Systematic Review of Diffusion and Perfusion Imaging in Acute Ischemic Stroke

S.L. Keir, MRCP; J.M. Wardlaw, FRCR

Background and Purpose—Recent advances in neuroimaging have raised hopes of early and accurate identification of ischemic brain and the discrimination of dead from salvageable tissue. We sought to determine whether the data published so far are enough to establish the roles of these techniques in everyday clinical practice.

Methods—A systematic review of studies of MR diffusion-weighted imaging (DWI), perfusion imaging (PI), or a combination of the two, in human stroke, excluding abstracts and case reports. One reviewer extracted information on the size of each study, its main purpose, methodological details, and results.

Results—We identified 47 studies of DWI, 18 studies of MR PI alone or in combination with another advanced imaging modality, and 19 studies of DWI and PI together. Although high proportions of the studies were prospective and gave good details of the imaging sequences used, the majority gave very limited details on patient selection and clinical characteristics or blinded imaging assessment. Pathophysiological changes were inferred from DWI/PI patterns that were not supported by other data.

Conclusions—Despite considerable enthusiasm for and promise of these techniques, there is not sufficient information available in these studies to enable us to draw firm conclusions about the sensitivity and specificity of these techniques for identification of either ischemic lesions not visible by other means or salvageable tissue. Future studies should include larger numbers of carefully described patients, assess the contribution of DWI over and above other imaging, obtain follow-up at an appropriate time interval to determine accurate clinical and neuroradiological outcomes, and assess DWI/PI abnormality with reperfusion in randomized treatment trials. Investigators should also be encouraged to combine their individual patient data in meta-analyses to obtain a more robust assessment of the value of DWI and PI from larger sample sizes. (Stroke. 2000;31:2723-2731.)

Key Words: magnetic resonance imaging, diffusion-weighted ■ magnetic resonance imaging, perfusion-weighted ■ stroke, acute ■ stroke, ischemic

Recent advances in brain imaging may likely underpin the future of stroke care. In particular, the introduction of MR and development of advanced imaging-like spectroscopy (MRS) and diffusion-weighted and perfusion imaging (DWI and PI, respectively) have raised hopes of greater ability to discriminate dead from salvageable tissue and thus to target new treatments. Several reviews1,2 have summarized the technical aspects and possible applications of these advances, and considerable enthusiasm has been expressed for the use of these techniques in stroke; relatively few, however, have highlighted the problems that need to be overcome to maximize the potential of these exciting techniques.3,4

DWI works on the principle that sensitizing a standard MR image to diffusion weighting identifies regions of abnormal water movement, occurring very early after onset of ischemia, resulting in increased (bright) signal intensity.5 It is thought that visualization of lesions at such an early stage might identify patients before acute stroke treatment. The degree of water movement abnormality (or tissue injury) may be quantified by calculating the apparent diffusion coefficient (ADC).2

PI can detect hypoperfused regions of brain either by monitoring the transit of a rapidly injected contrast agent6 or magnetically tagged water molecules in arterial blood7 through the brain. In regions distal to an arterial occlusion, the arrival of the contrast agent or tagged water molecules may be delayed. The resulting signal-time curve can be converted into a concentration-time curve, from which several functions that describe regional perfusion can be determined.

Combining DWI and PI may enable visualization of the abnormal area soon after the onset of symptoms (previously, one would expect CT or T2-weighted MR to show an infarct) and provide information that may discriminate dead from still-recoverable brain. In hyperacute stroke treatment, this could help identify which patients are most likely to benefit from potentially risky treatments.

