Adapting the Computed Tomography Criteria of Hemorrhagic Transformation to Stroke Magnetic Resonance Imaging

Lars Neeb a, b  Kersten Villringer a  Ivana Galinovic a  Florian Grosse-Dresselhaus a, b  Ramanan Ganeshan a, b  Daniel Gierhake a  Claudia Kunze a  Ulrike Grittner a, c  Jochen B. Fiebach a

a Center for Stroke Research Berlin (CSB), and Departments of b Neurology and c Biostatistics and Clinical Epidemiology, Charité – Universitätsmedizin Berlin, Berlin, Germany

Key Words
Acute ischemic stroke · Hemorrhagic transformation · Magnetic resonance imaging · Computed tomography · Thrombolysis · Imaging in stroke

Abstract
Background: The main safety aspect in the use of stroke thrombolysis and in clinical trials of new pharmaceutical or interventional stroke therapies is the incidence of hemorrhagic transformation (HT) after treatment. The computed tomography (CT)-based classification of the European Cooperative Acute Stroke Study (ECASS) distinguishes four categories of HTs. An HT can range from a harmless spot of blood accumulation to a symptomatic space-occupying parenchymal bleeding associated with a massive deterioration of symptoms and clinical prognosis. In magnetic resonance imaging (MRI) HTs are often categorized using the ECASS criteria although this classification has not been validated in MRI. We developed MRI-specific criteria for the categorization of HT and sought to assess its diagnostic reliability in a retrospective study.

Methods: Consecutive acute ischemic stroke patients, who had received a 3-tesla MRI before and 12–36 h after thrombolysis, were screened retrospectively for an HT of any kind in post-treatment MRI. Intravenous tissue plasminogen activator was given to all patients within 4.5 h. HT categorization was based on a simultaneous read of 3 different MRI sequences (fluid-attenuated inversion recovery, diffusion-weighted imaging and T2* gradient-recalled echo). Categorization of HT in MRI accounted for the various aspects of the imaging pattern as the shape of the bleeding area and signal intensity on each sequence. All data sets were independently categorized in a blinded fashion by 3 expert and 3 resident observers. Interobserver reliability of this classification was determined for all observers together and for each group separately by calculating Kendall’s coefficient of concordance (W). Re-
**Results:** Of the 186 patients screened, 39 patients (21%) had an HT in post-treatment MRI and were included for the categorization of HT by experts and residents. The overall agreement of HT categorization according to the modified classification was substantial for all observers ($W = 0.79$). The degrees of agreement between experts ($W = 0.81$) and between residents ($W = 0.87$) were almost perfect. For the distinction between parenchymal hematoma and hemorrhagic infarction, the interobserver agreement was almost perfect for all observers taken together ($W = 0.82$) as well as when experts ($W = 0.82$) and residents ($W = 0.91$) were analyzed separately. **Conclusion:** The ECASS CT classification of HT was successfully adapted for usage in MRI. It leads to a substantial to almost perfect interobserver agreement and can be used for safety assessment in clinical trials.

**Introduction**

Thrombolytic therapy with recombinant tissue plasminogen activator (rt-PA) up to 4.5 h after stroke onset remains the only available causal therapy for acute stroke to date [1]. The most critical risk of rt-PA therapy is a severe hemorrhagic transformation (HT) of the infarction. Its incidence is one of the main safety endpoints used in thrombolytic clinical trials. An HT can range from a harmless spot of blood accumulation to a space-occupying parenchymal bleeding associated with a massive deterioration of symptoms and clinical prognosis.

A clinically relevant, symptomatic HT can be found in approximately 6% of the rt-PA-treated patients [2]. To classify the different bleeding types, various criteria based on clinical and radiological findings have been developed [3–9]. In clinical practice and stroke trials, the radiological criteria developed for computed tomography (CT) in the European Cooperative Acute Stroke Study (ECASS) in the 1990s are mainly used for the separation of HT into four different subtypes. These criteria differentiate between hemorrhagic infarcts (HI) with small petechiae or more confluent petechiae (HI-1 and HI-2) and parenchymal hematomas (PH) $\leq 30\%$ of the infarcted area with some mild space-occupying effect (PH-1) and $>30\%$ of the infarcted area with a significant space-occupying effect or a clot remote from the infarcted area (PH-2) [3, 10]. As the occurrence of PH after thrombolytic therapy has been shown to be associated with a poor clinical outcome [10–13], an accurate and standardized specification of HT is crucial for determining the safety of stroke therapies in clinical trials.

