NeuroImage: Clinical 15 (2017) 274–283

Contents lists available at ScienceDirect

NeuroImage: Clinical

journal homepage: www.elsevier.com/locate/ynicl

Improved cerebral microbleeds detection using their magnetic signature on T2*-phase-contrast: A comparison study in a clinical setting

Takoua Kaaouana[a,b,c,d,e,f,*], Anne Bertrand[a,b,c,d,e,g], Fatma Ouamer[g], Bruno Law-ye[a,g], Nadya Pyatigorskaya[a,g], Ali Bouyahia[a,b,c,d,e,f], Nathalie Thiery[h], Carole Dufouil[h,i], Christine Delmaire[j], Didier Dormont[a,b,c,d,e,g], Ludovic de Rochefort[k], Marie Chupin[a,b,c,d,e]

[a] Sorbonne Univ, UPMC Univ Paris 06, UM 75, ICM, F-75013 Paris, France
[b] Inserm, U1127, ICM, F-75013 Paris, France
[c] CNRS, UMR 7225, ICM, F-75013 Paris, France
[d] ICM, Paris, France
[e] Inria, Aramis project-team, Centre Paris-Rocquencourt, France
[f] CATI, Paris, France
[g] Neuroradiology, CHRU Pitié Salpêtrière, Paris, France
[h] CIC-EC7, CHU de Bordeaux, Pole de Sante Publique, Bordeaux, France
[i] Univ. Bordeeaux, ISPED, Centre INSERM U897-Epidemiologie-Biostatistique & CIC-EC7, Bordeaux, France
[j] Neuroradiology, CHRU Roger Salengro, Lille, France
[k] Univ. Paris-Saclay, CNRS IR4M UMR8081, Orsay, France

ARTICLE INFO

Article history:
Received 27 April 2016
Received in revised form 31 July 2016
Accepted 3 August 2016
Available online 9 August 2016

Keywords:
Microbleeds
Magnetic susceptibility
Phase MRI
SWI

ABSTRACT

Introduction/purpose: In vivo detection of cerebral microbleeds (CMBs) from T2* gradient recalled echo (GRE) magnitude image suffers from low specificity, modest inter-rater reproducibility and is biased by its sensitivity to acquisition parameters. New methods were proposed for improving this identification, but they mostly rely on 3D acquisitions, not always feasible in clinical practice. A fast 2D phase processing technique for computing internal field maps (IFM) has been shown to make it possible to characterize CMBs through their magnetic signature in routine clinical setting, based on 2D multi-slice acquisitions. However, its clinical interest for CMBs identification with respect to more common images remained to be assessed. To do so, systematic experiments were undertaken to compare the ratings obtained by trained observers with several image types, T2* magnitude, Susceptibility Weighted Imaging reconstructions (SWI) and IFM built from the same T2*-weighted acquisition.
Materials/methods: 15 participants from the MEMENTO multi-center cohort were selected: six subjects with numerous CMBs (20 ± 6 CMBs), five subjects with a few CMBs (2 ± 1 CMBs) and four subjects without CMB. 2D multi-slice T2* GRE sequences were acquired on Philips and Siemens 3T systems. After pilot experiments, T2* magnitude, Susceptibility Weighted Imaging (SWI) minimum intensity projection (mIP) on three slices and IFM were considered for the rating experiments. A graphical user interface (GUI) was designed in order to consistently display images in random order. Six raters of various background and expertise independently selected "definite" or "possible" CMBs. Rating results were compared with respect to a specific consensus reference, on both lesion and subject type points of view.
Results: IFM yielded increased sensitivity and decreased false positives rate (FPR) for CMBs identification compared to T2* magnitude and SWI-mIP images. Inter-rater variability was decreased with IFM when identifying subjects with numerous lesions, with only a limited increase in rating time. IFM thus appears as an interesting candidate to improve CMBs identification in clinical setting.