A study to investigate a new imaging procedure should give clear details of its aim, the definition and characteriza-
tion of its study population (eg, severity and type of ischemic stroke), the technical details of the imaging tests, whether the analysis was prospective or retrospective, how the images were analyzed and read, and whether the reading was blinded to clinical details and other imaging. If comparisons within the study population are made, an objective statistical analysis should be used to compare the populations with a null hypothesis being clearly stated. Thus, before we can establish the role of these advanced imaging techniques, we need to be confident that studies have demonstrated the efficacy, feasibility, and the reliability of the information so derived and that they do indeed identify patients for specific reasons not possible by other means. We therefore undertook a systematic review of all published studies of DWI, PI, or both in patients with stroke to see how well the above criteria had been met, what information had been obtained so far on patient characterization and acute stroke treatment guidance, and where more information on DWI and PI might be needed, and so where future research with DWI and PI should be directed.

Methods

We used the Cochrane Database of Systematic Reviews methodology to perform this systematic review. This methodology has been developed for undertaking systematic reviews of randomized controlled treatment trials, but the same principles apply to the conduct of systematic reviews of diagnostic tests and observational studies. We have used these methods successfully previously to study risks for and diagnosis of ruptured intracranial aneurysms.

Search Strategy

We searched for all published articles on DWI, PI, or a combination of the two in the English and non-English language literature in MEDLINE and EMBASE from October 31, 1999, back to the earliest studies available (using the search terms “diffusion weighted,” “perfusion weighted,” “dynamic susceptibility,” “hemodynamically weighted,” expanded to maximize the number of hits, and combined these with the Cochrane library search strategy for stroke, hand-searched 6 relevant journals (Stroke, Radiology, American Journal of Neuroradiology, American Journal of Roentgenology, Magnetic Resonance in Medicine, Journal of Magnetic Resonance Imaging) from November 1999 to January 2000, and examined reference lists in the identified articles.

Inclusion criteria: published studies in which DWI, PI, or both in combination had been conducted in humans with stroke. We excluded case reports and studies that had so far only appeared in abstract form. Where there was any doubt about the inclusion or exclusion of a study, the paper was discussed and a consensus opinion was reached.

Data Extraction and Analysis

One reviewer (S.K.) extracted information on the sample size of each study, its main purpose (ie, use of imaging to predict outcome, technical development of the imaging technique, comparison with another imaging technique), the time window from onset of stroke symptoms to imaging, how the diagnosis of stroke had been made and by whom (ie, stroke physician, neurologist, general physician, or not stated), the patient inclusion criteria, whether the patients had received any stroke treatment and whether this was randomized, the DWI or PI scanning method, the image analysis method, whether interpretation of diffusion or perfusion images was blinded to clinical details and the results of other imaging modalities, whether any patients had been imaged but then excluded from further analysis and why, whether any data on clinical outcome had been collected and at what time after stroke, and the overall conclusion of the study.

Results

Forty-seven studies concerned specifically with DWI in humans with ischemic stroke were identified (total number of patients = 2436, median 34). Also identified were 14 studies concerning MR PI alone (n = 198, median 10); 4 combining MR PI with another advanced imaging modality, such as single-photon emission computed tomography (SPECT; n = 38, median 10); and 19 studies concerning the combination of DWI and MR PI (n = 563, median 21).

General Methodological Details of All Studies

Sample populations ranged from 3 to 224 patients. In these studies, although the technical information on the imaging sequences was good, the information on general methodological details (blinding, patient selection and exclusions, and the proportion of uninterpretable scans) was limited, and generally these details were not mentioned. For instance, in the DWI-only studies, only 9 of 47 (<20%) mentioned that scans were interpreted by researchers blinded to clinical details or other imaging. Only 13% gave specific details on patient inclusions and exclusions. Only 9 studies gave details on the number (and reasons for) inadequate scans that were excluded from further analysis, although it is likely that patients with poor-quality scans were excluded from the analysis in other studies. There were no details in any of the studies on patient tolerability of the investigation (Figures 1, 2, and 3).

Studies of DWI Alone

The 47 studies of DWI had the following primary purposes (some studies had >1 primary purpose): technical development of the imaging sequence or analysis method (17 studies); change in visibility of lesions on DWI or ADC over time (6 studies); correlation of DWI with measures of neurological severity (2 studies); comparison with CT or T2 MR to see whether DWI showed more lesions (11 studies); distinguishing old from new lesions (6 studies); use of DWI to demonstrate ischemic lesions in patients with TIA or lacunar stroke (4 studies); and feasibility of DWI in acutely ill patients (6 studies). Two studies addressed identification of hemorrhage on DWI and 1 addressed the important issue of negative DWI in patients with stroke-like deficits. These will be described in turn.

