a $T_{\text{max}} \geq 10$ threshold, the data suggest it is of potential interest and deserves further investigation in larger samples. It would seem ideal to perform both the reperfusion and final infarct assessments at an earlier time point to minimize the confounding effects of late futile reperfusion and infarct shrinkage.

\textbf{Sources of Funding}

Supported by National Health and Medical Research Council, Australia; National Stroke Foundation, Australia; and Heart Foundation of Australia.

\textbf{Disclosures}

None.

\textbf{References}


8. Calamante F, Christensen S, Desmond PM, Oستergaard L, Davis SM, Connelly A. The physiological significance of the time-to-maximum (Tmax) parameter in perfusion MRI. *Stroke*. 2010;41:1169–1174.


Moving Beyond a Single Perfusion Threshold to Define Penumbra: A Novel Probabilistic Mismatch Definition

Yoshinari Nagakane, Soren Christensen, Toshiyasu Ogata, Leonid Churilov, Henry Ma, Mark W. Parsons, Patricia M. Desmond, Christopher R. Levi, Kenneth S. Butcher, Stephen M. Davis, Geoffrey A. Donnan and for the EPITHET Investigators

Stroke. published online April 12, 2012:

Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2012 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/early/2012/04/12/STROKEAHA.111.643932

Data Supplement (unedited) at:
http://stroke.ahajournals.org/content/suppl/2013/10/02/STROKEAHA.111.643932.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/
ペナンブラを定義する単一灌流閾値からの移行——新しい確率的ミスマッチの定義

Yoshinari Nagakane, MD, PhD1; Soren Christensen, PhD2; Toshiyasu Ogata, MD, PhD3; Leonid Churilov, PhD1,4; Henry Ma, MBBS3; Mark W. Parsons, PhD6; Patricia M. Desmond, MD, FRACP4; Christopher R. Levi, FRACP4; Kenneth S. Butcher, MD, PhD5; Stephen M. Davis, MD, FRACP2; Geoffrey A. Donnan, MD, FRACP6; for the EPITHET Investigators

1National Stroke Research Institute, Florey Neuroscience Institutes, Austin Health, University of Melbourne, Australia; 2Department of Neurology, Royal Melbourne Hospital, University of Melbourne, Australia; 3Department of Mathematics and Statistics, University of Melbourne, Australia; 4Department of Neurology, Hunter Medical Research Institute, John Hunter Hospital, University of Newcastle, Australia; 5Department of Neurology, University of Alberta, Edmonton, Alberta, Canada; 6Department of Radiology, Royal Melbourne Hospital, University of Melbourne, Australia.

Abstract

背景および目的：灌流強調画像が拡散強調画像を超えると定義される病変容積のミスマッチが、虚血性ペナンブラ組織のマーカーとして用いられてきた。閾値設定による灌流病変の決定は、実用的なツールとして期待されてきた。陽電子断層撮影法を用いた複数の研究では、灌流と梗塞との関係に確率的な関係があることが示されている。今回、我々は脳卒中後3～6時間の組織プラスミノゲン活性化因子治療に関する無作為比較試験のデータセットを用いて、(1) 低灌流の重症度（Tmax によって測定）と梗塞のリスクとの関係を定量化し、(2) この関係を利用して二分法ではなく梗塞の確率に基づいた新しいミスマッチの定義を提示し、(3) 新しい定義によるミスマッチが認められた患者のサブグループにおいて治療の反応を評価した。

方法：Echoplanar Imaging Thrombolytic Evaluation Trial（EPITHET）の患者を本研究に組み入れた。ベースラインの灌流強調画像および90日時のT2強調画像を重ね合わせた。灌流強調画像の病変容積を10段階のTmax遅延に層別化し、梗塞のリスクを特定のTmax層で、90日までに梗塞へ進行する組織分画を定義した。

結果：62例の患者を評価した。梗塞のリスクは、コホート全体も含めたすべてのサブグループにおいて、Tmaxの増加関数であった。確率論的アプローチは、好ましい傾向が認められたTmax≧10閾値の層を除けば、すべてのTmax閾値に優っていた。

結論：梗塞のリスクおよび治療効果は、灌流異常の重症度が高いほど大きかった。これは重症度を重視したミスマッチの定義が、ペナンブラ組織の範囲をより正確に判定することを示唆している。

Stroke 2012; 43: 1548-1555

図1：異なるTmax病変プロフィールの実例。患者AおよびBは同等のTmax≧2病変容積を有している（それぞれ228および255mL）、が、Tmax重症度プロフィールを検討した場合、患者Aは非常に遅延したTmaxの層（カラーコードを参照）が患者Bよりもも多かった。患者Aは急性期に145mLであった拡散強調画像（DWI）病変が、90日時には188mLまで広がっていたのに対し、患者Bは3mLのDWI病変が90日時には5mLまで広がっていた。重症度プロフィールはすべての断面においてはまるが、ここでは各患者の代表的な1つの切片が示されている。

Stroke誌の図を一部省略して記載

図5：再開通が得られなかった患者（11例）における灌流強調画像（PWI）の容積が90日時の梗塞容積との差。バーは最高値および最低値を、ボックスは四分位範囲を、またボックス中の横線は中央値を示す。