1. Introduction

Cerebral microbleeds (CMBs), described as small foci of chronic blood products within brain parenchyma (Greenberg et al., 2009), were first seen in patients with small vessel diseases on specific MRI sequences that are sensitive to magnetic susceptibility. Correlation of these radiological findings with histopathological studies has generated

* Corresponding author at: Centre de Recherche de l'Institut du Cerveau et de la Moelle Epinière, UMR 7225 CNRS - UMRS 975 INSERM, Hôpital de la Salpêtrière, 47, Bd de l'Hôpital, 75013 Paris Cedex, France.
E-mail address: kaaouana.takoua@gmail.com (T. Kaaouana).

http://dx.doi.org/10.1016/j.nicl.2016.08.005

considerable interest (Cordonnier, 2011). CMBs are now commonly reported in the general population as well as in patients with specific disorders (Conijn et al., 2011). CMBs' prevalence is highly variable among reports: from 47% to 80% in patients with intra-cerebral hemorrhage (Lee et al., 2004; Naka et al., 2004), from 8% to 71% in patients with ischemic stroke (Naka et al., 2004; Tsushima et al., 2003), from 17% to 46% in patients with cognitive decline and dementia (C. Cordonnier et al., 2006; Hilal et al., 2014) and about 20% in healthy elderly population (Cordonnier, 2011). Deep brain CMBs are associated with hypertensive arteriopathy (HTA), while those close to the cortex are associated with cerebral amyloid angiopathy (CAA), although they may also be present in patients with isolated cerebral small vessel disease (Park et al., 2013).

CMBs identification using MRI remains complicated (Cordonnier, 2011; Greenberg et al., 2009). As made of hemosiderin, they can be described as strongly super paramagnetic iron–storage complexes (Cordonnier, 2011), whereas the surrounding brain parenchyma is diamagnetic. Thus, this magnetic susceptibility difference with surrounding brain parenchyma makes CMBs appear as magnetic inclusions, causing a local magnetic field inhomogeneity similar to the one that would be created by a unit dipole. At the voxel level, this inhomogeneity leads to intra-voxel phase dispersion and strong T2*-contrast. CMBs' detection is thus commonly based on Gradient Recalled Echo (GRE) T2*-weighted magnitude images, in which they appear as areas of signal loss. However, their appearance on these sequences is sensitive to imaging parameters such as echo time (TE) and B 0 field strength. Furthermore, blood vessels and cerebral micro-calcifications (CMCs) also have strong T2* effects and can be misidentified as CMBs. Physiologic calcifications commonly found in specific areas (e.g. pineal gland, choroid plexus, basal ganglia) can easily be identified but smaller calcifications at unexpected localization sometimes require a CT scan (Yamada et al., 1996).

Specific GRE-based solutions have been proposed to address these issues or related ones, including different acquisition protocols and/or different signal processing methods such as Susceptibility Weighted Imaging (SWI) (Cheng et al., 2013; Goos et al., 2011; Nandigam et al., 2009; Vernooij et al., 2008), Enhanced Susceptibility weighted angiography (ESWAN) (Guo et al., 2013), Quantitative susceptibility mapping (QSM) (Klohs et al., 2011; Liu et al., 2012; McAuley et al., 2010), internal field maps (IFM) (Guo et al., 2013; Kaaouana et al., 2015). SWI has already been considered for the diagnosis of CMBs. It is based on combining phase and magnitude images from 3D high resolution GRE acquisitions and aims at increasing detection sensitivity for paramagnetic structures such as veins or hemorrhages (Goos et al., 2011; Haacke et al., 2009; Haacke and Reichenbach, 2011; Nandigam et al., 2009; Reichenbach et al., 1997). A comparison between SWI and GRE T2* reported that conventional GRE T2* magnitude missed 67% of CMBs compared to SWI (Nandigam et al., 2009; Cheng et al., 2013; Vernooij et al., 2008). The QSM reconstruction technique, aiming at quantifying susceptibility, should make it possible to estimate the real lesion load. A validation study on 3D multi-echo GRE T2* acquisitions on ten patients suspected of having experienced a stroke reported that the total susceptibility (TS) of CMBs was more consistent than CMBs size measurement (Liu et al., 2012). Nevertheless, SWI, ESWAN and QSM techniques have been designed and evaluated with 3D multi-echo GRE T2* acquisitions which may not be available in large cohorts. Internal field maps (IFM) can be computed with standard parameters from phase images of routine 2D T2* GRE single echo acquisitions. A method based on 2D harmonic filtering (2DHF) (Kaaouana et al., 2015) has been shown to allow generating appropriate internal field maps for discriminating CMBs and visualizing the dipole field patterns created by CMBs; this magnetic signature could be used in CMBs' characterization.

