A supervised method for calculating perfusion/diffusion mismatch volume in acute ischemic stroke

Judy R. James\textsuperscript{a}, Karmen K. Yoder\textsuperscript{b}, Olaniyi Osuntokun\textsuperscript{c}, Andrew Kalnin\textsuperscript{b}, Askiel Bruno\textsuperscript{c}, Evan D. Morris\textsuperscript{a, b, *}

\textsuperscript{a}Department of Biomedical Engineering, Indiana University-Purdue University, (IUPUI) Indianapolis, IN, USA
\textsuperscript{b}Department of Radiology, Indiana University School of Medicine, Indianapolis, IN, USA
\textsuperscript{c}Department of Neurology, Indiana University School of Medicine, Indianapolis, IN, USA

Received 9 March 2005; accepted 18 May 2005

Abstract

Diffusion and perfusion (MR) imaging modalities identify overlapping but not identical areas of tissue as lesion following a stroke. It is thought that the ‘mismatch’ between modalities may represent tissue that could be recovered with proper (thrombolytic) treatment. We have designed a tool for semi-automated segmentation of the images and calculation of the mismatch volume. We present results from software phantoms and clinical data. Phantom results show our mismatch volume calculations are unbiased at realistic noise levels. Clinical data show that raters using our tool are consistent, fast (15 min per subject) and indistinguishable from an expert using manual segmentation.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: Stroke; Diffusion; Perfusion; Mismatch volume; Penumbra; Segmentation

1. Introduction

1.1. Stroke

Acute stroke (cerebral infarction) is caused by sudden oligemia (ischemia), which deprives brain cells of oxygen and nutrients. In the absence of blood flow, a pathophysiological chain of events begins that leads to cell death. There are reports that suggest that there is a level of cerebral blood flow
(20–30 mL/100 g/min) at which neurons stop functioning normally but have not yet undergone cell death [1]. Damage from ischemia might be reversible if the blood flow to the affected region could be restored to normal. Damaged tissue that remains viable but not functional (due to low blood flow) is referred to as the ‘ischemic penumbra’ [1], a term typically applied to the hypoperfused tissue surrounding the non-viable core of a stroke lesion [2,3]. The damaged cells in the penumbra can remain in a compromised state for several hours. Penumbral tissue, the main target of stroke therapy, can be saved if treated within a specified time frame (normally 3–6 h from the onset of stroke) [4–6]. The National Institute of Neurological Disorders and Stroke (NINDS) rtPA Stroke Study Group investigated the efficacy of thrombolytic therapy with recombinant tissue plasminogen activator (rtPA) administered within 3 h after stroke onset, and found that the rtPa therapy was beneficial for ischemic stroke patients [5].

Since CT and T2-weighted magnetic resonance (MR) images often cannot detect an infarct until 12–24 h after onset of ischemia, they are not useful for detecting early events in cerebral ischemia [7]. However, diffusion weighted (DW) and perfusion weighted (PW; dynamic susceptibility contrast [8]) MR imaging have been shown to detect acute stroke in early stages [9–12]. Diffusion weighted imaging (DWI) measures self-diffusion of water molecules in a tissue. During ischemia, it is believed that cell breakdown impedes movement of water molecules, resulting in increased DWI signal. Lesions appear bright (high signal) on DWI [13,14,16] (Fig. 1(A)). Perfusion weighted imaging (PWI) provides a measure of regional perfusion relative to the delivery of a contrast agent to a major blood vessel [14]. Perfusion is normally measured in terms of flow (milliliters per min per 100 g of brain tissue), volume (milliliters per 100 g of brain tissue) and mean transit time (s) [15]. Quantitative analysis of dynamic PW images yields three parametric maps: relative Cerebral Blood Flow (rCBF), relative Cerebral Blood Volume (rCBV) and Mean Transit Time (MTT) [13,16] (Fig. 1(B)). Stroke lesions are typically identifiable as regions having either low CBF or prolonged MTT. DW and PW images identify slightly different lesion extents, which may be useful in the characterization and therapeutic management of acute stroke.

Fig. 1. Stroke affected brain regions on DW (A) and PW images. Dynamic PW images (B) are modeled to create MTT map (C). Numbers define different areas of intensity in the images: (1) background, (2) normal tissue, (3) lesioned tissue.
1.2. Need to identify mismatch

