Introduction to the ACNR Stroke Series

The ischaemic penumbra is a concept familiar to all clinicians and scientists interested in ischaemic stroke. The concept of “potentially salvageable tissue” has revolutionised the practice of stroke medicine: intravenous thrombolysis is now routine, and other thrombus removal techniques are under active investigation. However, despite many years of investigation, how imaging the penumbra may help to select patients most likely to benefit from reperfusion treatment remains to be fully settled. In this article in our Stroke Series, we are very lucky to have a clear and authoritative account of the development of penumbral imaging - from initial experiences preclinical and through to the very latest advances incorporating advanced MR imaging – by Josef Alawneh and Jean-Claude Baron.

Jean-Claude Baron
Trained in clinical neurology at the Salpêtrière Hospital, Paris, in Medical Physics at the CEA, Orsay, and in functional brain imaging at Harvard, USA. In 1986 he was appointed Inserm Director of Research, Director of the Inserm Unit and Scientific Director of the PET Centre in Caen, France. In 2000 he was appointed Professor of Stroke Medicine and Consultant Neurologist at Cambridge. Since 2010, he has been back at Inserm at the Centre for Psychiatry and Neuroscience, Sainte-Anne Hospital, Paris, and until recently still managed his Cambridge group as Emeritus Professor. He was awarded the Johannes Wepfer award of the European Stroke Conference in 2005.

Correspondence to:
Dr Josef Alawneh
Neurology Department,
Addenbrookes Hospital
Hills Road
Cambridge, CB2 0QQ
Email: josefalawneh@doctors.org.uk

Conflict of interest statement:
The authors indicate that there are no conflicts of interest.

Provenance and peer review:
Commissioned and externally reviewed.

To cite:

Penumbral imaging in acute stroke: a triumph of hope over experience?

It's been more than thirty years since Astrup, Siesjö and Symon described in a model of middle cerebral artery (MCA) occlusion in baboons the presence of two separate cerebral blood flow thresholds below which electrical and energy failure occur respectively. They noted that between those two thresholds neuronal activity temporarily seizes but could potentially recover if perfusion is restored. This area resembled the halff-shaded zone around the centre of a complete solar eclipse and thus was called the 'ischaemic penumbra'. The potential clinical benefits of its salvage and the importance of developing tools to identify it were already noted since those early days.

The first studies that proved the existence of the penumbra in acute stroke in humans used Positron Emission Tomography (PET) which has remained the gold standard imaging technique until now. Furlan et al showed using 18-O-PET the presence of a critically hypoperfused tissue with high oxygen extraction fraction (OEF) that either recovered or progressed to infarction, whose volume correlated with acute neurological deficit and whose recovery was strongly associated with clinical improvement.

Therefore, this tissue fulfilled all the prespecified operational criteria for the penumbra, namely a) ischaemic brain tissue (CBF <20ml/100g min, i.e., ~40% of normal, and high OEF), b) that is functionally affected, c) is of uncertain fate, becoming part of the infarct or recovering depending on subsequent (early) reperfusion, and d) whose survival, if any, underlies clinical recovery.

The importance of treatments that can salvage the penumbra therefore became very apparent and it wasn't long before salvage of penumbra was demonstrated after IV thrombolysis using PET imaging. However, PET imaging has several limitations precluding its widespread use in acute stroke; these include unavailability, technical difficulty and high cost. This in turn shifted the interest to other imaging modalities potentially able to depict the penumbra in the acute stage. One of these modalities is Magnetic Resonance Imaging (MRI) and in particular the development of the Diffusion Weighted and Perfusion Weighted Imaging (DWI and PWI respectively). The DWI sequence is very sensitive in detecting severely ischaemic tissue within minutes after stroke onset, furthermore, DWI lesions strongly predict final infarct within a few hours of an acute stroke and thus these lesions have been regarded as representing the infarct core. PWI, on the other hand, can depict all hypoperfused tissue and thus can identify not only the infarct core but also the tissue at-risk of infarction. It became, therefore, common to regard the mismatch between DWI and PWI as penumbral tissue due to its location outside the infarct core but still being at risk of infarction. Moreover, several early studies have shown that salvage of the DWI/PWI mismatch is associated with neurological improvement, thus fulfilling another of the cardinal operational criteria of penumbra.