Overall, CMBs' detection has limited reproducibility and is observer dependent (de Bresser et al., 2013; Charidimou et al., 2012). This can be explained by confounding structures and artifacts as well differences in acquisition settings between studies. Previous studies comparing rating performance for advanced images vs standard magnitude images did in fact combine the advantages of 3D acquisitions with those of the new pre-processing techniques. Here, we aimed at evaluating the specific advantages of advanced image processing technique for CMB's identification by trained raters. In fact, 3D GRE T2*-weighted acquisitions may not always be feasible in clinical setting and 2D acquisitions make it possible to ensure uniformity in multi-center clinical studies while guaranteeing reasonable acquisition time. Thus, routine 2D multi-slice GRE acquisitions were used as inputs in this study from which several types of reconstructions were derived: magnitude images, SWI-like images, and preprocessed phase images (IFM).

The paper is organized as follows. In the first part, the dataset used for this comparison study is presented, followed by a description of the advanced pre-processing techniques and comparison experiments. Results are then detailed in the second part, regarding the reliability of the reference built specifically for this study and the performance of the method on both “single lesion” and “subject” points of view.

2. Material and methods

2.1. Evaluation dataset

The evaluation dataset was extracted from data acquired for the ongoing French national cohort named MEMENTO (Chene et al., 2014). MRI and PET acquisitions, provided by a network of 24 centers with MRI systems from different manufacturers, models and field strengths, are monitored by the CATI, the French National Platform for Multi-center Neuroimaging Studies (http://www.cati-neuroimaging.com/). A sub-sample of the first 382 MEMENTO participants were evaluated for CMBs (Kaaouana et al., 2015), and 77 subjects with CMBs were identified. This prevalence of 20% for this population (mean age was 55) was consistent with the prevalence reported for elderly subjects (Cordonnier, 2011). For the current evaluation study, 15 subjects were selected in order to analyze six subjects with numerous CMBs [13–30 CMB], five subjects with few CMBs [1–4 CMB] and four subjects without CMB. MRI data for these 15 subjects were acquired on either Siemens (four centers, Verio systems, seven subjects) or Philips (three centers, Achieva systems, eight subjects) 3T systems. The acquisition protocol was described in (Kaaouana et al., 2015).

2.2. Methods

CMBs are made of hemosiderin and are detected on GRE images because of the local magnetic susceptibility variation they yield. Besides, GRE phase images are proportional to magnetic field variations and thus sensitive to local susceptibility variations. Both magnitude and phase images will thus be considered in this study.

However, phase image analysis for local field variations is not straightforward. In fact, phase wraps appear on the images because phase is defined in the $[-\pi, \pi]$ interval and local variations are hidden in large scale field variations resulting from background effects, dominated by the magnetic susceptibility sharp edge of air-tissue interfaces.

The extraction of relevant internal field information thus requires two preliminary steps: phase unwrapping and background field removal (Haacke and Reichenbach, 2011). These two pre-processing steps are embedded in the two methods chosen for this study, SWI and IFM, as described below.