It is assumed that a lesion on a PW image represents both the ischemic core and the penumbra, whereas the DW lesion is believed to represent the core only [15]. When the volume of an acute ischemic lesion on PW (Vol\(_{\text{PW Lesion}}\)) is greater than the volume of the lesion on DW images (Vol\(_{\text{DW Lesion}}\)), it is referred to as a ‘mismatch’ (MM). Strokes with a ‘mismatch’ generally tend to show larger evolution of infarct size on PW images with time if perfusion of the infarct region is not re-established. Since the MM volume may represent the penumbra, which could be salvageable via anti-clotting drugs such as rtPA, it would be useful to be able to calculate the volume of MM accurately [17–19]. If Vol\(_{\text{DW Lesion}}\) > Vol\(_{\text{PW Lesion}}\) it is referred to as ‘reverse mismatch’ [20]. Cases of reverse mismatch have also been reported. These usually involve patients who are scanned more than 6 h after the onset of stroke. Smaller apparent PWI lesions may be the result of re-perfusion of the infarct region before scanning. However, at this late time point (> 6 h post-infarct), there is probably already irreversible tissue damage to the stroke region. This could be one reason for a larger lesion in DWI than in PWI [21]. The current clinical focus on treatment is to enable stroke patients with a mismatch to receive timely treatment as specified by the NINDS study group. Therefore, there exists an interest amongst neurologists in developing faster, more efficient, and more reliable techniques to determine MM volume. Ultimately, MM volume could serve as a clinical indication for the initiation of rtPA, and could serve as an endpoint in the evaluations of the efficacy of thrombolytic therapy [4,17,20–23].

1.3. Previous attempts at obtaining lesion size on MR images

Much of the research dealing with the segmentation of infarcts on medical images has relied on manual outlining of the lesions [4,8,17,22–24]. Coutts et al. published a comparison of qualitative methods of MM assessment for the purpose of evaluating the ischemic penumbra [20]. The Coutts study employed raters to verify presence and size of penumbra based on either visual impression alone or with the aid of a manual tracing of the lesion on DW and MTT images. The size of the MM (if found) was assigned to deciles by the raters, and the study did not utilize an actual quantitative evaluation of lesion size. Kluytmans et al. [14] reported the development of a semi-automated technique for determination of lesion volume, which used a region-growing technique that required a rater to place one or more seed points manually. This semi-automated approach also required manual editing by the user to confirm the boundaries. Manual editing of images tends to be time-consuming, and in the development of our tool, we sought to keep user operations to a minimum.

1.4. Key elements of our segmentation technique

We developed a supervised segmentation algorithm, ‘Mismatch Volume Calculation’ (MMVolCal) for segmenting stroke lesions on both DW and PW images to obtain MM volume in acute stroke. The procedure is intentionally not completely automated in order to take efficient advantage of a trained rater’s ability to differentiate lesion from artifacts. A rater’s intervention is needed only at two decision points in our segmentation scheme: first, to identify a lesion from normal tissue; and second, to filter out apparent artifacts by use of size-matched filters. Here, we demonstrate that this technique is at least as accurate as completely manual lesion segmentation performed by an expert. MMVolCal is capable of segmenting stroke lesions reliably on DWI and PWI in less than 15 min. Unlike other studies, we measure
MM as a continuous variable. We anticipated that the semi-automated approach of MMVolCal would yield consistent results across raters.

We present results of five raters’ evaluations of synthetic phantoms and real patient data sets to demonstrate that our technique yields accurate, reproducible values for MM volumes, and to show that these MM values are statistically indistinguishable from manually segmented volumes. We propose that MMVolCal could be used in larger studies that require quantification of stroke lesions to evaluate MM volume.

2. Methods

2.1. Patient population

The present study was part of an imaging protocol for acute stroke approved by the Indiana University Institutional Review Board. The study was performed by identifying patients retrospectively who were diagnosed with ischemic stroke. Clinical histories, including approximate time of onset of stroke symptoms, were obtained from patient records. Patients met the following inclusion criteria: (1) well-defined time of onset of acute stroke symptoms (2) non-lacunar stroke visible on DW (3) absence of cerebral hemorrhage (4) visually apparent lesion on MTT images. 11 patients (7M, 4F, Age: 54.27 ± 13.58 y, body weight: 76.6 ± 12.16 kg) received MR scans of the head, including DW and PW imaging, at Wishard Memorial and Indiana University Hospitals (Indianapolis, Indiana). Patients were scanned within 68 h (mean ± s.d., 39 ± 29 h) of onset of stroke symptoms.