The DWI/PWI mismatch concept, however, has
its limitations as well. One of these is that very early DWI lesions don’t always represent the infarct core and can permanently reverse in part or in whole following acute reperfusion, thus behaving like penumbral tissue, in addition, the precise penumbral perfusion thresholds on PWI imaging are still the matter of ongoing research. This has led researchers to search for other sequences that might depict the penumbra more precisely. One of these used the BOLD T2* MRI signal affected by deoxyhaemoglobin to map the OEF. This has been reported with varying success but hasn’t been able to take over from the DWI/PWI paradigm so far.

As an alternative to MRI, Computer Tomography Perfusion (CTP) imaging has been extensively used to define the penumbra. Compared to MRI, it is more easily accessible and faster than MR; moreover, it is suitable for uncooperative patients and those with possible MR contraindications. Recently the methodology using CTP has improved to provide both full brain coverage and CT angiogram following a single contrast injection. CTP thresholds to identify core and penumbra have been validated by comparing with PWI/DWI imaging before iv thrombolysis was the EPITHET study which aimed to test the validity of the mismatch concept in the 3-6 hours time window. However, the study design was to assess in retrospect whether the presence of mismatch was a predictor of good response to iv therapy versus placebo, i.e., randomisation was not based on MRI criteria. Patients with mismatch were defined as those with PWI/DWI ratio $\geq$1.2 and a PWI-DWI mismatch ratio $\geq$8s) PWI lesion $\geq$10mL. The study showed that reperfusion in patients with the ‘target mismatch’ profile was associated with favourable outcome when compared to reperfusion in patients with no mismatch.

The first randomised placebo controlled study that prospectively acquired MRI with DWI/PWI imaging before iv thrombolysis was the EPITHET study which aimed to test the validity of the mismatch concept in the 3-6 hours time window. However, the study design was to assess in retrospect whether the presence of mismatch was a predictor of good response to iv therapy versus placebo, i.e., randomisation was not based on MRI criteria. Patients with mismatch were defined as those with PWI/DWI ratio $\geq$1.2 and a PWI-DWI volume $\geq$10mL, with the PWI lesion defined as tissue with T1max $\geq$2s (Table 1). 101 patients were randomly assigned to receive alteplase or placebo. The primary endpoint was infarct growth between baseline DWI and day 90 T2 lesion in the mismatch patients only and this was non-significantly lower in the alteplase group. The study found, however, that alteplase was significantly associated with reperfusion in the mismatch group, and reperfusion in turn was significantly associated with better neurological outcome. However, as will be discussed below, post-hoc reanalysis of this dataset using more appropriate perfusion thresholds and endpoints showed strongly positive results. The need for larger randomised studies using well-designed penumbra imaging in the inclusion criteria became ever more apparent following this study.

Around the same time two pilot studies (DIAS and DEDAS) of a new thrombolytic drug were published. DIAS was a dose-finding phase II randomised trial designed to evaluate the safety and efficacy of intravenous desmoteplase, a novel plasminogen activator, administered within 3-9 hours of ischaemic stroke onset in patients with perfusion/diffusion mismatch on MRI (Table 1). 35 Part 1 of the trial was terminated prematurely due to the high rate of symptomatic intracranial haemorrhage (sICH). Following reduction in the administered dose, part 2 of the study showed lower rates of sICH using doses up to 125µg/kg. The conclusions were that desmoteplase in patients with mismatch was associated with a higher rate of reperfusion rates and better clinical outcome compared with placebo. This pilot study was followed by DEDAS 36 which also used perfusion/diffusion mismatch ratio $\geq$1.2 to assess the safety and efficacy of desmoteplase 90µg/kg and 125µg/kg in patients with acute ischaemic stroke 3-9 hours after onset. This small phase II study (n=37) confirmed the findings of DIAS that IV desmoteplase at those doses is safe and may improve clinical outcome.

The findings of the above two phase II desmoteplase studies led to the first large phase III study that used penumbra imaging in its inclusion criteria. DIAS II was a randomised, placebo-controlled, double blinded, dose-ranging study comparing desmoteplase 90 µg/kg and 125 µg/kg versus placebo in acute ischaemic stroke given 3-9 hours after onset. The primary end point was clinical response rate at day 90, defined as a composite of improvement in NIHSS score of $\geq$8 points or an NIHSS score $\geq$1, a modified Rankin scale score $\geq$0-2, and a Barthel index of 75-100. 193 patients were randomised, and 186 patients received treatment: 57 received 90µg/kg desmoteplase; 66 received 125µg/kg desmoteplase; and 63 received placebo. The clinical response rates at day 90 were 47% for 90µg/kg desmoteplase, 36% for 125µg/kg desmoteplase and 46% for placebo, showing no statistical benefit of desmoteplase given 3-9 hours after onset of stroke. Patients were included if they had a distinct ‘penumbra’ with a perfusion lesion/core ratio $\geq$1.2 based on MRI (PWI/DWI) or CT (CTP/CT) (Table 1).

