Cerebral microbleeds (CMBs) are small, round, or ovoid foci of signal loss (hypointensity) on MR sequences sensitive to paramagnetic iron. Pathologically, they represent small areas of hemosiderin deposition adjacent to small arteries.1 In patients with cerebral small vessel disease, CMBs are considered to be a useful marker for the presence and severity of the disease.2 CMBs restricted to the cerebral cortex or the cortico–subcortical junction (that is, in lobar locations) are a hallmark of the small vessel disease cerebral amyloid angiopathy (CAA), caused by deposition of amyloid in the media and adventitia of small arteries in the brain and leptomeninges.3 CAA is a common cause of lobar intracerebral hemorrhage, but may also have other manifestations, including vascular cognitive impairment, vasculitis,4 or transient neurological symptoms.5 The presence of lobar CMBs is a supporting criterion for the validated Boston criteria for CAA diagnosis.6 In CAA patients, the number of CMBs at baseline is a strong determinant of the risk of future intracerebral hemorrhage recurrence.7 Therefore, reliable identification of CMBs would aid in the diagnosis of CAA and determining the prognosis of the patient.

Susceptibility-Weighted Imaging Is More Reliable Than T2*-Weighted Gradient-Recalled Echo MRI for Detecting Microbleeds

Ah-Ling Cheng, MD; Saima Batool, DM; Cheryl R. McCreary, PhD; M.L. Lauzon, PhD; Richard Frayne, PhD; Mayank Goyal, MD; Eric E. Smith, MD, MPH

Background and Purpose—We investigated the sensitivity and reliability of MRI susceptibility-weighted imaging (SWI) compared with routine MRI T2*-weighted gradient-recalled echo (GRE) for cerebral microbleed (CMB) detection.

Methods—We used data from a prospective study of cerebral amyloid angiopathy (n=9; mean age, 71±8.3) and healthy non–cerebral amyloid angiopathy controls (n=22; mean age, 68±6.3). Three raters (labeled 1, 2, and 3) independently interpreted the GRE and SWI sequences (using the phase-filtered magnitude image) blinded to clinical information.

Results—In 9 cerebral amyloid angiopathy cases, the raters identified 1146 total CMBs on GRE and 1432 CMBs on SWI. In 22 healthy control subjects, the raters identified ≥1 CMBs in 6/22 on GRE (total 9 CMBs) and 5/22 on SWI (total 19 CMBs). Among cerebral amyloid angiopathy cases, the reliability between raters for CMB counts was good for SWI (intraclass correlation coefficient, 0.87) but only moderate for GRE (intraclass correlation coefficient, 0.52). In controls, agreement on the presence or absence of CMBs in controls was moderate to good on both SWI (κ coefficient ranged from 0.57 to 0.74 across the 3 combinations of rater pairs) and GRE (κ range, 0.31 to 0.70). A review of 114 hypointensities identified as possible CMBs indicated that increased detection and reliability on SWI was related to both increased contrast and higher resolution, allowing better discrimination of CMBs from the background and better anatomic differentiation from pial vessels.

Conclusions—SWI confers greater reliability as well as greater sensitivity for CMB detection compared with GRE, and should be the preferred sequence for quantifying CMB counts. (Stroke. 2013;44:2782-2786.)

Key Words: cerebral hemorrhage ■ magnetic resonance imaging ■ microbleeds
identification, based on a higher risk of false-positive identifications. Our literature review identified only 1 previous publication that reported good inter-rater agreement for CMBs on SWI and GRE in 20 patients; however, this study was done in a patient population where most did not have CMBs, and those that did had few CMBs per patient. To our knowledge, there are no previous studies of the reliability of SWI compared with GRE in patients with cerebral small vessel disease and larger numbers of CMBs, where interpretation of the images may be more complex.