2.2. Imaging protocol

MR images were obtained on a 1.5T MR Scanner (Genesis Signa, GE Medical System). DWI was acquired with an axial isotropic single shot Spin-Echo-Echo-Planar imaging (SE-EPI) sequence with a diffusion sensitivity \( b = 1000 \text{s/mm}^2 \) (repetition time (TR) = 7000–10 000 ms, echo time (TE) minimized; twenty-six 5 mm slices, pixel size 0.9375 mm × 0.9375 mm, 22 × 22 cm field of view, 128 × 128 acquisition matrix reconstructed to 256 × 256 matrix). Scan time for DW imaging was approximately 48 s. PW images were acquired dynamically with an axial T\(^*\)\(_2\) Spin-Echo-Echo-Planar (SE-EPI) sequence (TR = 2200 ms; TE = 100 ms) before and after injection of an intravenous bolus of paramagnetic contrast agent, Gadodiamide [dose = 0.13± 0.02 mmol/kg]. Gadodiamide (Gd-DTPA; Omniscan; Mallinckrodt, Inc., Hazelwood, MO, USA) was injected using a power injector at 4 cc/s. Injection followed 15 s of baseline data acquisition. Ten to twelve 11 mm thick slices (pixel size 2.3475 mm × 2.3475 mm, no inter slice gap, 24 × 16 cm field of view, 128 × 64 acquisition matrix reconstructed to 128 × 128 matrix) were acquired. PWI acquisition produced 400–480 images in approximately 40 time points. Scan time for perfusion imaging was less than 2 min. The total imaging time for acquisition of all MR scans, including standard anatomical sequences, was between 30 and 40 min.

2.3. Parametric image generation for dynamic perfusion images

A dynamic series of PW images were post-processed to create parametric maps of rCBF, rCBV, and MTT maps using the perfusion analysis tool included in MEDx 3.4.2 (image processing and analysis
software developed by Sensor Systems Inc., Sterling, VA, USA). Initially, dynamic susceptibility curves produced due to the passage of paramagnetic contrast agent were converted to dynamic tracer concentration data [25]. The perfusion analysis tool can be set to automatically detect voxels that are most likely to be arteries [26]. The average signal from a few of these selected ‘arterial’ voxels was used as the arterial input function (AIF) [8,15,25,27]. The dynamic tissue concentration curves at each voxel were then deconvolved with the AIF using the singular value decomposition method proposed by Østergaard [27] to create parametric perfusion maps (MTT, rCBF, rCBV).

Slices at each time point were observed for patient head motion by means of a visual inspection of the data while played in a cineloop. If significant motion was detected in any slice at any time point, motion correction [28] was applied to all the time points in the given data set prior to parametric image generation. Since it has been reported (and observed by us) that MTT maps show a more distinct perfusion lesion border than any other maps [29,30], MTT maps were used in our analysis. The perfusion analysis tool failed to estimate MTT at some pixels in the image. An MTT value is calculated as rCBV/rCBF (central volume principle). Pixels with no flow in an rCBF map lead to failed estimation of MTT values in the corresponding map. These pixels in an MTT map are routinely assigned a non-zero error code by the perfusion analysis tool and required additional pre-processing on our part prior to segmentation.

2.4. Phantom image generation

To test the performance of MMVolCal at various noise levels, four phantom images, (phantom A, phantom B, phantom C, and phantom D) of both DW and MTT were created with different noise-to-signal (N/S) ratios (Fig. 2). Phantom A, for both DW and MTT, had no noise in the image. Phantom B was created with a noise level comparable to actual patient DW and MTT data. Phantoms C and D had progressively greater noise levels than phantom B. Phantom D was much noisier than any observed patient data. Phantoms were created with quasi-anatomical tissue regions representing three different tissue types: normal tissue, lesioned area, cerebrospinal fluid (CSF), and also background (BG). The lesion outlines were traced on a representative patient’s images (DW, MTT) by an expert rater (neuroradiologist). These lesion outlines were used to create the phantoms in each type of image (DW or MTT). The mean intensities of normal tissue, the lesion and the CSF were estimated from the representative patient data. The BG (exterior to the brain) was assigned a mean intensity and noise level similar to that of the CSF.

2.4.1. Diffusion weighted (DW) phantom images

The pixel dimensions of the DW phantoms were the same size and bit depth as the original images; (0.9375 × 0.9375 × 5.0039) mm^3; 26 slices; 256 × 256 (signed 16 bit) (Fig. 2). The mean intensities and zero-mean gaussian noise levels assigned to the tissue regions of the DW phantoms (Fig. 2) are given in Table 1. The true volume of the lesion on the noiseless DW phantom was 118.57 mL.

2.4.2. Mean transit time (MTT) phantom images

The pixel dimensions of the MTT phantoms were the same as the original MTT maps (2.35 × 2.35 × 11) mm^3; 11 slices; 128 × 128 (real 32 bit) (Fig. 2). In MTT phantom images, BG was assigned a value of zero. To emulate the behavior of the perfusion analysis tool in MEDx, CSF was assigned a non-zero error code (signifying a failed attempt to produce an MTT estimate from dynamic susceptibility data) in the MTT phantoms. In addition, we randomly assigned non-zero error codes to pixels throughout the
Fig. 2. Sample slices of phantom DW and PW (MTT) data created from outlines of actual patient scans with increasing noise levels. Outlines were generated for lesion, normal tissue, and cerebrospinal fluid (CSF). Phantoms were created at increasing noise levels (A-D: A, noiseless; D, noisiest). The noise distribution in the different regions of the image is described in Table 1. MTT slices have “drop-out” pixels (black spots on phantom MTT images), which were added manually at random to mimic failed estimation of perfusion parameters (CBF, CBV, MTT) by the commercial perfusion software.