Seventeen (36%) studies (total n = 567) investigated DWI techniques or refinements of techniques of image acquisition or processing, or methods of calculation of the apparent diffusion coefficient (ADC). No useful comparisons could be made between them, as they all covered different technical aspects of the imaging and mentioned little about the type of patients included.

Six (13%) studies (including a total of 564 patients) studied the visibility of DWI lesions or changes in ADC values over time. One study repeatedly scanned the majority of the patients in the study and found reduced ADC values up to 85 days after stroke. In the rest of the studies, the proportion...
of patients who had repeated scanning was much more limited\textsuperscript{18–40} or impossible to determine.\textsuperscript{15,41} In these studies, the time over which the DWI lesion remained visible was inferred by using scans of different patients at different time points, rather than repeated imaging of the same patient. Using only “snapshots” of different individual patients means that it is difficult to be precise about when a particular lesion may disappear, or an ADC value change.

Two (4\%) studies (including 92 patients) investigated the correlation of early DWI (within 24\textsuperscript{12} and 60\textsuperscript{17} hours) with measures of clinical stroke severity at the time of imaging and at either 317 or 42 weeks.\textsuperscript{12} Both studies demonstrated that acute lesion volumes on DWI correlated with the National Institutes of Health Stroke Scale (NIHSS) scores both acutely and at follow-up. In both studies, images were read blinded to clinical details, and in one\textsuperscript{17} to initial T2 MR. However, as infarction visualized on all imaging modalities tested so far correlates with stroke severity, knowledge that DWI also correlates does not add significantly to the body of knowledge.

Eleven (23\%) studies (total n = 399) suggested that DWI demonstrated more lesions in the symptomatic anatomical area at an earlier time point than did conventional imaging. Although one study directly compared CT and DWI\textsuperscript{10}...
(n = 17), the majority of studies compared DWI to conventional MR.15,18,25,40,42–47 All of these studies indicated that more lesions were visible on DWI than on conventional imaging: in 5 studies, within 6 hours or less,10,15,40,43,45 in 2 studies within 48 hours,44,46 in 2 studies within 4 days,18,42 and in 1 study “the time period under investigation” (8 hours to 12 days).25 In only 2 of the studies the scans were read by observers blind to clinical details,15,18 in only 1 study blind to other imaging,43 and in only 1 the observers were blind to both clinical details and other imaging,10 ie, in fewer than one third of studies purporting to show that DWI was better than conventional imaging was an attempt made to reduce bias by blinding the DWI interpretation. None mentioned whether an attempt was made to randomize the order in which the DWI and conventional imaging were performed or stated the order in which imaging was actually performed (ie, whether DWI was always performed after conventional imaging, in which case DWI would always be likely to show more lesions, or vice versa). The sample sizes in these studies ranged from 9 to 103 patients, which is too small to determine accurately the benefit of DWI. The proportion of patients in which DWI was felt to be superior to conventional MR ranged widely, from 5% to 71%. In 1 study it was impossible to define in how many patients DWI had proved to be superior.47

Six of 399 studies (13%) found that in patients with multiple lesions demonstrated on MR, it was possible to distinguish new from old lesions with DWI, with the acute lesion appearing hyperintense and the old hypointense on the DWI.16,18,25,41,48,49 Timing of scanning from onset of symptoms ranged from less than seven hours to 11 days. In only 2 studies were the scans read blinded to clinical details.16,18 The proportion of patients in which DWI was said to distinguish new from old lesions better than did conventional MR ranged from four to 15 out of about 220 patients with multiple lesions, ie, less than 10% of the total number of patients included.