We hypothesized that higher resolution SWI would be more sensitive for detecting CMBs than typical routine GRE in patients with cerebral small vessel disease, but that inter-rater reliability for microbleed count would be lower on SWI, reflecting the more complex images obtained using SWI. To test these hypotheses, we retrospectively analyzed SWI and GRE data from patients with CAA and controls participating in a prospective funded study.

Methods

Patient Population

This is a retrospective analysis of MRI data collected prospectively as part of a study of MRI biomarkers of CAA. Patients with CAA (n=9; mean age, 71±8.3 years) were recruited from hospital clinics. All patients had previous MRI evidence of multiple lobar CMBs, without deep hemispheric CMBs, consistent with Boston criteria for probable CAA. Patients were excluded if they had occipital intracerebral hemorrhage (because the research MRI protocol included visual task-related MRI) and were not fluent in English or had moderate-to-severe dementia (to enable reliable neuropsychological testing, which was part of the research protocol). Healthy stroke-free nondemented controls (n=21; mean age, 68±6.3 years) were recruited by community advertising.

Patients and controls provided written consent to participation. The study was approved by our institutional ethics review board.

MRI Acquisition

MRI was performed on a 3.0T MR scanner (Sigma VH1, GE Healthcare, Waukesha, WI). GRE and SWI sequences were obtained in the same MRI session.

GRE and SWI sequence parameters were chosen to mirror typical parameters used for those sequences in clinical practice and research, differing in both contrast mechanism (GRE versus SWI) as well as spatial resolution (higher for SWI). The GRE sequence parameters were similar to our institutional routine clinical protocol and as follows: repetition time 1200 ms, echo time 20 ms, field of view 24 cm, matrix 256×256, and slice thickness 3.5 mm with no gap. The SWI sequence parameters were as follows: repetition time 30 ms, echo time 20 ms, field of view 24 cm, matrix 256×256, slice thickness 2 mm, reconstructed to a matrix of 512×512 with slice thickness 1 mm. SWI were generated using an implementation of the methods described by Haacke et al. Magnitude and phase images were acquired and saved using a fully flow-compensated 3D GRE sequence. Phase images were unwrapped and filtered using an 80x80-pixel low-pass filter to remove the low-spatial frequency components of the background to generate a negative phase mask. The susceptibility images were generated by applying the phase mask, exponentiated by a factor of 3, to the magnitude images.

MRI Interpretation

MRI interpretation of CMBs was conducted independently by 3 raters: (1) rater 1 was a neurologist with 9 years of experience in research in cerebral small vessel disease, (2) rater 2 was a radiologist with 1 year of practice, and (3) rater 3 was a final year radiology resident. The GRE and SWI for each patient were read at least 2 weeks apart and in random order, to minimize the bias that could result from reinterpretation of the first image. Readings were blinded to clinical information.

CMB interpretation was guided by a recent review2 and 2 previously published rating scales,14,15 which were read and discussed by all raters before beginning the study. GRE number and location were recorded as 0, 1 or ≥2 CMBs using the brain observer microbleed scale, with the exception that CMBs were recorded as present versus absent, without including an uncertain category. Before beginning the study, all raters counted CMBs on a test set of GRE obtained from 15 lobar intracerebral hemorrhage patients not participating in the research study (of which 11/15 had ≥2 CMB, with a total of 76 CMBs across all patients), and reviewed the results together to ensure baseline consistency. After initial data analysis showed that inter-rater reliability for total microbleed count in CAA was poor on GRE for rater 1 and rater 3, both raters reviewed their CAA patient reads together to determine agreement for each individual CMB, a consensus diagnosis (true CMB versus mimic) for each putative CMB identified by either rater, and the most likely reasons for disagreement when disagreement was present. For patients with many putative CMBs (>20), CMBs were sampled on the basis of reviewing 4 slices per patient at predetermined anatomic levels (cerebellum at the level of the dentate gyrus, thalamus and basal ganglia, centrum semiovale, and immediately superior to the lateral ventricles). Based on literature and experience, reasons for failing to diagnose a CMB were categorized as follows: (1) low-signal intensity (that is, too faint), (2) unable to exclude confidently that the hypointensity was a vessel, (3) both (1) and (2), (4) difficulty in distinguishing the hypointensity from a nearby chronic hematoma, or (5) other factors or uncertain.

