

ORIGINAL ARTICLE

Quantitative measurements of relative fluid-attenuated inversion recovery (FLAIR) signal intensities in acute stroke for the prediction of time from symptom onset

Bastian Cheng¹, Mathias Brinkmann¹, Nils D Forkert², Andras Treszl³, Martin Ebinger⁴, Martin Köhrmann⁵, Ona Wu⁶, Dong-Wha Kang⁷, David S Liebeskind⁸, Thomas Tourdias⁹, Oliver C Singer¹⁰, Soren Christensen¹¹, Marie Luby¹², Steven Warach¹², Jens Fiehler¹³, Jochen B Fiebach⁴, Christian Gerloff¹ and Götz Thomalla¹, for the STIR and VISTA Imaging Investigators¹⁴

In acute stroke magnetic resonance imaging, a ‘mismatch’ between visibility of an ischemic lesion on diffusion-weighted imaging (DWI) and missing corresponding parenchymal hyperintensities on fluid-attenuated inversion recovery (FLAIR) data sets was shown to identify patients with time from symptom onset ≤ 4.5 hours with high specificity. However, moderate sensitivity and suboptimal interpreter agreement are limitations of a visual rating of FLAIR lesion visibility. We tested refined image analysis methods in patients included in the previously published PREFLAIR study using refined visual analysis and quantitative measurements of relative FLAIR signal intensity (rSI) from a three-dimensional, segmented stroke lesion volume. A total of 399 patients were included. The rSI of FLAIR lesions showed a moderate correlation with time from symptom onset ($r = 0.382$, $P < 0.001$). A FLAIR rSI threshold of < 1.0721 predicted symptom onset ≤ 4.5 hours with slightly increased specificity (0.85 versus 0.78) but also slightly decreased sensitivity (0.47 versus 0.58) as compared with visual analysis. Refined visual analysis differentiating between ‘subtle’ and ‘obvious’ FLAIR hyperintensities and classification and regression tree algorithms combining information from visual and quantitative analysis also did not improve diagnostic accuracy. Our results raise doubts whether the prediction of stroke onset time by visual image judgment can be improved by quantitative rSI measurements.

Journal of Cerebral Blood Flow & Metabolism (2013) **33**, 76–84; doi:10.1038/jcbfm.2012.129; published online 10 October 2012

Keywords: acute stroke; diffusion-weighted imaging; fluid-attenuated inversion recovery; magnetic resonance imaging

INTRODUCTION

In acute ischemic stroke, treatment decisions are based on the time since symptom onset as a key variable. This essential information is, however, missing in as many as 25% of patients with stroke onset during sleep.^{1,2} The ‘stroke patient who woke up’³ is consequently precluded from thrombolysis in accordance with the international treatment guidelines.^{4,5} Recently, evolution of signal intensities in fluid-attenuated inversion recovery (FLAIR) sequences of magnetic resonance imaging (MRI) in acute stroke has attracted increased attention as a potential surrogate marker for time since stroke onset.^{6–8} Specifically, a visual mismatch between diffusion-weighted (DWI) and FLAIR MRI has been shown to identify patients likely to be within a time window of 4.5 hours.⁹ Although these findings show the specificity of a DWI/FLAIR mismatch as a ‘tissue clock’ to identify patients

potentially eligible for thrombolytic therapy, some limitations of the method remain. Agreement between observers as to the detection of a corresponding FLAIR lesion based on visual assessment varied to some extent in different studies from moderate to almost perfect ($\kappa = 0.65–0.97$).^{7,9,10} Also, whereas specificity and positive predictive value (PPV) in predicting ischemic lesions at < 4.5 hours were high, sensitivity and negative predictive value (NPV) remained low in previous studies (specificity: 0.78–0.90; PPV: 0.83–0.93; sensitivity: 0.38–0.74; NPV: 0.54–0.70).^{6,8,10} As a consequence, a considerable number of patients who are in a time window of ≤ 4.5 hours and thus may benefit from treatment with thrombolysis are missed by this approach.

Relying on qualitative visual assessment as a rating method of FLAIR images introduces variability and may contribute to low sensitivity and NPV. Judgment of lesion visibility may be

¹Klinik und Poliklinik für Neurologie, Kopf- und Neurozentrum, Universitätsklinikum Hamburg-Eppendorf, Hamburg, Germany; ²Institut für Computational Neuroscience, Universitätsklinikum Hamburg-Eppendorf, Hamburg, Germany; ³Institut für Medizinische Biometrie und Epidemiologie, Universitätsklinikum Hamburg-Eppendorf, Hamburg, Germany; ⁴Centrum für Schlaganfallforschung Berlin, Charité-Universitätsmedizin Berlin, Berlin, Germany; ⁵Klinik für Neurologie, Universität Erlangen-Nürnberg, Erlangen, Germany; ⁶Athinoula A Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, USA; ⁷Department of Neurology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, South Korea; ⁸Department of Neurology, University of California, Oakland, California, USA; ⁹University de Bordeaux, CHU de Bordeaux, Service de Neuroimagerie Diagnostique de Thérapeutique, Place Amélie Raba Léon, Bordeaux, France; ¹⁰Kliniken für Neuroradiologie, Universitätsklinikum Erlangen-Nürnberg, Erlangen, Germany; ¹¹Department of Neurology, Royal Melbourne Hospital, University of Melbourne, Parkville, Victoria, Australia; ¹²National Institute of Neurological Disorders and Stroke, Bethesda, Maryland, USA and ¹³Klinik und Poliklinik für Neuroradiologische Diagnostik und Intervention, Universitätsklinikum Hamburg-Eppendorf, Hamburg, Germany. Correspondence: Bastian Cheng, Klinik und Poliklinik für Neurologie, Kopf- und Neurozentrum, Universitätsklinikum Hamburg-Eppendorf, Martinistr. 52, Hamburg 20246, Germany. E-mail: b.cheng@uke.uni-hamburg.de

This work was supported by the National Institute of Neurological Disorders and Stroke (NINDS), and also by the European Union’s Seventh Framework Programme grant agreement No. 202213 and No. 223153 (European Stroke Network), the Volkswagen Foundation, and the Deutsche Forschungsgemeinschaft.

¹⁴See Appendix.

Received 2 July 2012; revised 11 August 2012; accepted 14 August 2012; published online 10 October 2012

influenced by external factors such as rater experience or differences regarding the window-level setup used when displaying the images, contributing to only moderate interobserver agreement. To overcome these limitations, quantitative measurement of relative signal intensities (rSIs) on FLAIR has been proposed, although yielding contradictory results.^{6,7}

We studied the value of FLAIR rSI measurement as a surrogate marker of acute ischemic brain lesion age. We hypothesized that adding information from quantitative FLAIR analysis to visual analysis would improve sensitivity and NPV of DWI and FLAIR for the identification of patients within 4.5 hours of symptom onset. To this end, we first identify and describe a group of patients with 'subtle' FLAIR lesions to be distinguished from the previously described groups with 'no' or 'obvious' corresponding FLAIR lesions by visual inspection. Second, we analyze FLAIR rSI within the DWI lesion using a three-dimensional segmentation approach. Finally, we evaluate the potential value of alternative classification algorithms incorporating information from quantitative FLAIR analysis to identify patients within the 4.5 hour time window.