Table 1
Mean ± s.d. of noise added to each region (normal, lesion, CSF) for DW and (MTT) phantoms

<table>
<thead>
<tr>
<th></th>
<th>DWI</th>
<th>MTT (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phantom A</td>
<td>200 ± 0</td>
<td>10 ± 0</td>
</tr>
<tr>
<td>Phantom B</td>
<td>200 ± 25</td>
<td>10 ± 2</td>
</tr>
<tr>
<td>Phantom C</td>
<td>200 ± 37.5</td>
<td>10 ± 4</td>
</tr>
<tr>
<td>Phantom D</td>
<td>200 ± 50</td>
<td>10 ± 8</td>
</tr>
<tr>
<td>Normal</td>
<td>480 ± 0</td>
<td>20 ± 0</td>
</tr>
<tr>
<td>Lesion</td>
<td>480 ± 55</td>
<td>20 ± 5</td>
</tr>
<tr>
<td>CSF and BG</td>
<td>480 ± 82.5</td>
<td>20 ± 10</td>
</tr>
<tr>
<td>CSF</td>
<td>480 ± 110</td>
<td>20 ± 20</td>
</tr>
</tbody>
</table>

image to simulate failure of MEDx to estimate MTT in voxels with undetectable responses to injection of contrast agent. As with DW, four MTT phantoms were created with different levels of additive noise (Fig. 2). Mean intensities and noise levels added to the tissue types are listed in Table 1. The volume of the lesion on the noiseless MTT phantom was 63.95 mL. The negative values produced in the phantom after addition of zero-mean gaussian noise were also assigned non-zero-error codes to make the phantoms resemble parametric images produced by MEDx.
2.5. Design strategy

Our objective was to build a tool, MMVolCal, that would segment stroke lesions on DW and PW images semi-automatically and calculate the MM volume rapidly, with minimal input from a trained rater.

2.5.1. Interface design

MMVolCal was implemented as a Tcl/Tk script in MEDx 3.4.2 (using MEDx image script functions) with an interactive graphical user interface (GUI) (Appendix A). This tool must be run from within MEDx 3.4.2 or higher. The GUI was arranged vertically as a single column of buttons and entry widgets. Each processing step was clearly numbered and colored on the GUI to enforce uniformity of use by the raters. An outline of the segmented lesion overlaid on the original image was provided as a feedback to the user after the completion of each processing step. The user was allowed to ‘UNDO’ and repeat a completed processing step (e.g., thresholding or filtering) if he/she was unsatisfied with the preview of the lesion outline. To limit the variability of the result, only the most recent step in the chain (Fig. 3) was UNDO-able.

2.5.2. Quality control

A formatted report file in html was created from each analysis to facilitate a subsequent quality control examination. This allowed the investigators to check for mis-operation of the tool that might invalidate a result. The report files generated for each case are viewable in any browser and do not require MEDx for viewing. The report recorded each selection and operation performed by the user, as well as the final segmentation of the lesion overlaid on the original images. Regions of interest (ROI) placed by the rater.
to identify lesion areas were also recorded. Each report contained sufficient information to completely reconstruct an analysis (Appendix B).

2.6. **Image processing/segmentation**

DW images were processed directly following acquisition. The acquired dynamic perfusion images were processed to create perfusion maps (rCBF, rCBV, MTT). MTT maps were chosen for analysis. The DW and MTT images were processed as described in the flow chart (Figs. 3 and 4). Phantom data were processed identically to the patient data. Prior to segmentation of the images, patient data were deidentified. Raters were not allowed access to any patient images or clinical data that could influence their assessment of lesion boundaries.

2.7. **Program implementation**

1. **Loading and preparing data for processing:**
   (a) DW images did not require any pre-processing prior to segmentation.
   (b) Pre-processing of the MTT data:

   All negative values in the image were set to zero. The non-zero error codes produced in the MTT maps were assigned maximum MTT values on the assumption that poor perfusion (large MTT, low or no flow) caused the error in those pixels (Fig. 5, steps 2–3). Although the lesion is visible at step 3, the MTT maps were still very noisy and did not provide a clear lesion boundary. Therefore, the MTT maps were also smoothed and de-speckled slice by slice using a 2D median filter (3 x 3 kernel) (Fig. 4(B)-1'; Fig. 5 step 4).