Statistical Analysis

Descriptive statistics on the total number of CMBs per patient are presented as medians and interquartile ranges. For CAA patients, who usually had many CMBs on both GRE and SWI, we expressed the difference between SWI and GRE as percentages of the number seen on GRE, with significance testing of these percentages by Wilcoxon signed-rank test. To determine inter-rater reliability of CMB counts in CAA, we used the intraclass correlation coefficient (ICC). Because control subjects had few or no CMBs, we calculated inter-rater agreement for the presence or absence of ≥1 CMBs using the κ coefficient. The κ coefficient represents an agreement rate adjusted for chance agreement and ranges from −1 (for perfect disagreement) to +1 (for perfect agreement). For both ICC and κ, 0.4 to 0.6 is generally considered to indicate moderate agreement, 0.6 to 0.8 is considered good agreement, and >0.8 is considered excellent agreement. Dice’s coefficient, which ranges from zero to 1, was used as a
In the 9 CAA cases, the raters identified up to 1146 total CMBs on GRE and up to 1432 CMBs on SWI (Table 1). Overall, CMB counts were highest for rater 1, intermediate for rater 2, and lowest for rater 3, except that rater 3 identified more CMBs on SWI than rater 2. Compared with GRE, rater 1 identified median 13% more (interquartile range, −6% to 49% more; P = 0.25), rater 2 identified median 30% more (interquartile range, 0% to 65% more; P = 0.04), and rater 3 detected median 184% more CMBs (interquartile range, 73% to 343% more; P = 0.008) on SWI.

Inter-rater reliability for CMB counts in CAA, on GRE and SWI, across all 3 raters and according to each pairwise combination of raters, are shown in Table 2. The ICC was higher for SWI (0.88; 95% confidence interval, 0.75–0.96) than for GRE (0.52; 95% confidence interval, 0.26–0.82), in part because there was poor agreement between rater 1 and rater 3 on GRE.

To explore reasons for disagreement between rater 1 and rater 3, we re-reviewed a total of 114 hypointensities identified as putative CMBs by either rater 1 or 3 on GRE or SWI. On GRE, both raters agreed that the hypointensity was either a putative CMB or was not a CMB in 74/114 (65%; Figure 2); the Dice coefficient was 0.69. Most disagreement was in cases where rater 1 identified a putative CMB but rater 3 did not (38/114, 33%), usually because the hypointensity was judged by rater 3 to possibly be a vessel (28/38, 74%). After consensus re-review of these 38 cases by both raters, 27 (71%) were judged to be true CMBs. On SWI, both raters agreed that the hypointensity was either a putative CMB or was not a CMB in 72/114 (63%; Figure 3); the Dice coefficient was 0.75. As with GRE, disagreement was mostly in cases where rater 1 identified a putative CMB but rater 3 did not (37/114, 32%), usually because the hypointensity was judged by rater 3 to possibly be a vessel (27/37, 73%). After consensus re-review of these 37 cases by both raters, 29 (78%) were judged to be true CMBs. Compared with GRE, disagreement on SWI was less frequently related to faintness of the hypointensity (15/38, 40% on GRE versus 8/37, 22% on SWI). Among the 114 hypointensities identified as putative CMB on either GRE or SWI, raters 1 and 3 identified 24 more true CMBs by consensus on SWI (95/114 putative CMBs) than on GRE (71/114); retrospectively, the reason was thought to be because of higher spatial resolution, with more confident discrimination of spherical or ovoid CMBs from linear vessels, or higher signal contrast on SWI (Figure 4).