PATIENTS AND METHODS

Patients

We analyzed MRI and clinical data from patients included in the PRE-FLAIR study (PREdictive value of FLAIR and DWI for the identification of acute ischemic stroke patients ≤ 3 and ≤ 4.5 hours of symptom onset—a multi-center study). PRE-FLAIR was a multicenter observational study evaluating the combined use of FLAIR and DWI for the identification of patients with acute ischemic stroke within 4.5 hours of symptom onset.¹¹ All patients were studied within 12 hours of witnessed stroke onset. Patient age, gender, side of ischemic lesion, time from symptom onset to MRI, and severity of neurologic deficit on admission assessed by the NIHSS (National Institutes of Health Stroke Scale) were recorded. Details of imaging parameters and clinical characteristics were reported in the original publication.¹¹ For this secondary analysis, we selected the group of DWI-positive patients previously included in the PRE-FLAIR study ($n = 516$). No further restriction to stroke lesion territory or lesion size was applied. Patients were excluded in whom postprocessing and quantitative analysis of FLAIR images as described below was not possible because of technical reasons or insufficient image quality. This study has been approved by the local ethics committee (Hamburg, AZ M51/2000). Either written or verbal informed consent was obtained for all patients, as required by local legislation.

Qualitative Analysis of Acute Ischemic Lesion Visibility on Fluid-Attenuated Inversion Recovery

The methods and results from visual rating according to the original study protocol have been reported previously.¹¹ Briefly described, two independent raters (BC and ME), masked to clinical information, judged whether and in which region they were able to identify FLAIR lesions that correspond to acute lesions on DWI. For the current analysis, visible corresponding FLAIR lesions were further subdivided into 'subtle' or 'obvious' (Figure 1). This judgment was based on the subjective evaluation of the corresponding parenchymal hyperintensities in FLAIR. Consensus was obtained in cases of conflicting judgment. Additionally, one observer (BC) assessed for the presence of leukoaraiosis on FLAIR images using an adapted scale of Fazekas and Schmidt.¹²

Quantitative Analysis of Fluid-Attenuated Inversion Recovery Signal Intensity Ratio

For FLAIR quantitative rSI measurements, DWI, and FLAIR images were analyzed using the in-house-developed software tool AnToNla (Analysis Tool for Neuro Image Data, Institute for Computational Neuroscience, Hamburg, Germany). For quantitative

assessment of FLAIR signal intensity ratios, the FLAIR data set of each patient was registered to the corresponding DWI data set using a rigid transformation, maximization of the mutual information, and linear interpolation. The rigid transformation was used because affine or even nonlinear differences were not expected. After registration, the lesion was segmented semiautomatically in each DWI data set: first, a volume of interest (VOI) surrounding the DWI lesion ('lesion VOI') with a generous margin at each affected slice was manually drawn. A second VOI was then drawn roughly covering the unaffected hemisphere, avoiding sulci and ventricles. This second VOI was then used to calculate the mean (μ) and standard deviation (σ) of the signal intensities in the DWI data set. These parameters were used for a refinement of the lesion VOI by rejecting all voxels within this VOI exhibiting a signal intensity less than $\mu + 2\sigma$. The retained voxels constituted a new refined three-dimensional stroke lesion VOI ('refined lesion VOI'). This 'refined lesion VOI' was projected onto the contralateral hemisphere ('mirror VOI'). Therefore, the interhemispheric fissure was manually defined by straight lines in two axial slices as distant as possible. These two lines were then used to define a three-dimensional plane, which is used to mirror the corresponding refined DWI lesion VOI onto the contralateral hemisphere (Figure 2). All VOIs were visually checked and corrected if necessary, e.g., in case of inclusion lateral ventricles in the 'mirror VOI'. Additionally, we checked for the VOI not to include distal hyperintense vessels in sulci adjacent to stroke lesions.

After this, the two VOIs were used to calculate the mean signal intensities (SIs) within the 'refined lesion VOI' and the 'mirror VOI' using the registered FLAIR data set. Finally, the FLAIR SI ratio (rSI) was calculated by ($rSI = SI \text{ 'lesion VOI'}/SI \text{ 'mirror VOI'}$). Lesion volume was measured from the DWI 'lesion VOI'.

Statistical Analysis

Interobserver agreement for the qualitative rating of FLAIR lesion visibility was calculated. We tested for a linear trend in FLAIR rSI regarding the groups of visually rated FLAIR lesion visibility (no lesion, 'subtle' lesion, and 'obvious' lesion). Group comparisons depending on the visibility of FLAIR lesions were performed using a one-way analysis of variance. A Bonferroni *post hoc* test was applied to detect significant group differences. The Pearson's χ^2 -test was used for group comparison of categorical variables. Groups were also compared regarding each area under the curve of FLAIR rSI for the identification of patients with time since symptom onset ≤ 4.5 hours. An optimal cutoff value for FLAIR rSI was determined from the receiver operating characteristic (ROC) curve including all patients by calculating Youden's indices. The correlation between FLAIR rSI and time since symptom onset was calculated using Pearson's correlation coefficient. For further analysis, a decision tree was created using the CART (classification and regression tree) algorithm. CART is a recursive partitioning method that explores the structures of a set of data and visualizes decision rules for predicting a categorical outcome.¹³ Binary splits ('nodes') are made on the predictor variables that best differentiate the outcome variable. We used classification inference trees within the 'ctree' function (available in the R package 'party'), which bases its node splitting on statistical tests.¹⁴ In our model, time from symptom onset (\leq or > 4.5 hours) was chosen as the binary outcome variable, whereas data from visual rating of FLAIR signal intensity, FLAIR rSI, and acute DWI lesion volume were used as predictor variables.

Sensitivity, specificity, NPV, and PPV for the identification of patients with stroke onset ≤ 4.5 hours were calculated based on visual analysis differing in the classification of patients with 'subtle' lesions: in model 1, subtle lesions were judged as visible (i.e., 'FLAIR-positive'). This 'liberal' rating corresponds to the rating in the original PRE-FLAIR study. In model 2, subtle lesions were judged as absent (i.e., 'FLAIR-negative') as a more 'conservative' rating. Predictive values were also calculated for the best FLAIR rSI