   2. **Placement of ROIs on normal and lesion areas:** Raters were prompted to position 2D ROI (single boxes not connected to each other across axial slices) on what they considered to be normal tissue in the image. The user placed the 2D ROI boxes on the first, middle, and the last slice of the image. Next, users were prompted to position a 3D ROI (boxes linked together on multiple slices) to generously encompass the stroke lesion (Fig. 4(A) and (B), step 2). The raters were asked to indicate the first and last slices of the image on which they observed a lesion. Both the normal and lesion ROIs could be adjusted manually or ‘UNDOne’. The 3D ROIs could be adjusted simultaneously on all slices. When positioning the 3D ROI over a lesion, the rater could include normal tissue, lesioned tissue and BG in the ROI (Fig. 4(A)-2, (B)-2). Once the 3D ROI was positioned appropriately around the lesion by the rater, all tissue outside the 3D ROI was erased (Fig. 4(A)-2', (B)-2'). Some residual capture of artifacts within the defined lesion volume was inevitable. Final ROIs were saved in appropriate files and directories for further analysis or review.

   3. **Thresholding:** A histogram analysis was performed to measure the intensity distribution in the normal and lesion ROIs. The maximum intensity detected in the 2D ‘normal’ ROIs was automatically applied as the minimum threshold value for the entire image on the assumption that the lesion is characterized by intensity values greater than normal. The rater was allowed to adjust the threshold numerically based on the real-time preview of the border of the lesion at a given threshold (Fig. 4(A)-4, (B)-4).

   4. **Morphological filtering:** After thresholding, some bright spots disconnected from the boundary of the main lesion may remain, which, at this point in the process, are still counted toward the total lesion volume, as illustrated in the DW image depicted in Fig. 4(A)-3, 4. However, these disconnected segments may in fact be artifacts (not lesion). These putative artifacts within the defined lesion volume
Fig. 4. Sample DW (A) and PW (MTT) (B) images processed by MMVolCal. Steps in this figure correspond to steps in the flowchart in Fig. 3. DW and MTT data shown here do not correspond to the same anatomical space.
Fig. 5. Pre-processing of MTT maps. (1) Original MTT, (2) Binary mask of non-zero error codes are created, (3) Binary mask is assigned high MTT value and mask is added to original image, and (4) Median filtering creates the final MTT image. This pre-processing (steps 1–4) is performed prior to MTT segmentation, and aids in revealing lesion boundaries more clearly.

were eliminated by a morphological opening operation (erosion followed by dilation) without altering the size of the putative main lesion. The opening process tends to smooth the contour of the lesion, break narrow connections between large entities and eliminate small (i.e., smaller than the filter kernel) disconnected bright spots (Fig. 4 (A)-5,6, (B)-5,6).

2D morphological filters were used to remove artifacts in DW and MTT images. The choice of filter was dependent on the shape and contour of the thresholded image. The GUI provided a choice of filters to the rater, with instructions to “use the smallest possible filter to remove suspected artifacts”. The kernel sizes of the filters were: small filter (3 × 3 pixels), medium filter (5 × 5 pixels), large filter (7 × 7 pixels). A no filter option was also available to the rater. Filtering was ‘UNDO-able’.

(5) Lesion volume calculation: The number of voxels within the final outline of the lesion (Fig. 4(A)-6, (B)-6) was counted and multiplied by the appropriate voxel volume to determine the overall lesion volume for both DW and MTT lesions (Vol(DW Lesion) and Vol(PW Lesion), respectively). MM volume was calculated as [Vol(PW Lesion) − Vol(DW Lesion)].

2.8. Statistical analysis:

Repeated measures ANOVA was used to (1), test the internal consistency of MMVolCal by testing for effects of rater on MTT, DW and MM volumes, (2) determine if volumes calculated by raters using MMVolCal were different from lesion volumes obtained from manual outlining, and (3) test the effect of noise level in images on the accuracy of MMVolCal.

Pearson’s correlation coefficient was used to determine how well results from individual raters using MMVolCal corresponded to the lesion volumes obtained from manual outlining.

3. Results

3.1. Performance of MMVolCal on synthetic data

Sample slices of phantom DW and MTT images after segmentation are shown in Fig. 6, with the final lesion outlines superimposed on the original slices. The white outline around the lesion (Fig. 6 (A–D)) represents the boundary of the lesion defined by MMVolCal. In Fig. 6 (C–D), in addition to the main lesion boundary, there were also smaller outlines within the lesion. These inner outlines represent voxels
Fig. 6. Segmentation of lesions on synthetic (phantom) DW and MTT images depicted in Fig. 2. The white outlines on the images, both around and inside the lesions, represent the final lesion outlines produced by MMVolCal, from which lesion volumes are calculated.

that are not included in the overall lesion volume. Images with higher noise levels yield smaller apparent lesion volumes. For the phantom image data set, this effect of noise is less severe in high-noise MTT images than in high-noise DW images.