Four (8%) studies (n = 159) concentrated on the clinicotopography of a specific stroke subtype such as lacunar stroke13,45,18 or transient ischemic attack.50 The studies concentrating on lacunar symptoms demonstrated that DWI could identify appropriate subcortical areas of ischemia. The one study concerned with transient ischemic attacks demonstrated that 48% of patients whose symptoms resolved within 24 hours had relevant lesions on DWI within 24 hours of symptom onset. Patients whose symptoms resolved by 24 hours had smaller and less obvious lesions on DWI than patients whose symptoms had lasted longer.50 Nonetheless, as with CT scanning,51 this means that DWI cannot be used to discriminate between symptoms that will turn out to be TIAs and those that will turn out to be strokes.

Six (13%) studies (n = 518), two of which were retrospective analyses, addressed the feasibility of undertaking an advanced imaging protocol in acutely ill stroke patients and suggested that it was possible.20,26,32,42,46,50 None of the imaging was read blinded, which is clinically realistic, but data on patient inclusion and exclusions, as well as clinical characteristics was often limited, ie, information to identify patients who would be poor subjects for DWI was not given. Two studies (which included 20 patients with cerebral hemorrhage) demonstrated that it was possible to identify acute hemorrhage, a subset of patients previously thought to present difficulty on DWI,52,53 although as neither study was blinded to other imaging results, the validity of this conclusion must be questioned.

One retrospective study (n = 27) raised the important issue of patients with stroke-like deficits and negative DWI imaging27 and explained that specificity had not been addressed because their entire (and substantial) case series (782 pa-
tients) had not been reviewed. Such a task would indeed have been a major undertaking, but it would have given us information on sensitivity and specificity of DWI with an accuracy that smaller studies which have attempted to address this issue have not achieved. Other studies have attempted to define specificity and sensitivity of DWI, but selection bias make it difficult to extrapolate these values to the general population with stroke-like symptoms.

**Studies of PI Alone**

One of the 14 studies of PI alone attempted to evaluate the clinical usefulness of PI (n=15); the other 13 studies (n=183) concentrated on the technical aspects of demonstrating abnormalities of regional cerebral perfusion. All but one, which attempted to quantify the arterial input function (a constant required for the calculation of absolute blood flow), assessed relative rather than absolute cerebral blood flow. All studies used the gadolinium bolus tracking method of measuring blood transit time. A variety of perfusion abnormalities were demonstrated; one study noted a heterogeneous distribution of rCBF in the expected regions of interest in all examinations; documented hyperperfusion as well as delayed or absent perfusion in the region of interest; and documented no perfusion deficits in 4 of the 11 subjects despite marked clinical signs and repeat MR within 48 hours confirming infarction.

Of the 4 studies that combined MR PI with another advanced imaging modality, perfusion MRI was correlated with SPECT findings and with xenon CT in the expected regions of interest. All found that MR perfusion techniques correlated well with the perfusion deficit demonstrated on these other modalities. Sample sizes were small, and it was not made clear whether perfusion images were read by researchers blinded to other forms of imaging.

**DWI and PI in Combination**

Nineteen studies were identified, with a total sample size of 563 (median 21). The main purpose of 3 of the studies was to document the clinicotopography of the combined imaging; one of these documented important data on the variety of patterns of DWI and PI seen. However, sample size was small, and it was not clear whether DWI and PI were read with blinding to clinical details. One study compared the combination to CT in relation to acute determination of intracerebral hemorrhage, and with xenon CT in the expected regions of interest. All found that MR perfusion techniques correlated well with the diffusion deficit demonstrated on these other modalities. Sample sizes were small, and it was not clear whether perfusion images were read by researchers blinded to other forms of imaging.

**Discussion**

Stroke medicine is now in an era of acute intervention, and practical imaging techniques that help confirm an ischemic
insult within minutes of the event and determine the amount of salvageable brain tissue remaining would obviously be invaluable tools. Hence the great flurry of enthusiasm around the exciting results of DWI and PI studies. The studies imply that DWI identifies definitely (permanently) damaged tissue and that DWI/PI mismatch identifies tissue at risk but still salvageable. Is this correct? Acute stroke imaging has come a long way in the last decade, but what has it really told us and where do we go from here?