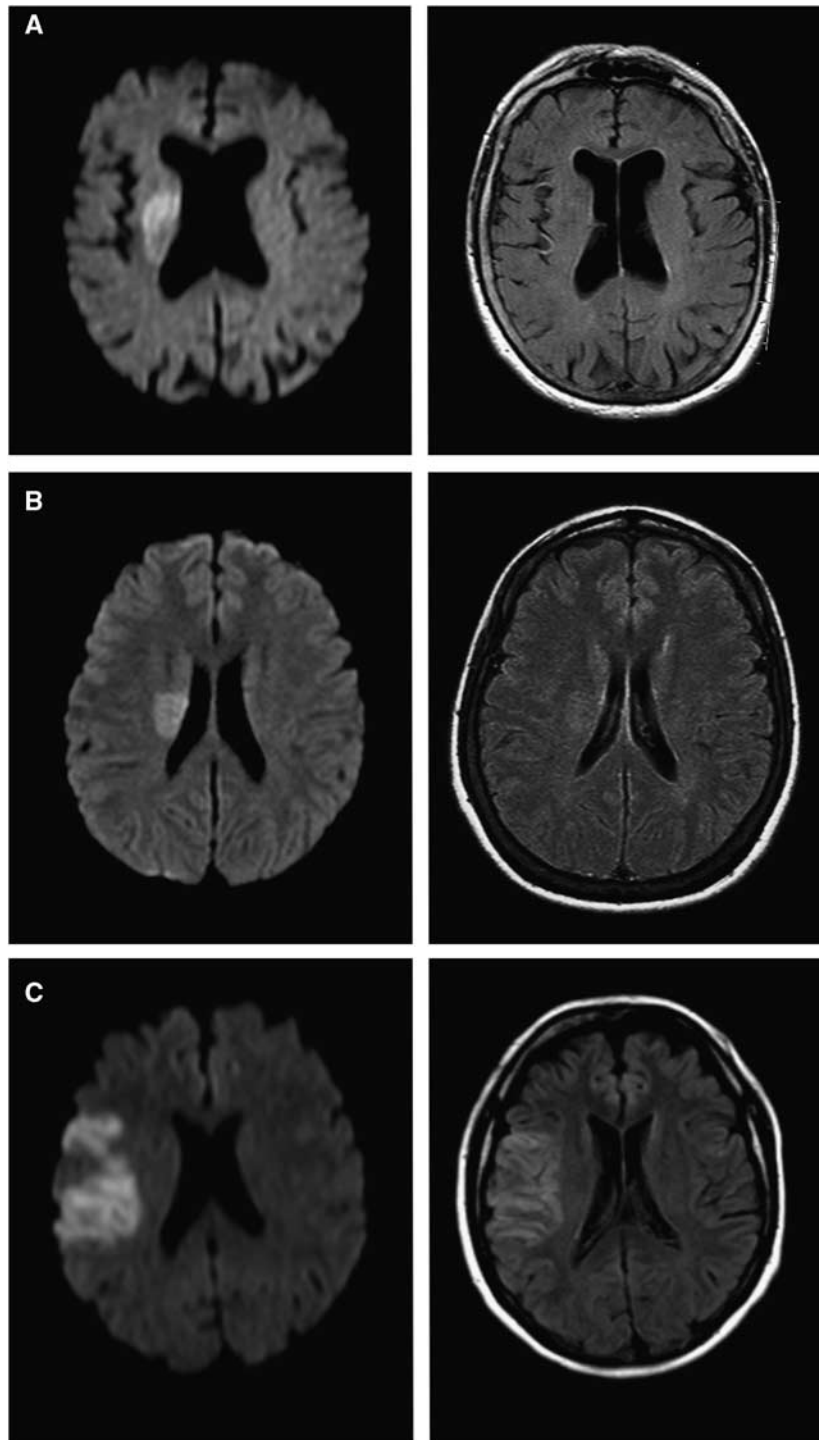


Figure 1. Examples of diffusion-weighted imaging (DWI) and fluid-attenuated inversion recovery (FLAIR) images. DWI (left column) and FLAIR (right column) images. Acute ischemic lesions are easily identified on all DWI. FLAIR images show (A) no corresponding FLAIR lesions; (B) 'subtle' FLAIR lesions; and (C) 'obvious' FLAIR lesions.

threshold determined by the ROC curve (model 3) and for six predictive models resulting from CART analysis depending on categorization of patients in individual leafs (models 4–9).

Statistical analysis was performed with the statistical package R (version 2.11.1, R Foundation for Statistical Computing, Vienna, Austria) and SPSS (version 19.0, Chicago, IL, USA). $P < 0.05$ was chosen as the significance level (two tailed).

RESULTS

Of the 516 patients included in the final analysis in PRE-FLAIR, 117 had to be excluded for the following reasons: insufficient image quality for rSI analysis ($n = 22$), bilateral stroke lesions interfering with measurements of contralateral reference SI ($n = 18$), and poor alignment of FLAIR and DWI images after rigid registration precluding projection of DWI VOI to FLAIR ($n = 77$).