In the 21 controls, the raters identified up to 9 total CMBs on GRE and up to 19 CMBs on SWI. Most controls had no CMBs. On GRE, rater 1 identified CMBs in 3/21, rater 2 in 6/21, and rater 3 in 5/21. The pairwise \( \kappa \) values for CMB presence or absence on GRE ranged from 0.31 to 0.70. On SWI, rater 1 identified CMBs in 4/21, rater 2 in 5/21, and rater 3 in 5/21. The pairwise \( \kappa \) values for CMB presence or absence on SWI ranged from 0.57 to 0.74.

Because SWI has higher sensitivity for CMBs, it is possible that patients with probable CAA by Boston criteria could be reclassified as unlikely to be CAA, if additional CMBs were detected in nonlobar locations. However, only 1 of 9 CAA patients had 2 possible nonlobar CMBs detected by rater 2 on SWI but not GRE, which were not identified as CMBs by the other raters.

<table>
<thead>
<tr>
<th>Table 1. Number of CMBs Identified by Each Rater on GRE and SWI in 9 CAA Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequence</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>GRE</td>
</tr>
<tr>
<td>SWI</td>
</tr>
</tbody>
</table>

| CA indicates cerebral amyloid angiopathy; CMBs, cerebral microbleeds; GRE, T2*-weighted gradient-recalled echo; and SWI, susceptibility-weighted imaging. |

<table>
<thead>
<tr>
<th>Table 2. Intraclass Correlations for Total Numbers of Microbleeds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequence</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>GRE</td>
</tr>
<tr>
<td>SWI</td>
</tr>
</tbody>
</table>

\( r \) indicates confidence interval; GRE, T2*-weighted gradient-recalled echo; and SWI, susceptibility-weighted imaging.
other 2 raters. Similarly, it is possible that SWI could detect additional CMBs in asymptomatic controls, including the pattern of multiple lobar CMBs potentially suggestive of CAA. However, there was only 1 instance where 1 rater (rater 2) identified a single extra lobar-only CMB.

**Discussion**

In this study, we show that SWI has higher sensitivity for microbleed detection and that, contrary to our prespecified study hypothesis, reliability of microbleed counts was better, not worse, on SWI than GRE. Additionally, we found that the sensitivity and reliability of SWI for CMB detection depended on rater experience. The rater with the most experience (rater 1) identified more CMBs in CAA patients on either GRE or SWI and identified the fewest extra CMBs on SWI (only 13% more). By contrast, the rater with the least experience (rater 3) was much less likely to identify CMBs on GRE, but identified the most extra CMBs on SWI. Inter-rater reliability was only moderate on GRE but excellent on SWI.

A systematic re-review of GRE and SWI hypointensities suggested that the increased signal contrast and higher spatial resolution of SWI increased the confidence of the less experienced rater that the hypointensity in question was indeed a CMB, whereas the more experienced rater was more confident that faint nonlinear hypointensities on GRE represented a true CMB. The consensus was that most hypointensities identified by rater 1 as CMB were indeed true CMB, suggesting that most were true-positive CMBs, not false-positive misidentification of CMB mimics.

These results add to the current literature by showing that inter-rater reliability for CMB identification can be better on SWI than on GRE. Previous studies of CMBs on GRE report ICCs of 0.91 to 0.97 for total CMB counts, and values generally between 0.40 to 0.75 for CMB presence or absence. By contrast, the ICC on GRE in our study was lower than for these other studies, probably reflecting differences in the populations studied. Our 9 CAA patients had a very large number of CMBs (Table 1) that exceeded the numbers of CMBs per patient reported in other studies; therefore, interpretation of the images was time-consuming and difficult, which might have affected the reliability. Raters completed a CMB identification training set (based on GRE only, not SWI) before beginning the study, and used a CMB report form and method previously reported to increase inter-rater reliability. Despite these steps, inter-rater reliability for some raters and pairs of sequences was only moderate. Our findings emphasize the need for standardized assessments with training sets, particularly when using raters with little prior experience in CMB interpretation.