Due to its diagnostic advantages and its increasing availability, cerebral magnetic resonance imaging (MRI) has been established in clinical stroke trials and also in diagnostic routine in the last years. MRI detects an acute hemorrhage after stroke as well as [14–16] or even better [17, 18] than CT. HTs in MRI are often classified using the CT-based ECASS criteria. However, this classification has not been validated in MRI. No MRI-specific criteria exist that consider how the different types of HT present in different MRI sequences and which sequences are best used for their categorization.

Based on the four different subtypes of the ECASS criteria (HI-1, HI-2, PH-1 and PH-2), we developed specific criteria for the categorization of HT in MRI that account for the various aspects of the imaging pattern as the shape of the bleeding area and signal intensity on different types of images. To validate the diagnostic reliability of the new criteria, we retrospectively evaluated 39 patients with an HT who underwent an MRI scan 12–36 h after thrombolysis. HTs in MRI were categorized by 3 board-certified radiologists and 3 observers with at least 1 year of stroke MRI experience. The results of the different observers and groups were compared to determine the interobserver agreement of this new classification.
Methods

Patients
This is a retrospective, single-center study conducted at the Center for Stroke Research Berlin at the Campus Benjamin Franklin of the Charité University Hospital. Consecutive acute stroke patients recruited between March 2008 and July 2012 from the 1000Plus study [19] who fulfilled the following criteria were eligible: (1) they had to have received a systemic thrombolysis for acute stroke treatment within 4.5 h after symptom onset and (2) they had to have undergone 3-tesla MRI examination before and within 12–36 h after the end of rt-PA infusion. The standard MRI protocol included T2* gradient-recalled echo (T2*GRE), fluid-attenuated inversion recovery (FLAIR) and diffusion-weighted imaging (DWI). All patients had given informed consent for the creation of imaging data and consented that the data would be used for substudies within the 1000Plus project. To be included in image analysis, follow-up MRI scans had to show a new HT that was identified and documented in the clinical report.

Imaging Protocol
All patients participating in the 1000Plus study went through a standardized MRI protocol as described before [19]. All examinations were conducted on a 3.0-tesla MRI scanner (Tim Trio; Siemens AG, Erlangen, Germany). In brief, T2*-weighted sequence parameters were: 25 slices; slice thickness = 5 mm; interslice gap = 0.5 mm; field of view (FOV) = 220 mm; matrix = 256 × 192; repetition time (TR) = 620 ms; echo time (TE) = 20 ms; flip angle 20°. DWI was conducted with 50 slices; slice thickness = 2.5 mm; interslice gap = 0.0 mm; FOV = 230 mm; matrix = 192 × 192; TR = 7,600 ms; TE = 93 ms; b = 0 s/mm² (b0) and 1,000 s/mm² (b1). FLAIR parameters were: 25 slices; slice thickness = 5 mm; interslice gap = 0.5 mm; matrix = 256 × 256; FOV = 220 mm; TR = 8,000 ms; TE = 100 ms; flip angle 130°.

Image Analysis
The two data sets of each patient (pre- and post-thrombolysis including DWI, T2*GRE and FLAIR) were anonymized and randomly distributed. Each data set was evaluated by 6 blinded observers which formed two groups of 3 observers. The first group consisted of neurology residents with 1-year experience in stroke imaging (residents) and the second group was comprised of board-certified radiologists with 5–16 years of experience in stroke- and neuroimaging (experts).

Each observer scored the HT seen on follow-up MRI according to the newly developed criteria listed below.
(1) Small petechial hemorrhage (HI-1): In T2*GRE, isolated hypointensities (in spots or lines) smaller than 10 mm that appear hypo- or isointense in FLAIR and DWI.
(2) Confluent petechia within the infarcted area (HI-2): In T2*GRE, confluent hypointensities >10 mm that appear hypo- or isointense in FLAIR with signal intensity in DWI close to the initial infarct signal (hypo- to hyperintense).
(3) Parenchymal hemorrhage <1/3 of the infarcted area (PH-1): In T2*GRE, round-shaped hypointensity, sometimes with central hyperintensity and hyperintense perifocal edema. In FLAIR, the lesion appears hyperintense with an iso- or hypointense border between bleeding and perifocal edema. Signal intensity in DWI is strongly hypointense with iso- to hyperintense areas.
(4) Parenchymal hemorrhage >1/3 of the infarcted area (PH-2): In T2*GRE, round-shaped hypointensity with possible central hyperintensity and hyperintense perifocal edema. In FLAIR, the lesion is hyperintense with an iso- or hypointense border between bleeding and perifocal edema. The lesion in DWI is very hypointense with iso- to hyperintense portions. All hemorrhagic clots distant from the infarcted area are rated as PH-2.
Examples for each bleeding category are provided in figure 1.