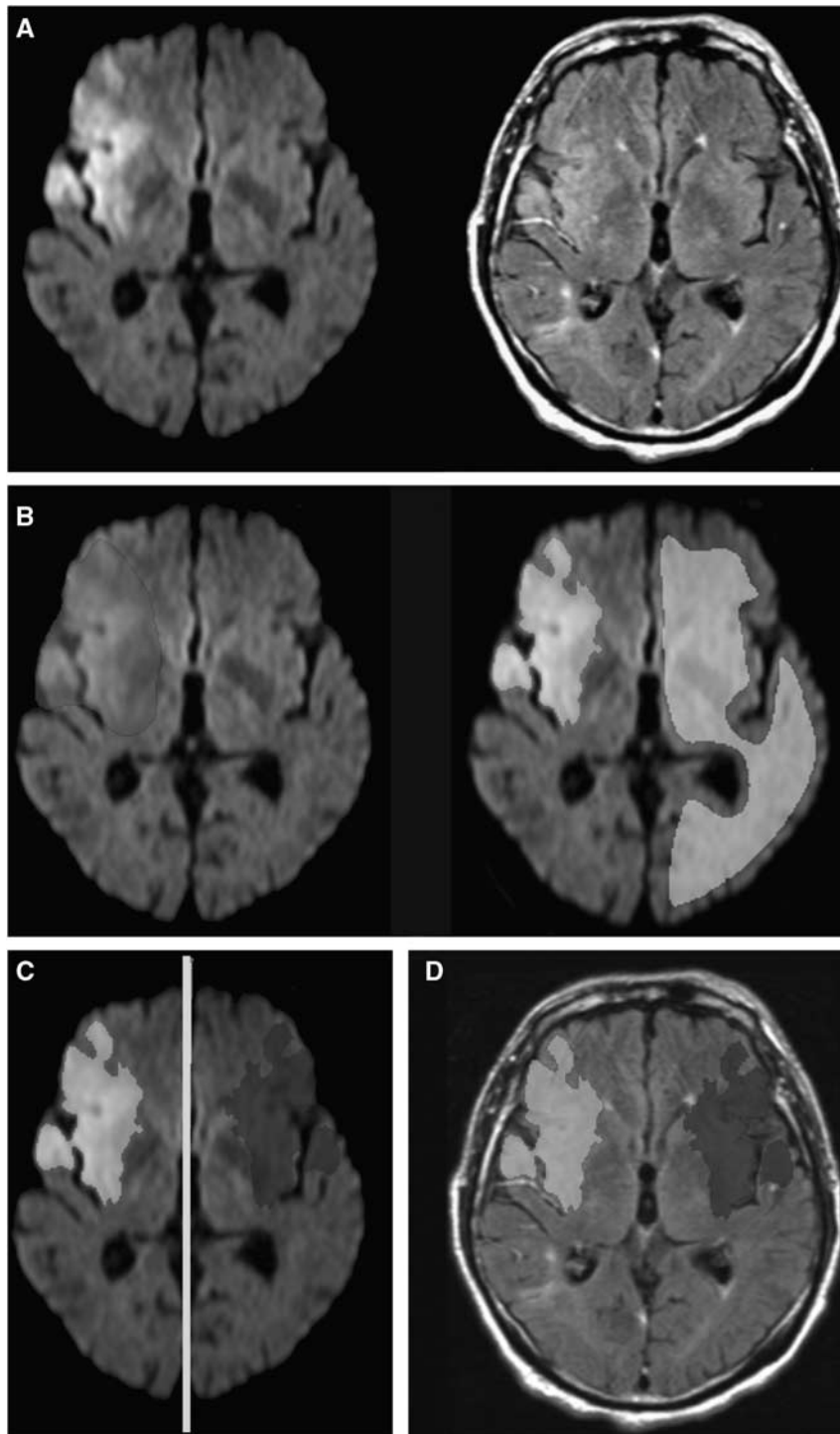


Figure 2. Postprocessing of magnetic resonance (MR) images for quantitative measurements. (A) Rigid registration of fluid-attenuated inversion recovery (FLAIR) images to the diffusion-weighted (DW) image; (B) definition of the ischemic lesion volume of interest (VOI): manual delineation of the lesion VOI with generous margin (red) and manual delineation of contralateral reference VOI (yellow), calculation of the refined ischemic lesion VOI ('refined lesion VOI', green); (C) projection of the 'refined lesion VOI' onto the contralateral hemisphere ('mirror VOI', blue); (D) projection of both VOIs onto previously registered FLAIR images. The color reproduction of this figure is available at the *Journal of Cerebral Blood Flow and Metabolism* journal online.

Thus, 399 of 516 patients (77.3%) were included in the final analysis.

In qualitative consensus rating, FLAIR images were classified as 'no lesion' in 172 (43.1%) patients, 'subtle lesion' in 116 (29.1%) patients, and 'obvious lesion' in 111 (27.8%) patients. Interobserver

agreement for qualitative judgment of FLAIR lesion visibility was 72.7% ($\kappa = 0.52$; 95% confidence interval (CI) 0.46–0.89) for 'no lesions', 41.4% ($\kappa = 0.31$; 0.24–0.39) for 'subtle lesions', and 72.1% ($\kappa = 0.53$; 0.45–0.61) for 'obvious lesions'. Group comparisons of clinical and imaging characteristics are presented in Table 1. Time

Table 1. Group comparisons of patients with 'no', 'subtle', and 'obvious' FLAIR lesion judged by visual rating

	No FLAIR lesion (n = 172)	'Subtle' FLAIR lesion (n = 116)	'Obvious' FLAIR lesion (n = 111)	P-value
Age (years)	69.7 (67.6–71.9)	61.1 (58.3–63.9)	63.5 (60.6–66.4)	<0.001
Female gender	80 (47%)	53 (46%)	53 (48%)	0.952
Right side of infarction	81 (47%)	46 (40%)	54 (49%)	0.499
NIHSS on admission	8 (7.1–8.9)	8.8 (7.6–10.1)	7.3 (6.3–8.5)	0.221
Time to MRI (min)	145.1 (131.2–160.6)	221.9 (198.9–247.5)	330.6 (294.2–371.5)	<0.001
Lesion volume (ml)	4.5 (3.5–5.6)	10.5 (7.7–14.1)	8.2 (5.9–11.4)	<0.001
Severe leukoaraiosis ^a	66 (38%)	26 (22%)	30 (27%)	0.010

FLAIR, fluid-attenuated inversion recovery; MRI, magnetic resonance imaging; NIHSS, National Institutes of Health Stroke Scale.

^aLeukoaraiosis defined as a score of > 1 on either of the subscales (deep white matter changes or periventricular white matter changes) of the adapted scale by Fazekas and Schmidt. Data are mean (95% confidence interval (CI)) or n (%).

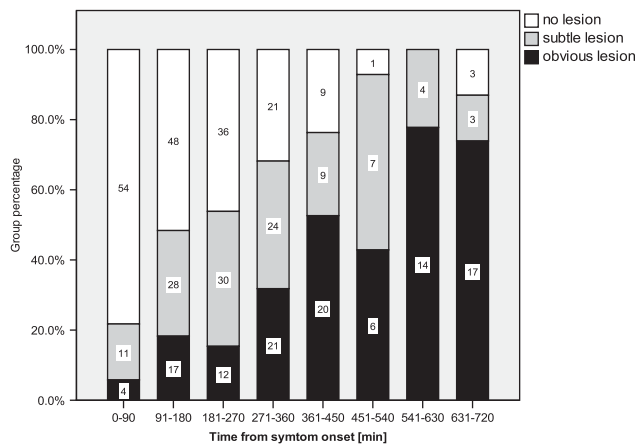


Figure 3. Fluid-attenuated inversion recovery (FLAIR) lesion visibility in relation to time from symptom onset. Numbers are patients within each time interval.

since symptom onset was significantly longer in the 'obvious' lesion group compared with both 'subtle' and no lesion groups. Additionally, patients with no FLAIR lesions were significantly older, had smaller DWI lesion volumes, and had more severe leukoaraiosis compared with patients with 'subtle' or 'obvious' lesions. However, the latter two groups did not differ significantly regarding these parameters in *post hoc* comparisons.

The frequency of patients with 'no lesion' and patients with 'obvious lesion' showed a clear relation to time from symptom onset (Figure 3). In contrast, the number of patients with 'subtle lesion' appeared rather equally distributed across the entire range of times from onset studied.

Across all patients, FLAIR rSI in the acute ischemic lesion showed a relative increase of 11.9% as compared with the contralateral hemisphere (mean FLAIR rSI 1.119; 95% CI 1.108–1.130). A significant linear increase of FLAIR rSI was found among the three groups defined by FLAIR lesion visibility ($P < 0.001$): mean FLAIR rSI was lowest in patients with 'no lesions' (1.073; 1.060–1.085), followed by 'subtle lesions' (1.118; 1.105–1.132), and 'obvious lesions' (1.193; 95% CI 1.165–1.220).

The FLAIR rSI showed a moderate but significant correlation with time from symptom onset ($r = 0.382$, $P < 0.001$) (Figure 4). Looking at the three groups separately, significant correlations between FLAIR rSI and time from symptom onset were found for the 'no lesion' and 'obvious lesion' groups ($r = 0.254$; 95% CI 0.099–0.38; $P < 0.001$ and $r = 0.274$; 95% CI 0.093–0.438; $P = 0.002$, respectively). However, no significant correlation was found in the 'subtle lesion' group ($r = 0.12$; 95% CI -0.063 to 0.295, $P = 0.45$).

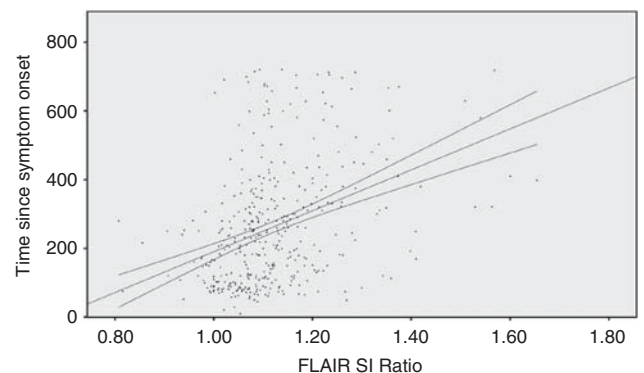


Figure 4. Scatterplot of fluid-attenuated inversion recovery (FLAIR) signal intensity (SI) ratios plotted against time from symptom onset (regression line and 95% mean confidence interval are shown).