Previous studies suggest that CMBs can be identified better on SWI than GRE; however, these studies are mostly limited to illustrative case examples or small series of patients with few CMBs. A previous study of 14 CAA patients found that higher MRI field strength (3T compared with 1.5T) and SWI (versus GRE) were associated with better CMB contrast-to-noise ratio. Inter-rater reliability for CMB count on SWI was high (0.91) but was not compared with inter-rater reliability for CMB count on GRE. A single rater found 3x as many CMBs on SWI as on GRE, but this was based on only 3 patients.

In 141 dementia clinic patients, CMBs were detected more frequently on SWI compared with GRE (40% versus 23%) and higher counts were seen on SWI (284 versus 219, a 30% increase). Reliability among 3 raters was reported in a subset of 20 patients, with good agreement for presence or absence of CMB on both GRE (weighted Cohen’s κ, 0.83) and SWI (weighted Cohen’s κ, 0.86). However, characteristics of the 20 patients were not described in detail, including the number with CMBs or the number of CMBs per patient, nor were the individual rater’s readings reported. The number of CMBs per patient was low—among patients with CMBs, the median number was 1 on GRE and 2 on SWI. Therefore, the reliability estimates from this study are mostly applicable to patient populations with relatively low numbers of CMBs, where the most common question is whether CMBs are present or not. By contrast, our study reports in detail on inter-rater reliability in a population with high numbers of CMBs, where the most
common goal is to assess the total number of CMBs reliably. For this purpose, our study data suggests that using SWI may be superior to GRE. However, the clinical relevance of identifying different numbers of CMBs remains relatively unknown. Our study has limitations, which should be considered when interpreting the findings. Although the number of patients was small, the number of CMBs reviewed was high (>1000). Our results were obtained on a 3T MRI and may not directly translate to imaging at 1.5T, as field strength is known to affect sensitivity for CMB detection. Because our MRI acquisition differed in both signal contrast mechanism (GRE versus SWI) and spatial resolution (lower on GRE than SWI), we are unable to determine their independent effects on sensitivity and reliability of CMB detection. We chose this study design, comparing a typical clinical GRE sequence at our institution to a typical research SWI sequence, to make the results most directly relevant to our clinical and research practice. Although we selected raters with a range of experience with CMB detection, our findings may not necessarily generalize to other groups of raters.

Based on our findings, we have adopted SWI as the sequence of choice at our institution for identifying CMBs caused by cerebral small vessel disease, based on its greater sensitivity and higher reliability compared with routine GRE. Although our evidence supports use of SWI as the current best standard for CMB research, we note that research methods in development may allow further improvements in the sensitivity, reliability, and efficiency of CMB detection. These developing methods include automated or semiautomated computer algorithms for CMB identification and new acquisition and MRI sequence processing algorithms, including quantitative SWI.23

Sources of Funding
The study was funded by the Canadian Stroke Network, Heart and Stroke Foundation of Canada, and the Alzheimer Society of Canada. Dr McCreary receives salary support from the Kathy Taylor Chair in Vascular Dementia Research at the University of Calgary (chair holder: Dr Eric Smith). Dr Lauzon receives salary support from the Hopewell Professorship in Brain Imaging. Dr Frayne is a Canada Research Chair. Dr Smith is supported by external salary awards from Alberta Innovates – Health Solutions and a New Investigator Award from the Canadian Institutes for Health Research.

Disclosures
Dr Frayne and Smith are investigators on a grant funded by Alberta Innovates – Health Solutions and Pfizer to develop new methods for microbleed detection using susceptibility-weighted imaging. The other authors have no conflicts to report.

References
Susceptibility-Weighted Imaging is More Reliable Than T2*-Weighted Gradient-Recalled Echo MRI for Detecting Microbleeds
Ah-Ling Cheng, Saima Batool, Cheryl R. McCreary, M.L. Lauzon, Richard Frayne, Mayank Goyal and Eric E. Smith

Stroke. 2013;44:2782-2786; originally published online August 6, 2013;
doi: 10.1161/STROKEAHA.113.002267

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://stroke.ahajournals.org/content/44/10/2782