We had hoped to determine the sensitivity and specificity of DWI for identifying acute ischemic stroke and of DWI/PI for identifying salvageable tissue, but that was not possible from the given data. Although it is likely that DWI is more sensitive to acute ischemic stroke than conventional MR or CT, sample sizes were so small, or insufficient details of clinical features, inclusion and exclusion criteria, and unsuccessful examinations were given to enable an accurate estimation of efficacy to be made. As the results of this systematic review show, publications in this field so far tend not to include important methodological details and so may be leading stroke researchers into jumping to conclusions not supported by the data. In many studies, the advanced imaging was not read blinded to the more routine imaging, and there were no details on the order in which imaging was performed. In the studies comparing DWI with conventional MR, if DWI was always performed after conventional imaging (as with studies to date that have compared conventional MR with CT), it is likely DWI would always show more lesions. There was little information on patients unable to complete DWI or PI, and most studies lacked clinical details of patient selection and case mix, with the latter information being very important for several reasons.

First, some conditions that mimic stroke clinically can also manifest abnormalities on DWI. Few studies mentioned whether any of their prospectively identified patients later turned out not to have had a stroke—though it would be unusual, in a prospective sample even as small as 40, not to find the occasional patient thought initially to have had a stroke who turned out to have a nonvascular cause of the symptoms. Conversely, the study by Ay et al and other case reports have highlighted the potential for DWI to be negative in patients with ischemic stroke. Although PI may be of some use in these patients, clarification about this issue is sorely needed.

Second, because stroke is a heterogeneous condition and the case mix of patients is likely to differ between hospitals (even within a small geographical area), this will have led to differences between studies in the type of stroke patients included. However, unless sufficient details of the patients’ clinical characteristics at baseline (eg, age, gender, some measure of neurological deficit, prestroke morbidity) and clinical status at a recognized, valid outcome point (eg, 3 months) by a validated outcome score are given, it is impossible to gauge the generalizability of the individual study results and hence the relevance to the stroke population in general. One month after stroke is far too early to determine clinical outcome and still too early for radiological outcome. There may still be either swelling (overestimating final infarct size) or fogging (grossly underestimating final infarct size). In fact, use of any radiological outcome is difficult if a volume measure or image coregistration are used, because (in addition to swelling and fogging) loss of volume in the affected area and ex vacuo effects in surrounding tissues from 4 to 6 weeks onward will lead to underestimation of final infarct volume.

Third, the clinical severity of the stroke is a profound determinant of clinical outcome. It correlates strongly with the site and extent of the lesion on CT and structural MR and therefore must be taken into consideration when any new imaging techniques are described. Most studies failed to do this adequately: the NIHSS (the most frequently used neurological score) captures part, but by no means all, of the severity of the patient’s stroke; ie, it predicts a proportion but by no means all of the patient’s outcome. Stroke severity must be fully adjusted before any additional information contributed by an imaging technique to the identification of particular patients, or the determination of outcome, can be identified. Stroke severity also influences the time to admission to hospital (patients with severe strokes are admitted sooner after stroke than those with milder strokes) and hence will influence the time to imaging. This, in turn, will probably influence the ADC values found in individual patients. The influence of stroke severity cannot be overemphasized: failure to take account of it may have contributed to perceived differences between studies that compared ADC changes over time in which “snapshots” of different patients at different times were used instead of the same patients scanned serially at different times. A similar effect most probably affects the findings in studies of DWI with PI, in which the proportion of patients with diffusion/perfusion mismatch and its size will depend on the severity of stroke and the time delay to imaging.