The area under curve for FLAIR rSI to identify time from symptom onset ≤ 4.5 hours was 0.714 (95% CI 0.663–0.764). The ROC analysis identified a FLAIR rSI of 1.072 as the best cutoff value for predicting the symptom onset ≤ 4.5 hours with a Youden's index of 0.317 (see Table 2 for corresponding predictive values). The respective FLAIR rSI cutoff values for time windows of ≤ 3 and ≤ 6 hours are 1.0751 and 1.0972.

The CART analysis identified visual image categorization and FLAIR rSI as significant predictor of time from symptom onset, whereas DWI lesion volume was not retained in the model (Figure 5). The FLAIR rSI thresholds identified to achieve refined classification were 1.050 for the 'no lesion' group (leaves A and B) and 1.036 for the 'obvious lesion' group (leaves D and E). For the group of patients with 'subtle lesion', no improvement of classification by a FLAIR rSI threshold was found (leaf C).

Table 2 summarizes the predictive values for the different prediction models. As compared with the original 'liberal' visual rating algorithm (model 1), the modified 'conservative' visual rating (model 2) clearly increased sensitivity at the cost of a marked decrease in specificity. Prediction by FLAIR rSI analysis using rSI of 1.0721 as cutoff (model 3) resulted in a slight decrease of sensitivity countered by a comparably slight increase in specificity. Similar 'balancing' changes in sensitivity and specificity were observed in models 4–9, where results from CART were used for a variety of possible allocations to a binary ('FLAIR-positive' or 'FLAIR-negative') categorization. Across all predictive models, overall accuracy, i.e., the fraction of correctly classified patients, remained largely comparable (0.55–0.73) with widely overlapping 95% CIs.

For the alternative time windows of ≤ 3 hours, predictive values were: 0.54 (sensitivity; 95% CI 0.46–0.62), 0.75 (specificity; 0.70–

Table 2. Predictive values for the identification of patients within 4.5 hours of symptom onset

Rating method	Classification		Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	Accuracy (95% CI)
	FLAIR-negative	FLAIR-positive					
Model 1: visual rating ('liberal')	No lesions	'Subtle' and 'obvious' lesions	0.58 (0.52–0.65)	0.78 (0.72–0.85)	0.80 (0.74–0.86)	0.56 (0.50–0.62)	0.66 (0.61–0.70)
Model 2: visual rating ('conservative')	No and 'subtle' lesions	'Obvious' lesions	0.86 (0.82–0.91)	0.48 (0.41–0.56)	0.71 (0.66–0.76)	0.70 (0.62–0.79)	0.71 (0.67–0.75)
Model 3: quantitative (FLAIR rSI)	FLAIR rSI < 1.0721	FLAIR rSI > 1.0721	0.47 (0.40–0.53)	0.85 (0.80–0.91)	0.82 (0.79–0.89)	0.52 (0.46–0.58)	0.61 (0.57–0.66)
Model 4: CART	Leafs A–D	Leaf E	0.91 (0.87–0.94)	0.47 (0.39–0.55)	0.72 (0.67–0.77)	0.78 (0.69–0.86)	0.73 (0.69–0.77)
Model 5: CART	Leafs A, B, D	Leafs C, E	0.62 (0.56–0.68)	0.77 (0.71–0.84)	0.80 (0.74–0.86)	0.58 (0.51–0.65)	0.68 (0.63–0.72)
Model 6: CART	Leafs A, C, D	Leafs B, E	0.60 (0.54–0.66)	0.65 (0.58–0.73)	0.72 (0.66–0.78)	0.52 (0.46–0.57)	0.62 (0.57–0.67)
Model 7: CART	Leafs A, D	Leafs B, D, E	0.33 (0.27–0.39)	0.95 (0.92–0.98)	0.90 (0.84–0.97)	0.52 (0.46–0.57)	0.57 (0.54–0.59)
Model 8: CART	Leafs A, C	Leafs B, D, E	0.55 (0.49–0.62)	0.66 (0.59–0.70)	0.71 (0.64–0.77)	0.50 (0.44–0.57)	0.60 (0.55–0.65)
Model 9: CART	Leaf A	Leafs B, C, D, E	0.27 (0.21–0.33)	0.96 (0.93–0.99)	0.91 (0.85–0.98)	0.47 (0.42–0.53)	0.55 (0.52–0.57)

CART, Classification and Regression Tree; CI, confidence interval; FLAIR, fluid-attenuated inversion recovery; NPV, negative predictive value; PPV, positive predictive value; rSI, relative signal intensity.

Classification describes the algorithm applied to results from individual rating methods to binary groups of 'FLAIR-positive' and 'FLAIR-negative' patients. Accuracy describes the fraction of correctly classified patients.

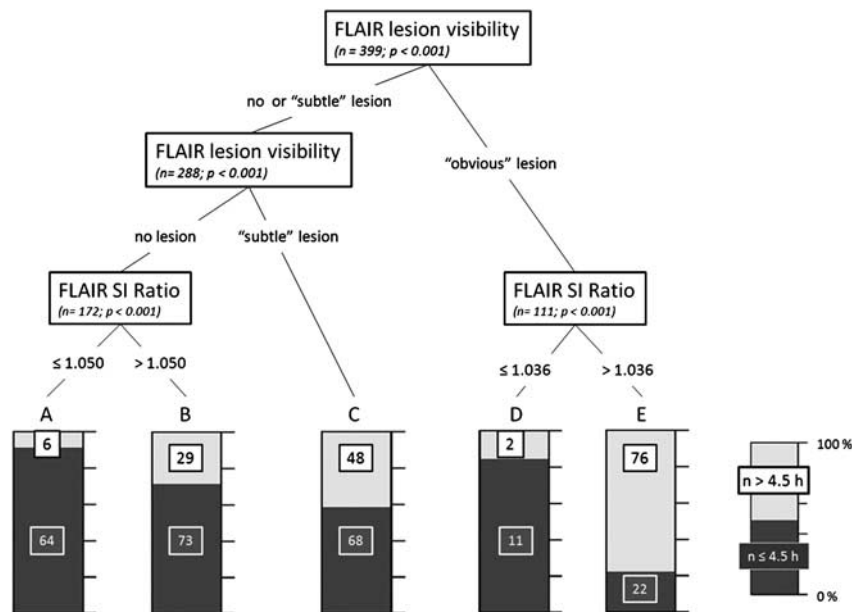


Figure 5. Classification and Regression Tree. Classification regarding time from symptom onset \leq or $>$ 4.5 hours; leafs A–E show individual patient numbers categorized by significant splitting predictor variables; numbers in individual leafs denote patients with time from symptom onset \leq 4.5 hours (black) or $>$ 4.5 hours (white). Variables included in the model: fluid-attenuated inversion recovery (FLAIR) lesion visibility, FLAIR signal intensity (SI) ratio, and diffusion-weighted imaging (DWI) lesion volume.