All HTs seen on the MRI sequences were evaluated individually by each observer for one of these categories. To categorize the bleeding type, all 3 MRI sequences of each data set were read simultaneously. The interobserver agreement of this classification was assessed for each group and for all observers together.

**Statistical Analysis**

The interobserver agreement for the categorization of HT in the 3 MRI sequences according to our MRI-adopted classification was assessed by calculating Kendall’s coefficient of concordance \((W)\) for all observers together and for each group separately. Kendall’s coefficient is used for testing degrees of agreement among multiple observers (≥2) for a numerically coded ordinal response. \(W\) values of 0.6–0.8 indicated a substantial and of 0.81–1 an almost perfect degree of agreement.

**Results**

At the time of this study the 1000Plus database included 762 patients who had the clinical and radiological (positive DWI) diagnosis of an ischemic stroke and received a follow-up MRI 1 day after stroke onset according to the 1000Plus protocol. 186 of these 762 patients (24.4%) had received systemic thrombolysis. Some kind of HT was determined in 39 of the 186 thrombolysed patients (21%) in post-treatment MRI in any of the rated sequences (DWI, T2*GRE and FLAIR) by a radiologist experienced in stroke imaging. All of these 39 patients were included in the study.

**Interobserver Agreement of the Categorization of HTs**

Statistical analysis of the overall concordance between all observers using Kendall’s coefficient of concordance \((W)\) revealed a substantial interobserver agreement \((W = 0.79)\) for the new MRI-specific categorization of HTs. Assessing the two groups – residents and experts – separately, we detected an almost perfect degree of agreement between experts \((W = 0.81)\) and between residents \((W = 0.87)\). Looking at the interobserver agreement regarding the distinction between PH and HI, we saw an almost perfect agreement for all observers together \((W = 0.82)\) as well as for the individual groups of residents \((W = 0.91)\) and experts \((W = 0.82)\) (table 1).
**Fig. 1.** Examples of different types of HT in MRI in $T2^*$GRE, DWI and FLAIR according to the adapted ECASS classification in MRI. **a** HI-1: $T2^*$GRE with isolated hypointensities smaller than 10 mm which appear hypointense in DWI and hypo- or isointense in FLAIR. **b** HI-2: $T2^*$GRE confluent hypointensities >10 mm that appear also hypointense in DWI and hypo- or isointense in FLAIR. **c, d** Examples of PH-1 und PH-2 that can be seen in $T2^*$GRE as round hypointensities with central hyperintensities. In DWI, they show a low signal with hyperintense portions which appear hyperintense in FLAIR and are surrounded by a hypointense area between bleeding and edema.
Discussion

The purpose of this study was to test the reliability of a newly developed adaption of the established CT-based ECASS criteria for HT of a stroke in a combination of 3 MRI sequences (FLAIR, T2*GRE and DWI). For a possible comparison with older studies, we used the same categories of HT (HI-1, HI-2 and PH-1, PH-2) as the ECASS classification. HTs were categorized according to MRI based on size, shape, distribution and signal intensity of the hemorrhage in the different sequences. Our results show that this MRI-specific categorization of HT has a very good to almost perfect interobserver reliability for separating different types of HT. The interobserver agreement for the distinction between harmless HI and symptomatic PH which are negatively correlated with the patients' outcome [10–13] was almost perfect. These findings provide the first evidence of the reliability of our criteria and their potential use in clinical trials and clinical practice.

The gold standard for assessing intracerebral hemorrhage has been noncontrast CT. The most common classification of HT of a stroke is the ECASS classification which has been developed for CT imaging [3]. However, previous studies have shown that the reliability for bleeding detection after systemic thrombolysis in CT was poor to moderate and that the sensitivity and reproducibility of the detection of an HT in MRI in the T2*GRE sequence were higher than in CT, T2 and FLAIR MRI sequences within the first hours of HT [17, 18].