### Table 1: Criteria defining patients with mismatch

<table>
<thead>
<tr>
<th>STUDY</th>
<th>N</th>
<th>Time to treatment (hrs)</th>
<th>Perfusion lesion/core ratio</th>
<th>Perfusion lesion definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEFUSE</td>
<td>74</td>
<td>5.5</td>
<td>$\geq$1.2 PWI/DWI</td>
<td>Tmax $\geq$2s</td>
</tr>
<tr>
<td>EPITHET</td>
<td>101</td>
<td>4.9 alteplase</td>
<td>$\geq$1.2 PWI/DWI</td>
<td>Tmax $\geq$2s</td>
</tr>
<tr>
<td>DIAS 104</td>
<td>37</td>
<td>7.4</td>
<td>$\geq$1.2 PWI/DWI</td>
<td>Perfusion lesion not defined</td>
</tr>
<tr>
<td>DIAS II</td>
<td>186</td>
<td>6.5</td>
<td>$\geq$1.2 (PWI/DWI, some with CTP)</td>
<td>Perfusion lesion not defined on CTP or PWI</td>
</tr>
<tr>
<td>Tenecteplase versus alteplase</td>
<td>75</td>
<td>2.7 alteplase</td>
<td>$\geq$1.2 CTP/Core</td>
<td>Transit time maps/core based on CBV (both without threshold)</td>
</tr>
<tr>
<td>DEFUSE II</td>
<td>99</td>
<td>4.5 (time to baseline MRI)</td>
<td>$\geq$1.8 PWI/DWI</td>
<td>Tmax $\geq$6s</td>
</tr>
<tr>
<td>MR RESCUE</td>
<td>118</td>
<td>5.5 (time to enrolment)</td>
<td>$\geq$1.4 PWI/DWI And CTP/Core</td>
<td>Tmax $\geq$6s</td>
</tr>
</tbody>
</table>
Although the desmoteplase trials deliberately applied still unproven imaging markers to a time-window beyond those investigated in observational studies, a strategy that in retrospect was unlikely to meet with success, the negative results of DIAS II came as a surprise to the stroke community at large, and raised several questions on the imaging definition of penumbra as will be discussed below.

To enrich the population studied by recruiting a more homogeneous group, a randomised trial of tenecteplase versus alteplase for acute ischaemic stroke used the presence of penumbra on perfusion imaging as an inclusion criterion, in order to enrich the population studied by recruiting a more homogeneous group. This novel design allowed them a considerable reduction of the required sample size. This was a Phase 2B 3-arms trial that assigned 75 patients to receive alteplase (0.9mg/kg) or tenecteplase (0.1mg/kg or 0.25mg/kg) less than six hours after stroke onset. Inclusion criteria included the presence of perfusion lesion at least 20% greater than core, with volume of at least 20 ml and an associated vessel occlusion on CTA. Perfusion lesions were based on CTP transit-time maps, and core defined using cerebral blood volume maps. The coprimary end points were the proportion of the perfusion lesion that was reperfused at 24 hours on MRI PWI and the extent of clinical improvement at 24 hours assessed on the NIHSS score. Together, the two tenecteplase groups had greater reperfusion (p<0.004) and clinical improvement (p<0.001) at 24 hours than the alteplase group with no significant differences in serious adverse events. In this study, time to treatment was 2.7 hours in the alteplase group and 3.1 hours in the tenecteplase group which is the earliest treated in the tenecteplase trial. This would imply that significant amounts of the tissue is expected to be largest. This was despite the suggestion from observational studies that penumbral imaging even within the 3-hour window enhanced the clinical benefit from IV t-PA. The challenge faced by EPITHET and DIAS II, therefore, was to capture this shrinking penumbra beyond 3 hours and prove that its salvage was of clinical benefit. While this has proved to be very hard, the negative results from the primary analysis of these two studies have provided us with important lessons for the future.