0.81), 0.59 (PPV; 0.51–0.67), 0.72 (NPV; 0.66–0.77), and diagnostic accuracy: 0.67. For time since witnessed onset \leq 6 hours: 0.43 (Sensitivity; 0.38–0.49), 0.87 (Specificity; 0.80–0.94), 0.92 (PPV; 0.87–0.96), 0.32 (NPV; 0.26–0.38), diagnostic accuracy: 0.54.

DISCUSSION

To summarize the main results of this study, quantitative analysis of FLAIR relative signal intensities predicted the time from symptom onset \leq 4.5 hours with a similar specificity as visual analysis, therefore validating the latter as an appropriate rating technique. However, contrary to our hypothesis, the addition of quantitative FLAIR signal intensity measurements did not improve sensitivity. Neither refined visual analysis nor quantitative FLAIR

analysis nor the combination of both in multivariate predictive models improved identification of patients within the 4.5-hour time window. Observed differences in predictive values between the differing approaches only reflected a 'trade-off' between specificity and sensitivity.

Although judgment of a DWI/FLAIR mismatch is straightforward in cases of marked presence of corresponding FLAIR lesions, subtle changes in FLAIR signal intensity pose a challenge to the observer.¹⁵ This is reflected by the poor interpreter agreement for the visual classification of 'subtle lesions'. Previous studies have acknowledged this finding and applied different algorithms to assign these patients to two unambiguous groups of 'obvious' or 'absent' lesions.^{6,7} in the PRE-FLAIR study, patients were classified as 'FLAIR-positive' (i.e., an acute ischemic lesion visible in FLAIR

corresponding to a lesion visible on DWI) if any, although subtle, parenchymal hyperintensity was visible, thus subsuming 'obvious' and 'subtle' FLAIR lesions. This 'liberal' rating scheme yielded a specificity of 0.78 and sensitivity of 0.62 for the identification of patients within 4.5 hours of stroke in the original study, and corresponding values of 0.78 (specificity) and 0.58 (sensitivity) in the subgroup included in this analysis. Low sensitivity was identified as a major drawback of this approach, translating into a large proportion of acute ischemic lesions misclassified as >4.5 hours of age. Given the observation that rather subtle hyperintensity can be observed already within the first hours of cerebral ischemic on FLAIR together with the nearly linear increase of T2 signal during the first hours of stroke,¹⁶ we used a refined 'conservative' rating algorithm disregarding subtle FLAIR hyperintensities. This approach resulted in a considerable increase in sensitivity, but countered by a comparable decrease in specificity. Translated into potential clinical applications, this implies an increased number of patients being wrongly classified as ≤ 4.5 hours of symptom onset while actually being beyond 4.5 hours and thus beyond the approved time window for thrombolysis; this is an evidently undesirable effect considering the use of DWI/FLAIR mismatch to select patients with unknown time of symptom onset for thrombolytic therapy.

As a second approach to overcome the limited sensitivity of the original visual analysis, we tested quantitative analysis of FLAIR signal intensities. However, neither classification based on rSI measurements nor a multivariate classification model combining visual judgment and quantitative analysis resulted in an improvement of overall prediction. The optimal FLAIR rSI threshold identified by ROC analysis resulted in a comparably low sensitivity as the initial visual analysis. Any increase of this threshold to increase sensitivity at the same time results in a decrease in specificity. The same effect was seen for the different predictive models based on CART analysis. Depending on the allocation of the individual leafs partitioned by the algorithm, either sensitivity or specificity was increased, whereas overall accuracy remained largely unchanged.

A low sensitivity of any classification algorithm for the identification of patients ≤ 4.5 hours of symptom onset leads to patients being misclassified to the group of patients with late (>4.5 hours) symptom onset. In the context of a potential clinical application or clinical trial, these patients would therefore be excluded from a potentially beneficial therapy aiming at an early time window. Low sensitivity stems from a significant number of patients in the time window of ≤ 4.5 hours showing visible, i.e., 'obvious' (42/329; 12.7%), or 'subtle' (85/329; 25.8%) FLAIR lesions. This phenomenon can be attributed to the evolution of FLAIR signal intensities after ischemic stroke. In FLAIR sequences, an inversion recovery pulse is applied to null signal from cerebrospinal fluid before measuring a T2-weighted signal.^{17,18} The contributing T2 signal chiefly increases because of net increase in tissue water, i.e., cytotoxic and vasogenic edema after ischemic stroke.¹⁹ Increased T2 signal is noticed within the first 2 hours after stroke,^{20,21} and early visibility of FLAIR lesions have been described in previous studies.^{6,7,22} In terms of quantitative measurements, T2-signal intensities have shown to linearly increase within the time from stroke onset,^{16,23} but this relationship was not reflected in a strong linear correlation of FLAIR rSI in our study. This might be attributed to the fact that contrary to quantitative T2 measurements, signal intensities of FLAIR sequences are not equally quantifiable because of the application of different inversion recovery pulses. Additionally, FLAIR signal intensities are influenced by irregularly distributed phenomena such as diffuse white matter lesions, and cerebrospinal fluid-related or T2 shine-through artifacts, as well as sensitivity to patient movement.^{24,25} This is exemplarily shown by patients with FLAIR rSI < 1, where signal intensities were higher in the unaffected hemisphere despite manual correction of VOI

positions, an implausible finding if FLAIR SI increase would only result from tissue water increase resulting from cerebral ischemia.

Previous studies using FLAIR rSI measurements for the identification of time from symptom onset have yielded contradictory results: Ebinger *et al*⁶ reported no significant correlation ($r = -0.15$; $P = 0.128$) in 94 acute stroke patients, whereas a significant correlation was found in a study by Petkova *et al*⁷ including 130 acute stroke patients ($r = 0.63$, $P < 0.001$). In this study, a FLAIR rSI threshold of >1.07 (7% relative increase) predicted time from symptom onset >3 hours with high specificity and sensitivity (0.92 and 0.90, respectively). Interestingly, this threshold is quite close to the cutoff value that we found in our study for identifying patients within 4.5 hours (rSI = 1.072). Another recent study reported a FLAIR SI ratio of <1.15 (15% relative increase) on measurements in the area of brightest FLAIR hyperintensities optimal cutoff to identify patients within 4.5 hours of symptom onset with a PPV of 0.90.²⁶ Discrepancies in the results of rSI measurement and the predictive value of certain FLAIR rSI thresholds may partly be explained by methodology and patient characteristics. For the purpose of this study, FLAIR lesions were delineated on multiple slices based on back-projected DWI lesions similar to the approach of Petkova *et al*.⁷ However, analysis was performed on data from multiple centers involving a heterogeneous set of FLAIR sequence parameters, which might contribute to a decreased homogeneity of relative FLAIR signal intensities across the entire sample and thus explain a less clear association with time as reported by Petkova *et al*.⁷

Recently, a different approach of FLAIR SI measurements was suggested that measures signal intensities in manually placed ROI corresponding to the area of brightest FLAIR hyperintensity.²⁶ It is a frequent observation that FLAIR hyperintensities in acute stroke lesions are not evenly distributed but show a rather patchy pattern that may reflect regional heterogeneity of tissue changes. It is not surprising that an analysis focusing on smaller areas within the ischemic lesion with the most intense FLAIR changes results in higher FLAIR SI ratios. Of course, reader-measured SI ratios based on manually placed ROI are observer dependent and prone to variability due to individual differences in ROI placement. Conversely, they can be performed within minutes and thus may be more suitable to be used to guide treatment decisions in the acute setting. This is suggested by Song *et al*²⁶ as a pragmatic two-step procedure to evaluate FLAIR to identify patients with unknown time of symptom onset suitable for enrollment in a clinical trial of stroke thrombolysis.²⁶⁻²⁸

The results of our study suggest the conclusions that the accuracy of FLAIR as surrogate marker of lesion age has limitations, which are inherent to either the method or to the sample studied and which cannot be overcome by more sophisticated ways of analyzing images. There appears to be a group of patients in whom the assessment of FLAIR hyperintensities falls short of a temporal allocation of the ischemic lesion no matter whether images are assessed by visual judgment or quantitative analysis. The group of patients with lesions on FLAIR classified as 'subtle' by visual rating was more or less equally distributed over the time range studied. This observation was corroborated by the lack of a significant correlation between FLAIR rSI ($r = 0.382$) and time from symptom onset and the failure of CART to provide a further classification based on FLAIR rSI measurements in this group of patients.