The true volume of the lesion on the noiseless DW and MTT phantoms was 118.57 and 63.95 mL, respectively. Thus, the true MM volume is $-54.62$ mL. The mean MTT lesion volume calculated across phantoms A–D by the five raters using MMVolCal was $45.75 \pm 4.63$ mL. The mean DW phantom lesion volume across the four phantoms with the MMVolCal was $104.89 \pm 12.59$ mL. The mean MM volumes generated from the four phantoms from the five raters was $-54.14 \pm 8.31$ mL (Fig. 7).

There was no significant effect of noise level on lesion volumes obtained for DW phantom images. As noise levels increased, MM volumes tended to decrease. This decrease parallels a progressive decrease in measured phantom DWI lesion volumes, resulting from an increasing number of noise-related artifacts that are being excluded from the DWI lesion. In other words, the error and bias in the phantom MM volume appear to be more affected by the noise in DW images than by the noise in MTT images. As we will describe below, the opposite appears to be the case for the real data set analyzed with MMVolCal.

3.2. Performance of MMVolCal on patient data

There was a wide range of lesion volumes and MM volumes across the 11 patients. The MTT lesion volumes varied from $3.64 \pm 2.56$ to $91.05 \pm 42.76$ mL (mean $\pm$ s.d. determined from MMVolCal volumes).
Fig. 7. Performance of MMVolCal on phantom data with noise levels A-D (please see Fig. 2 and Table 1). The mean PW (MTT), DW and MM volumes obtained via MMVolCal from five raters are displayed. Data are presented as mean ± s.d. The first triad of bars at the very left of the x-axis indicates the true volumes of PW (MTT), DW and MM lesions, respectively. These true volumes were defined in the noiseless phantoms.

Fig. 8. MM volumes calculated with MMVolCal demonstrate a good correspondence to the ‘gold standard’ manually traced volumes. MM volumes calculated with MMVolCal from the five raters (y-axis, raters are denoted by different symbols) are plotted against the ‘gold standard’ MM volumes (x-axis). The ideal relationship (perfect correspondence) between the MMVolCal volumes and the ‘gold standard’ volumes is plotted as a dotted line passing through the origin. The points cluster closely around this ideal performance.

The DWI lesion volumes ranged from $1.47 \pm 0.92$ to $166.63 \pm 68.43$ mL. MM volumes ranged from $-133.91 \pm 70.43$ to $4.22 \pm 33.51$ mL. The data showed a higher variability in MTT lesion volumes than in DWI lesion volumes. Thus, the variability in the final MM volumes derived with MMVolCal is likely related to variability in the MTT lesion volumes.
Table 2
Pearson’s $r$-values for correlations between the respective DW, MTT, and MM volumes calculated with MMVolCal and the manually outlined lesion volumes (‘gold standard’). Only real subject lesion data were included in the correlational analyses.

<table>
<thead>
<tr>
<th></th>
<th>Rater 1</th>
<th>Rater 2</th>
<th>Rater 3</th>
<th>Rater 4</th>
<th>Rater 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>DW</td>
<td>Gold</td>
<td>0.982**</td>
<td>0.961**</td>
<td>0.942**</td>
<td>0.974**</td>
</tr>
<tr>
<td></td>
<td>standard</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTT</td>
<td>Gold</td>
<td>0.919**</td>
<td>0.695*</td>
<td>0.812**</td>
<td>0.619*</td>
</tr>
<tr>
<td></td>
<td>standard</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MM</td>
<td>Gold</td>
<td>0.836**</td>
<td>0.845**</td>
<td>0.676*</td>
<td>0.844**</td>
</tr>
<tr>
<td></td>
<td>standard</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

** $p < 0.01$.
* $p < 0.05$.

Fig. 9. Effect of contrast agent dose per body weight on the noise/signal ratio (s.d./mean signal intensity) from normal brain tissue in PW images. Higher doses per body weight of contrast agent increase the PW image quality.

The MTT, DWI, and MM volumes derived from MMVolCal were compared to manually traced lesion volumes completed by a neuroradiologist. With regard to patient data, manually traced volumes were considered to be the ‘gold standard’ to which results from any other method are to be compared. Fig. 8, which plots volumes from MMVolCal against ‘gold standard’ results, shows that MMVolCal yields results that are extremely consistent with those of the ‘gold standard’. Repeated measures ANOVA showed that MTT, DWI and MM volumes derived from MMVolCal were not statistically different from the respective ‘gold standard’ volumes. Furthermore, there was no effect of rater on MTT, DWI, and MM volumes calculated with MMVolCal. Table 2 demonstrates that MTT, DW and MM volumes calculated by individual raters using MMVolCal are highly correlated with the ‘gold standard’ (i.e., manually traced volumes).
3.3. Effect of administered contrast agent dose on noise in PW images

In order to understand the possible sources of error, we examined the effect of contrast agent on the N/S ratio in PW images. Only PW images (and not DW) depend on the injection of contrast. Each patient’s mean tissue activity curve was examined after perfusion analysis. The time point at which the curve peaks was noted. The s.d. of the intensity value in the image (noise) divided by the mean in the brain tissue (signal; excluding background) was estimated at the selected time point for each patient and plotted versus dose/ body weight. Although not significant, Fig. 9 shows a trend toward a negative correlation ($r^2 = 0.2031; p > 0.05$) between administered dose per body weight and the noise in the MTT images. That is, noise decreases with greater dose per body weight of contrast agent.