The reliability of the ECASS classification in MRI compared to CT has been assessed previously in a retrospective study in 43 patients with post-thrombolysis HT in CT or at least 1 of 3 MRI sequences (FLAIR, DWI and T2*GRE). A good inter- and intraobserver agreement of the classification was found for T2*GRE, DWI and CT and a moderate agreement was found for FLAIR. An excellent interobserver agreement for T2*GRE and a moderate agreement for CT, DWI and FLAIR could be archived for more severe bleedings classified as parenchymal hematoma. The authors concluded that the T2*GRE sequence is the most reproducible sequence for detection and categorization of post-thrombolysis HT and should be used to assess HTs in thrombolytic therapy trials [18].

However, due to the different imaging techniques and the higher sensitivity of T2*GRE sequence in detecting an HT within the first hours, the ECASS criteria cannot be transferred to a single MRI sequence. By definition, PH is associated with a space-occupying effect. PH-1 in MRI does not necessarily induce a mass effect, especially in hematomas located far from the ventricles. On the contrary, some petechial hemorrhages in HT are associated with a mass effect based on the underlying infarction [18]. Confluent hypointensities on T2*GRE do not necessarily represent hematomas. T2*-weighted images show a blooming effect of signal loss, and the use of the T2*GRE sequence as the only sequence for HT categorization could lead to an upward grading [17]. In the study by Renou et al. [18], both modalities (CT and MRI) were obtained from different patients, which did not make an assessment of the accuracy of the use of the ECASS criteria in MRI compared to the gold-standard CT possible. Direct comparisons of intracranial bleeding categorization in CT and MRI using T2*GRE performed in the same patients did confirm the risk potential for blood volume overestimation for this sequence due to susceptibility artifacts of T2* and thus introduced a bias in rating [16, 17]. DWI correlated best with the hematoma size on CT [16] and should be used for any size measurement. The increasing availability of 3-tesla MRI with higher sensitivity to hemosiderin deposits on T2*-weighted images and the introduction of susceptibility-weighted imaging leads to an even higher risk of rating harmless HI-2 as PH. The occurrence of PH but not HI within the first 36 h of ischemic stroke is associated with a poorer prognosis in most of the cases [10]. As the rate of HI and PH is used as a surrogate parameter for safety in clinical stroke trials, falsely rated parenchymal bleedings would endanger an adequate safety assessment in the clinical development of new drugs or endovascular procedures.
These studies show that the use of the CT-based ECASS criteria in MRI is not justified for the categorization of HT without adaptation. We believe that a combined evaluation of FLAIR, DWI and T2*GRE as in our criteria is best used for the detection and categorization of HTs in MRI. The results of our study show that experienced but also less experienced observers are able to produce a reliable categorization of different bleeding types with these newly developed criteria. In comparison to the study of Renou et al. [18], the combination of the 3 sequences used here showed a substantially higher interobserver agreement than the use of an individual sequence. An advantage of our study is that all patients underwent a pre- and post-treatment MRI scan, thus preventing a misinterpretation of chronic hemosiderin deposits in T2*GRE for acute HT when they had previously been detected on pre-treatment examination.

The use of MRI in clinical trials protects placebo patients from X-ray exposure without an individual benefit. Nevertheless, to validate the use of the presented classification, a comparison with the gold-standard native CT is required, which we could not provide in this retrospective study. The proof of the reproducibility of HT categorization in MRI and CT would allow to compare the safety profile of drugs or interventions tested in newer MRI-based studies to older CT data. To do so, we initiated a prospective study comparing the specificity and sensitivity of the adapted MRI classification to the ECASS classification in native CT in stroke patients 24–36 h after systemic thrombolysis with an additional MRI examination after routine post-thrombolysis CT.

Conclusion

The ECASS CT classification of HT was successfully modified for usage in MRI. It results in a substantial to almost perfect interobserver agreement and can be used for safety assessment in clinical trials.

Acknowledgments

The research leading to these results received funding from the Federal Ministry of Education and Research via the grant Center for Stroke Research Berlin (01 EO 0801).

Disclosure Statement

J.B.F. reports having received consulting, lecture and advisory board fees from Siemens, Bio-Imaging Technologies, Boehringer Ingelheim and Lundbeck. L.N., K.V., I.G., F.G.-D., R.G., D.G., C.K. and U.G. have no conflicts of interest to disclose.

References