A closer look at the clinical characteristics of this subgroup might provide some clue as to the reasons for the failure of FLAIR as surrogate marker of lesion age in these patients. However, patients with 'subtle' FLAIR lesions differed significantly from the other two groups only in time from symptom onset, and we detected no variables that exceptionally characterized patients with 'subtle' lesions.

Of note, DWI lesion volume was not retained as a significant parameter in the CART analysis. Ischemic lesion volume has

previously been shown to influence the conspicuity of a corresponding FLAIR lesion in visual analysis.^{6–9} Accordingly, Petkova *et al*⁷ integrated lesion size in their classification algorithm, assigning patients with visually rated ‘large subtle lesions’ to a FLAIR-negative group and vice versa. In our sample, the inclusion of lesion volume into a multivariate classification analysis did not improve classification beyond the results of combined visual and quantitative rating.

The assessment of FLAIR together with DWI as a surrogate marker of lesion age has a profound clinical impact, as this approach might pave the way to an effective treatment of stroke patients with unknown time window, e.g., patients waking up with stroke symptoms. DWI/FLAIR mismatch has been suggested to identify patients with unknown symptom onset likely to benefit from stroke.^{8,9} Just recently, DWI/FLAIR mismatch has been reported to be a frequent finding in wake-up stroke patients.²⁹ Currently, at least two clinical trials are underway that apply DWI/FLAIR mismatch to enroll stroke patients with unknown symptom onset for treatment with thrombolysis: WAKE-UP, a European, randomized, controlled trial of efficacy and safety of intravenous thrombolysis with alteplase in MRI-selected patients (NCT01525290), and MR-WITNESS, an American safety study of intravenous thrombolysis with alteplase in MRI-selected patients (NCT01282242).

In conclusion, quantitative measurements of FLAIR rSI confirmed the results of visual judgment but did not improve diagnostic accuracy in identifying patients within a time window of 4.5 hours of symptom onset. We identified a group of patients displaying ‘subtle’ lesions in whom visual categorization remains challenging and is not improved by the addition of quantitative measurements. Given these results, it remains doubtful if prediction of stroke onset time can be achieved by a universal FLAIR rSI cutoff value superiorly to visual judgments. However, quantitative measurements could have a value in supporting visual categorization because of similar specificity for the prediction of the thrombolytic time window. Advanced neuroimaging markers such as quantitative T2 measurements,¹⁶ sodium MR imaging,³⁰ or T_{1ρ} measurements³¹ are other possible candidates for imaging surrogate parameters of time since symptom onset, and further studies are needed to test applicability in clinical practice.

DISCLOSURE/CONFLICT OF INTEREST

Jochen B Fiebach has received fees as a board member, consultant, or lecturer from Boehringer Ingelheim, Lundbeck, Siemens, Syngis, and Synarc. Christian Gerloff has received fees as a consultant or lecture fees from Bayer Vital, Boehringer Ingelheim, EBS technologies, Glaxo Smith Kline, Lundbeck, Pfizer, Sanofi Aventis, Silk Road Medical, and UCB. David S Liebeskind has received fees as a consultant for CoAxia and Concentric Medical. A Gregory Sorensen was supported by grants from the National Institutes of Health and has received fees as a board member, consultant, or lecturer from ACRIN, Biogen, Genzyme, Mitsubishi, and has received royalties from GE and Olea. Götz Thomalla has received a research grant from the Else Kröner-Fresenius-Stiftung. Thomas Tourdias has received a national grant from the French Government (PHRC). Ona Wu was supported in part by grants from the National Institutes of Health (R01NS059775, R01NS063925, and P50NS051343) and has received royalties and licensing fees from GE, Olea, and Imaging Biometrics. The other authors report no conflict of interest.

ACKNOWLEDGEMENTS

PRE-FLAIR has received funding from the Else Kröner-Fresenius-Stiftung (2009_A36). In Berlin, data were collected within the 1,000+ study, which has received funding from the Federal Ministry of Education and Research via the grant Center for Stroke Research Berlin (01 EO 0801). EPITHEM was a phase II prospective, randomized, double-blinded, placebo-controlled, multinational trial funded by the National Health and Medical Research Council (Australia), the National Stroke Foundation (Australia), and the Heart Foundation of Australia. VIRAGE is a national multicenter study supported by French national grant PHRC.