4. Discussion

MM volumes calculated from DW and PW images may be useful as a clinical indicator of whether or not to administer drugs to acute stroke patients. MM is commonly determined by visual inspection and/or manual segmentation of a lesion outline on DW and PW (i.e., MTT) images. However, both visual inspection and manual segmentation are very time-consuming. In a clinical setting, MM would ideally be determined within 3–6 h of onset of stroke symptoms. MMVolCal accurately segments stroke lesions on DW and MTT images and calculates MM volume in less than 15 min.

MM volumes calculated with MMVolCal were overestimated at high noise levels in the phantom data. To resolve this issue, we recommend pre-processing the DW images by smoothing the noise in DW images, as is currently done with MTT images.

The accuracy and reliability of MMVolCal were tested by comparing the DWI lesion, MTT lesion, and MM volumes from patient data calculated by MMVolCal to the results from the ‘gold standard’ manual segmentation by a neuroradiologist. There was quite good agreement between the volumes calculated by the MMVolCal and the ‘gold standard’ volumes (Fig. 8).

Inter-rater variability in lesion volumes calculated with MMVolCal could be attributed to differences in the morphological filters chosen by the raters (see example in Appendix B). Apparently, some raters judged possible artifacts to be part of the lesion, whereas others did not. Nevertheless, it should be noted that the lesion volumes were not statistically different among raters.

Much of the apparent variability in MM volumes obtained from patient data could be attributed to the variability in the MTT lesion volumes. MTT images were much noisier than DW images, which contributed to the higher variance in MTT lesion volumes generated by MMVolCal. We attribute the noise in the MTT images to motion artifacts and/or inadequate administration of contrast agent dose. Since the PW maps were corrected for motion, an insufficient dose of contrast agent may cause excessive noise in the MTT maps. Insufficient dosing could be caused by specific features of the bolus injection, including amount of contrast agent injected or the injection rate. In addition, the protocol used here administered a standard volume of contrast agent, regardless of body weight. We speculate that if all patients are given either a specific dose of contrast agent per body weight or a consistent “double dose” of a standard volume of contrast agent, the noise in MTT images could be reduced, and the noise-to-signal ratio could be decreased. In turn, this would improve PWI lesion definition and thus, reduce the variability in the calculation of MM volume.
The majority of patients in this study had reverse MM (or, negative MM). As mentioned in the introduction, a positive MM most likely would be observed if the patients were scanned within 6 h of onset of stroke symptoms. The patients included in our retrospective study were scanned later than the desired cutoff of 6 h post-infarct ($39 \pm 29$ h). We believe that a natural re-perfusion of the infarct region may have occurred before these patients were scanned, thus decreasing the MTT volumes and creating reverse MMs. It must be noted, of course, that MMVolCal will perform equally well whether MM volume is positive or negative.

5. Conclusion

We have developed a semi-automated method (MMVolCal) to quantitate stroke lesion volumes and calculate stroke mismatch volumes accurately, reliably, and quickly. MMVolCal could prove to be a useful tool for several applications, including assisting in clinical decisions about thrombolytic therapy and providing a reproducible and robust method to quantify MM in single site and multi-center research studies.

6. Summary

Perfusion weighted (PW) and diffusion weighted (DW) magnetic resonance imaging scans (MRs) capture slightly different aspects of brain lesions that occur as a result of acute stroke. If the perfusion lesion is larger than the diffusion lesion, the difference in volume between the lesions is referred to as ‘Mismatch’ volume. Mismatch lesion volume is thought to represent damaged tissue that can be salvaged with thrombolytic therapy (the ‘penumbra’), but only if such treatment is administered within hours after stroke onset. Accurate and reliable determination of mismatch volume would be extremely valuable for both clinical and research settings.

Unfortunately, the current methods to determine lesion volume typically involve manual tracing or editing of the lesion on MR images. Manual methods to determine lesion volumes are usually time-consuming, not fully quantitative, and are subject to great variability across raters (and clinical sites). A semi-automated method to quickly, accurately and reproducibly calculate mismatch volume could prove to be an extremely useful technique. Such a method could aid in the diagnostic process for rapidly determining whether or not thrombolytic therapy is a viable option. Reliable determination of mismatch volume would also be an important endpoint for research studies, such as ones seeking to establish the relationship between mismatch and the penumbra.