One of the cardinal questions these two studies had to address was how to define penumbral tissue using MR, and both EPITHET and DIAS II relied on the PWI/DWI mismatch concept. EPITHET used Tmax ≥2 seconds as the perfusion threshold to define penumbral tissue and DWI lesion for the core. DIAS II used any perfusion abnormality seen without specifying a threshold as the penumbral threshold and the DWI lesion again as the core. In retrospect, it is now clear that these definitions were too liberal and resulted in the inclusion of significant oligaemic tissue that was not dysfunctional nor at risk. Observational studies that directly compared MRI perfusion parameters to PET reported Tmax values around 5.5 seconds as the closest to the penumbral threshold, implying that significant amounts of the tissue regarded as penumbra in those studies were more oligaemic. In DIAS II the perfusion/core mismatch ratio in percentage was reported to be ≤50% more than 6 hours after stroke onset; this, in deed, is unlikely to represent true penumbral tissue at those late time points. Moreover, up to 67% of patients screened were thought to have significant mismatch volume penumbral imaging in their inclusion criteria. A first observation is that while the results of the pilot and observational studies were promising, those of the main randomised studies were negative. This raised important questions regarding the methodology used, and even some went as far as questioning the penumbral concept itself. The possible reasons for these negative results will be explored here by attempting to answer the question: Why did EPITHET, DIAS II and MR RESCUE fail to support the notion that using penumbral imaging is clinically useful?

To start with, as already noted, when the first alteplase study was conducted, imaging of the penumbra was not widely available and thus wasn’t required for recruitment. Subsequently, once alteplase was licensed for use within 3 hours of stroke onset, the emphasis was to get plain CT imaging swiftly and very little room was left for any research on penumbra imaging within that 3 hour window. The first thrombolysis studies using penumbral imaging (EPITHET and DIAS II), therefore, recruited patients beyond the 3 hours, looking into extending that window. This deprived us, however, and perhaps for ever, of any evidence of the clinical benefit from penumbra imaging in the first 3 hours post stroke onset, at a time when penumbral tissue is expected to be the largest. This was despite the suggestion from observational studies that penumbral imaging even within the 3-hour window enhanced the clinical benefit from IV t-PA. The challenge faced by EPITHET and DIAS II, therefore, was to capture this shrinking penumbra beyond 3 hours and prove that its salvage was of clinical benefit. While this has proved to be very hard, the negative results from the primary analysis of these two studies have provided us with important lessons for the future.

One of the cardinal questions these two studies had to address was how to define penumbral tissue using MR, and both EPITHET and DIAS II relied on the PWI/DWI mismatch concept. EPITHET used Tmax ≥2 seconds as the perfusion threshold to define penumbral tissue and DWI lesion for the core. DIAS II used any perfusion abnormality seen without specifying a threshold as the penumbral threshold and the DWI lesion again as the core. In retrospect, it is now clear that these definitions were too liberal and resulted in the inclusion of significant oligaemic tissue that was not dysfunctional nor at risk. Observational studies that directly compared MRI perfusion parameters to PET reported Tmax values around 5.5 seconds as the closest to the penumbral threshold, implying that significant amounts of the tissue regarded as penumbra in those studies were more oligaemic. In DIAS II the perfusion/core mismatch ratio in percentage was reported to be ≤50% more than 6 hours after stroke onset; this, in deed, is unlikely to represent true penumbral tissue at those late time points. Moreover, up to 67% of patients screened were thought to have significant mismatch volume...
at 6 hours post onset, again a very large proportion which might have included patients with oligemia but not true mismatch. To address the above, a post hoc analysis was recently published that refined the definition of mismatch, first in DIAS II, and then in a pooled cohort from DIAS, DEDEAS and DIAS II. Mismatch was re-defined as the difference between the calculated volume of hypoperfusion based on MTT maps (no defined threshold) and volume of DWI, but tested only in patients with mismatch volume >60mL. The difference in favourable response rate to desmoteplase between placebo and treatment group increased with the new definition but was statistically significant only in the pooled analysis. The authors concluded that desmoteplase appears beneficial in patients with large mismatch volume but that this will need to be tested in prospective studies. Along the same lines, in EPITHET significant oligaemic tissue was regarded as penumbra and it is suspected that a sizeable number of patients were misclassified. This is supported by the fact that the vast majority of patients (86%) screened were judged to have significant mismatch volume 3-6 hours post onset, again a clear overestimation. And when penumbra was redefined in a post hoc around 1.4 but failed to demonstrate any positive effect. DEFUSE II went up to 1.8 and had more promising results; however, this was an observational study with no randomised control arm. In order to address the question of which should be the optimum ratio to use, a reanalysis of the DEFUSE data was carried out, which evaluated the odds ratio for a favourable clinical response in mismatch patients with reperfusion compared with no reperfusion for various mismatch ratio thresholds. A mismatch ratio of 2.6 provided the highest sensitivity and specificity for identifying patients in whom reperfusion was associated with favourable response. Such a high mismatch ratio has not been tested in any randomised study yet, but clearly would considerably reduce the number of recruitable patients.