REFERENCES

- 1 Omama S, Yoshida Y, Ogawa A, Onoda T, Okayama A. Differences in circadian variation of cerebral infarction, intracerebral haemorrhage and subarachnoid haemorrhage by situation at onset. *J Neurol Neurosurg Psychiatry* 2006; **77**: 1345–1349.
- 2 Serena J, Davalos A, Segura T, Mostacero E, Castillo J. Stroke on awakening: looking for a more rational management. *Cerebrovasc Dis* 2003; **16**: 128–133.
- 3 Fink JN, Kumar S, Horkan C, Linfante I, Selim MH, Caplan LR *et al*. The stroke patient who woke up: clinical and radiological features, including diffusion and perfusion MRI. *Stroke* 2002; **33**: 988–993.
- 4 Adams Jr. HP, del Zoppo G, Alberts MJ, Bhatt DL, Brass L, Furlan A *et al*. Guidelines for the early management of adults with ischemic stroke: a guideline from the American Heart Association/American Stroke Association Stroke Council, Clinical Cardiology Council, Cardiovascular Radiology and Intervention Council, and the Atherosclerotic Peripheral Vascular Disease and Quality of Care Outcomes in Research Interdisciplinary Working Groups: the American Academy of Neurology affirms the value of this guideline as an educational tool for neurologists. *Stroke* 2007; **38**: 1655–1711.
- 5 Committee. ESOEECEW. Guidelines for management of ischaemic stroke and transient ischaemic attack 2008. *Cerebrovasc Dis* 2008; **25**: 457–507.
- 6 Ebinger M, Galinovic I, Rozanski M, Brunecker P, Endres M, Fiebach JB. Fluid-attenuated inversion recovery evolution within 12 hours from stroke onset: a reliable tissue clock? *Stroke* 2010; **41**: 250–255.
- 7 Petkova M, Rodrigo S, Lamy C, Oppenheim G, Touze E, Mas JL *et al*. MR imaging helps predict time from symptom onset in patients with acute stroke: implications for patients with unknown onset time. *Radiology* 2010; **257**: 782–792.
- 8 Thomalla G, Rossbach P, Rosenkranz M, Siemonsen S, Krutzmann A, Fiehler J *et al*. Negative fluid-attenuated inversion recovery imaging identifies acute ischemic stroke at 3 hours or less. *Ann Neurol* 2009; **65**: 724–732.
- 9 Thomalla G, Cheng B, Ebinger M, Hao Q, Tourdias T, Wu O *et al*. DWI-FLAIR mismatch for the identification of patients with acute ischaemic stroke within 4.5 h of symptom onset (PRE-FLAIR): a multicentre observational study. *Lancet Neurol* 2011; **10**: 978–986.
- 10 Aoki J, Kimura K, Iguchi Y, Shibazaki K, Sakai K, Iwanaga T. FLAIR can estimate the onset time in acute ischemic stroke patients. *J Neurol Sci* 2010; **293**: 39–44.
- 11 Thomalla G, Cheng B, Ebinger M, Hao Q, Tourdias T, Wu O *et al*. DWI-FLAIR mismatch for the identification of patients with acute ischaemic stroke within 4.5 h of symptom onset (PRE-FLAIR): a multicentre observational study. *Lancet Neurol* **10**: 978–986.
- 12 Kapeller P, Barber R, Vermeulen RJ, Ader H, Scheltens P, Freidl W *et al*. Visual rating of age-related white matter changes on magnetic resonance imaging: scale comparison, interrater agreement, and correlations with quantitative measurements. *Stroke* 2003; **34**: 441–445.
- 13 Breiman L, Friedman JH, Olshen RA, Stone CJ. *Classification and Regression Trees*. Chapman & Hall: New York, NY, 1984.
- 14 Hothorn T, Hornik K, Zeileis A. Unbiased Recursive Partitioning: A Conditional Inference Framework. *J Comput Graph Stat* 2006; **15**: 651–674.
- 15 Ziegler A, Ebinger M, Fiebach JB, Audebert HJ, Leistner S. Judgment of FLAIR signal change in DWI-FLAIR mismatch determination is a challenge to clinicians. *J Neurol* 2012; **259**: 971–973.
- 16 Siemonsen S, Mouridsen K, Holst B, Ries T, Finsterbusch J, Thomalla G *et al*. Quantitative t2 values predict time from symptom onset in acute stroke patients. *Stroke* 2009; **40**: 1612–1616.
- 17 De Coene B, Hajnal JV, Gatehouse P, Longmore DB, White SJ, Oatridge A *et al*. MR of the brain using fluid-attenuated inversion recovery (FLAIR) pulse sequences. *AJNR Am J Neuroradiol* 1992; **13**: 1555–1564.
- 18 Hajnal JV, Bryant DJ, Kasuboski L, Pattany PM, De Coene B, Lewis PD *et al*. Use of fluid attenuated inversion recovery (FLAIR) pulse sequences in MRI of the brain. *J Comput Assist Tomogr* 1992; **16**: 841–844.
- 19 Watanabe O, West CR, Bremer A. Experimental regional cerebral ischemia in the middle cerebral artery territory in primates. Part 2: Effects on brain water and electrolytes in the early phase of MCA stroke. *Stroke* 1977; **8**: 71–76.
- 20 Hoehn-Berlage M, Eis M, Back T, Kohno K, Yamashita K. Changes of relaxation times (T1, T2) and apparent diffusion coefficient after permanent middle cerebral artery occlusion in the rat: temporal evolution, regional extent, and comparison with histology. *Magn Reson Med* 1995; **34**: 824–834.
- 21 Venkatesan R, Lin W, Gurleyik K, He YY, Paczynski RP, Powers WJ *et al*. Absolute measurements of water content using magnetic resonance imaging: preliminary findings in an in vivo focal ischemic rat model. *Magn Reson Med* 2000; **43**: 146–150.
- 22 Aoki J, Kimura K, Iguchi Y, Shibazaki K, Iwanaga T, Watanabe M *et al*. Intravenous thrombolysis based on diffusion-weighted imaging and fluid-attenuated inversion recovery mismatch in acute stroke patients with unknown onset time. *Cerebrovasc Dis* 2011; **31**: 435–441.

- 23 Jokivarsi KT, Hiltunen Y, Grohn H, Tuunanen P, Grohn OH, Kauppinen RA. Estimation of the onset time of cerebral ischemia using T1rho and T2 MRI in rats. *Stroke* **41**: 2335–2340.
- 24 Gawne-Cain ML, Silver NC, Moseley IF, Miller DH, . Fast FLAIR of the brain: the range of appearances in normal subjects and its application to quantification of white-matter disease. *Neuroradiology* 1997; **39**: 243–249.
- 25 William Michael B. Fast fluid attenuated inversion recovery (FLAIR) imaging and associated artefacts in magnetic resonance imaging (MRI). *Radiography* 2007; **13**: 283–290.
- 26 Song SS, Latour LL, Ritter CH, Wu O, Tighiouart M, Hernandez DA *et al*. A pragmatic approach using magnetic resonance imaging to treat ischemic strokes of unknown onset time in a thrombolytic trial. *Stroke* 2012; **43**: 2331–2335.
- 27 Ebinger M, Kufner A, Galinovic I, Brunecker P, Malzahn U, Nolte CH *et al*. Fluid-attenuated inversion recovery images and stroke outcome after thrombolysis. *Stroke* 2011; **43**: 539–542.
- 28 Schwamm LH, Sorensen AG, Wu O. MR WITNESS Investigators. MR WITNESS–Safety trial of IV rt-PA in patients with unwitnessed stroke onset. 2010. *International Stroke Conference*. San Antonio (TX) 2010.
- 29 Huisa BN, Liebeskind DS, Raman R, Hao Q, Meyer BC, Meyer DM *et al*. Diffusion-weighted imaging-fluid attenuated inversion recovery mismatch in nocturnal stroke patients with unknown time of onset. *J Stroke Cerebrovasc Dis* 2012; e-pub ahead of print.
- 30 Thulborn KR, Gindin TS, Davis D, Erb P. Comprehensive MR imaging protocol for stroke management: tissue sodium concentration as a measure of tissue viability in nonhuman primate studies and in clinical studies. *Radiology* 1999; **213**: 156–166.
- 31 Grohn OHJ, Kettunen MI, Makela HI, Penttonen M, Pitkanen A, Lukkarinen JA *et al*. Early detection of irreversible cerebral ischemia in the rat using dispersion of the magnetic resonance imaging relaxation time, T1rho. *J Cereb Blood Flow Metab* 2000; **20**: 1457–1466.

APPENDIX

STIR and VISTA Imaging steering committee members: Steven Warach (chair), Gregory Albers, Stephen Davis, Geoffrey Donnan,

Marc Fisher, Anthony Furlan, James Grotta, Werner Hacke, Dong-Wha Kang, Chelsea Kidwell, Walter Koroshetz, Kennedy R Lees, Michael Lev, David S Liebeskind, A Gregory Sorensen, Vincent Thijs, Götz Thomalla, Joanna Wardlaw, and Max Wintermark.