We present the design, implementation, and testing of a new software tool, ‘Mismatch Volume Calculation Tool’ (MMVolCal), that quickly and reliably segments stroke lesions on PW and DW MR scans with only minimal intervention by a trained rater. Five raters tested MMVolCal on four phantom data sets (of varying noise levels) and 11 acute stroke patients. Performance on phantom images determined that the tool was highly accurate. Manual outlining of lesions was performed on the clinical data by an expert neuroradiologist (the ‘gold standard’) to which the performance of MMVolCal was compared. Results with MMVolCal were highly reproducible across raters, and all raters were indistinguishable from the gold standard. In addition, raters using MMVolCal required less than 15 min per patient data set to complete an analysis. This efficiency is ideal when results are required quickly, or when many subjects
Fig. A1.
are involved in a large research study that requires analysis of mismatch volume. The ease of use and reproducibility of MMVolCal should make it useful for further investigation of the exact meaning and utility of the mismatch volume within a single clinical setting or across multiple research sites.

Acknowledgements

The authors would like to thank Mrs. Erin DeBrota, Drs. Letisha K. Ali, Yonghua Xu, and Kenneth VanOwen, who served as raters. We would also like to thank Marc Normandin for his technical assistance with image preparation.

Appendix A.

Mismatch Volume Calculation Tool (MMVolCal) interface (Fig. A1).
Appendix B.

Sample report pages from MMVolCal. The output for two separate raters is shown for one patient’s DWI data set (Fig. B1).

References


Judy R. James received her MS in Biomedical Engineering (2004) from Indiana University–Purdue University (IUPUI), Indianapolis, IN. She is currently pursuing her Ph.D. from Purdue University, Lafayette, Indiana in the Department of Health Science with a specialization in Medical Physics.

Judy has been working as a research assistant in the Department of Radiology, Indiana University–School of Medicine, IN since her MS. Her research interest includes image processing (morphology, segmentation, registration etc), visualization and analysis for various medical applications. She is currently working on NMR spectroscopy and imaging of liver models.

Karmen Yoder received her Ph.D. in Medical Neurobiology with a minor in pharmacology from the Indiana University School of Medicine (IUSM, Indianapolis, IN) in 2002. From 1999 to 2001 Dr. Yoder was the recipient of a Scottish Rite Schizophrenia Research Fellowship. From 2002 to 2004, Dr. Yoder received postdoctoral training in departments of Psychiatry and Radiology at IUSM, concentrating on radioligand PET methods and modeling techniques. Currently, she is an Assistant Research Professor in the Department of Radiology at IUSM. The focus of Dr. Yoder’s research is to employ functional imaging techniques such as PET and MR to document disease processes in the fields of psychiatry and neurology.

Olaniyi Osuntokun is a member of the Royal College of Physicians (part 1) and received the Masters of Science in Clinical Neurosciences (with distinction) from the Institute of Neurology, Queen’s Square, University of London. There he received the first prize for best research paper in clinical neuroscience in 1999. He was a resident in neurology at Hope Hospital in Salford, UK from 1996 to 1999. Since 1999 he has been a resident in the combined Psychiatry and Neurology program at the Indiana University School of Medicine.

Andrew Kalnin is an Associate Professor in the Department of Radiology at the Indiana University School of Medicine. He completed his medical training, Radiology Residency and Fellowship in Neuroradiology at the New York University School of
Medicine. His primary research interests are in functional MRI, including neurosurgical navigation and other clinical applications of fMRI, as well as diffusion tensor imaging.

**Askiel Bruno** is Professor of Neurology and Director of Stroke Program at Indiana University School of Medicine. Dr. Bruno received his MD degree from New York University and completed his Neurology Residency at the West Los Angeles Veterans Affairs Medical Center. Dr. Bruno also completed a Stroke Fellowship at the University of Iowa College of Medicine. Dr. Bruno serves as a reviewer for several popular medical journals and is an active member of major neurological and vascular international societies. His primary research interests include acute stroke treatment and prevention of recurrent stroke. He serves as an advisor on multiple randomized clinical stroke trials.

**Evan Morris** has been working in functional imaging, primarily modeling dynamic PET and MR data, since completing the Ph.D. in Biomedical Engineering at Case Western Reserve University in 1991. He worked on kinetic modeling of the dopamine system for various research projects while a member of the PET groups at Massachusetts General Hospital and National Institute on Drug Abuse in the 1990s. Working at Sensor Systems, Inc., he was the primary designer and engineer of a module for producing Perfusion maps from contrast enhanced dynamic susceptibility change MR data in the MEDx image analysis software package. He is presently an assistant professor of Biomedical Engineering and Radiology at Indiana University, Purdue University, Indianapolis and his research interests include using PET to estimate temporal patterns of neurotransmitter fluctuation in the brain in response to drugs as well as algorithms for direct reconstruction of parametric images from PET sinograms.