Indeed, an important aspect of these studies is the number needed to recruit to be able to show significant difference in clinical response. The aim of penumbra imaging is to screen patients and select those who are most likely to benefit from thrombolysis and thus reduce the total number of participants in the studies. However, this hasn’t happened on the ground as shown above; and thus larger numbers of patients possibly were eventually face the same issues. This was noted in the MR RESCUE trial which used penumbra imaging before intervention. Here Tmax >6h threshold was used to define mismatch; mismatch ratio of around 1.4 was used which is marginally but definitely better than the 1.2 previously used; and the mean time to enrolment was 5.5 hours. The total number of patients recruited was 118 but those with a favourable penumbra pattern were only 68 (34 embolometry and 34 standard care group) thus significantly reducing the power of this study. An additional confounder in this study was that a percentage of the patients had IV thrombolysis before enrolment. This may be relevant to the fact that despite allocation to embolometry, recanalisation and reperfusion rates were similar between the treatment and control group. Reperfusion is the most important factor leading to salvage of the penumbra, and without it, no matter the volume of penumbra, good recovery is unlikely. Taking all the above limiting factors together, plus the fact that this study started many years ago and therefore used largely outdated thrombectomy devices and that both MR and CT perfusion were used to determine the favourable penumbral pattern, maybe it is not surprising that it was negative.

...although already widely used in clinical routine, penumbra imaging for the selection of candidates for reperfusion therapy both within and beyond the current licensed time-window for IV thrombolysis has not been validated yet

There are a few additional difficulties that studies aiming at imaging the penumbra have to face. One is the postprocessing used after acquiring the imaging. Various methods can be used, leading to different maps, each providing potentially different information. There has been no consensus yet on the best methodology, software, maps and thresholds to use for each particular method; however, there are recommendations regularly published addressing these topics (Stroke imaging research roadmap) which would be advisable to adhere to in order to improve consistency between studies. This remains a topic of active research and there is still some way to go until these questions are resolved. However, proprietary issues are involved here as well, as, at the moment each manufacturer offers commercial software which works as a ‘black box’ that delivers pretty blobby maps of penumbra and core of unclear physiological validity. In addition, in order for post-processing tools to be useable in clinical trials and also in future routine practice, they must consistently deliver reliable core/penumbra maps within a minute or two of data acquisition at most. Recently developed tools are being tested currently.

While the future of penumbra imaging in acute stroke within or beyond treatment

analysis using a threshold of Tmax>6s and proper image co-registration, the median mismatch volume fell from 126mL to 37mL and only 53% of patients were now found to have had true penumbra. This lead to several patients being reclassified, and a reanalysis with the new definition of penumbra including pooled patients from DEFUSE showed that favourable tissue and clinical outcome from reperfusion in the target mismatch group became statistically significant.

The other important question the above two studies had to address was what mismatch ratio is to be considered significant. This is the ratio of total affected tissue divided by core. Increasing the ratio would decrease the absolute volume of mismatch and also the chances of observing clinical improvement should that mismatch tissue survive; however, the more you increase the ratio, the more patients will be excluded as having no mismatch. Both DIAS II and EPITHET used 1.2 as their cutoff ratio, meaning that the affected but not core tissue‘penumra’ had to be at least 20% larger than the DWI lesion ‘core’, which in retrospect seems small. As the results of both trials were negative, subsequent studies used larger ratios in order to maximise the chances of capturing and saving clinically relevant mismatch. MR RESCUE used a ratio needed to demonstrate any benefit as the screening process has not been optimum. For example the NINDS study, the first to demonstrate benefit from rPA within the first 3 hours recruited 333 patients at a time-point where penumbra is expected to be large. ECASS III had to recruit 821 patients to demonstrate thrombolysis benefit between 3 and 4.5 hours post onset. EPITHET recruited only 101 and DIAS II 186 patients. While these numbers are relatively large compared to previous imaging studies, they are quite small compared to previous thrombolysis studies; particularly if you take into account that recruitment was between 3-6 hours at a time where penumbra volume has significantly declined. On the other hand, the above-mentioned tenecteplase trial requiring presence of penumbra as an inclusion criterion needed only 25 patients per arm to document efficacy. Calculating the sample size needed for those trials is challenging due to difficulty in estimating an accurate treatment effect. It is been estimated that about 330 patients will be needed in a study using penumbral imaging before thrombolysis treatment.

We addressed above the difficulties related to penumbra imaging in the two large IV thrombolysis studies. It is also expected that trials using an endovascular approach would...