Although many studies gave ample details of the imaging sequences used, there are problems with the precise DWI or PI parameters to be measured, moreso with PI. Most MR PI techniques measure relative and not absolute perfusion changes, and positron emission tomography studies have demonstrated drawbacks with this approach. Relative perfusion values may not precisely demonstrate which tissue is at greatest risk of infarction, added to which is the unresolved debate about the best parameter of the PI curve to use in the analysis. Attempting to ascribe a physiological mechanism to PI findings on the basis of our current knowledge is to risk mistaking association for causation. Very detailed analysis of small data sets (eg, 30 patients) may be used to generate hypotheses to test in new studies. However, conclusions from such analyses are data dependent in themselves, and placing too much reliance on the results might be misleading, particularly in the absence of any spontaneous or pharmacologically induced tissue recovery to identify which tissue actually is salvageable. An example of this type of error approximately 10 years ago was the overemphasis of the association between hemorrhagic transformation and cardio-embolic stroke. Then, it was incorrectly assumed that cardio-embolic stroke caused hemorrhagic transformation, because of an observed association, when in fact the size of the infarct rather than the mechanism per se was responsible.
In all cases, particularly in studies of DWI with PI so far, the sample size has been small. Our totals may overestimate the true number imaged, because some patients were included in several different publications. Small sample size makes the studies vulnerable to the play of chance in the mix of patients included and the study results. Examples of how profound an effect the play of chance can have can be found in the systematic review of early trials of thrombolysis for acute myocardial infarction (where in studies with sample sizes up to several hundred had diametrically opposite results) and in an exercise to demonstrate the effect of chance on small sample sizes (DICE [Don’t Ignore Chance Effects] therapy). Investigators should be encouraged to combine their existing individual patient data from different individual studies and participate in new multicenter studies whenever possible, thereby achieving much larger sample sizes and overcoming some of the case-mix problems outlined above. The numbers of patients required to obtain more robust data on the efficacy of DWI and/or PI in individual studies are not large, however. Only 84 patients would be required to show (with 90% power) that DWI demonstrated lesions in 90%, as opposed to CT showing lesions in 60% of patients. Finally, acutely ill patients are poor MR subjects: they are restless, they may be confused and unable to lie still for prolonged scan times, and observing their clinical state is difficult. Lying supine when the ability to protect the airway is impaired increases the risk of aspiration (roughly 50% of all stroke patients have acutely impaired swallowing reflex). DWI and PI have to genuinely provide information over and above that which is readily available on a plain CT, and which alters clinical management, and this information has to outweigh any loss of treatment efficacy resulting from added time delay to start of treatment, if their incorporation into patient management is to be justified. Such a balance has not yet been studied.

So, where do we go from here? Larger studies of DWI with PI are needed, with clear patient clinical details (inclusions, exclusions, and failures), blinding of analyses, valid clinical outcome measures, and comparison with a routine imaging such as CT so that an assessment of added clinical value can be made. Patient status during imaging should be assessed to identify adverse events, such as hypoxia, which could then be treated or avoided. DWI/PI should be evaluated in randomized trials of new or existing acute stroke treatments because that will continue to increase the knowledge pool on the practical application of complex imaging, and only then will we actually find out which tissue is still viable and recoverable and which is not. Then and only then will we know, when faced with an individual acute stroke patient, at a certain time after the stroke, with particular clinical features and appearances on DWI/PI, what the best clinical management will be.

Acknowledgments

Dr Keir received funding from the National Coordinating Centre for Health Technology Assessment, a branch of the National Health Service Research and Development body. The statements, opinions, and results in this article are those of the authors and do not reflect the policy or position of the NCCHTA.

References

6. Dr Keir received funding from the National Coordinating Centre for Health Technology Assessment, a branch of the National Health Service Research and Development body. The statements, opinions, and results in this article are those of the authors and do not reflect the policy or position of the NCCHTA.


Systematic Review of Diffusion and Perfusion Imaging in Acute Ischemic Stroke
S. L. Keir and J. M. Wardlaw

Stroke. 2000;31:2723-2731
doi: 10.1161/01.STR.31.11.2723

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/31/11/2723

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/