2.3. Susceptibility Weighted Imaging (SWI)

SWI relies on combining phase and magnitude images (Haacke and Reichenbach, 2011), phase information being used to enhance blood-related contrast on magnitude image (e.g. veins). Phase images are first high-pass filtered (HPF) to extract local information and then transformed in a phase mask with values in the [0; 1] interval. In this study, the HPF phase image was obtained through the following steps. The complex-valued image was first generated from magnitude and phase

images. It was then low-pass filtered slice by slice with a two dimensional Gaussian filter in Fourier domain. The HPF phase image was then estimated as the phase component of the ratio between complex-valued and low-pass-filtered images. The HPF phase image was then transformed in a consistent phase mask which was then applied N times on the original magnitude image to highlight voxels with a high phase value. The standard-deviation of the Gaussian filter, σ, and the N parameter were empirically chosen in order to optimize the setting for CMBs detection on 2D data. The σ parameter was set to 36 pixels, as recommended in (Haacke and Reichenbach, 2011), and N was set to 8.

2.4. Internal field map (IFM)

A 2D-based method for phase unwrapping and harmonic filtering, based on solving Poisson equation (Song et al., 1995), has been proposed as an efficient mean for obtaining the IFM (Kaaouana et al., 2015) on 2D acquisitions. This method was shown to solve the potential slice-to-slice phase inconsistency that may occur in 2D multi-slice T2* GRE datasets. The magnetic field observed inside the brain, B, can be decomposed as the sum of the magnetic field due to internal sources, B_{in}, and the one induced by external sources, B_{out}. From Maxwell's equations, B_{out} is harmonic inside the brain ($\Delta B_{out} = 0$), resulting in $\Delta B = \Delta B_{in}$ (Δ denotes the Laplacian). Consequently, field variations due to external sources can be filtered out through a second order derivative, followed by a second order integration using adequate boundary conditions. In the process, the Laplacian of the field B_{in}, which locally depends on susceptibility distribution, is set to 0 outside the brain to remove external susceptibility effects. Paramagnetic dot-like inclusions such as CMBs appear as a dipolar field on the resulting IFM (see Fig. 1 and (Kaaouana et al., 2015)). CMBs can be better discriminated by this magnetic signature.

2.5. Evaluation experiments

In order to determine the influence of the type of image for CMBs' identification, rating was performed by several raters in different conditions. A pilot experiment was first carried-out in order to determine the optimal settings for the rating experiments (see Appendix 1 for more details). Three types of images were finally considered as good candidates for CMBs rating in clinical setting: T2* magnitude image, SWI-mIP image (mIP being done on three slices with a resulting slab thickness of 7.5 mm) and IFM image.

2.6. Experiment

Rating was performed independently by six raters (one session each) with various levels of expertise: a trained clinical research assistant, a trained engineer, two junior neuroradiologists and two senior neuroradiologists. For each rater, all scans were scored in a single session. All observers were blind to image type, clinical information and other ratings. In order to facilitate detection, interactive visualization tuning was embedded in a specific Graphical User Interface (GUI) built with GUIDE in MATLAB (see Fig. 6 Appendix 1). As in MARS (S. M. Gregoire et al., 2009) and BOMBS (Charlotte Cordonnier et al., 2009) rating scales, a certainty score was included and CMBs could be categorized as “definite” or “possible”. The aim of this categorization was both to facilitate the rating in case of uncertainty and to differentiate the amount of variability that came from clear and dubious CMBs. CMBs were defined as small round areas of signal loss on axial slices without 3D connectivity that would characterize vessels on sequences that are sensitive to magnetic susceptibility. Their size may vary from 2 to 10 mm (Greenberg et al., 2009).

2.7. Building-up of the reference

After the series of ratings was completed, a reference was built by two trained neuroradiologists, one of whom did not participate to the comparison experiments. All the CMBs that were detected by any rater on any image during comparison experiments were reassessed independently by each trained neuroradiologist. Here, the three image types were displayed simultaneously through a specific GUI (reference GUI). The resulting two sets of 15 reference images were then combined to create a set of 15 consensus images using a given explicit scoring rule (Table 1). Briefly, if a lesion was identified by both observers and at least one observer considers it as definite, then it was rated as “definite CMB”

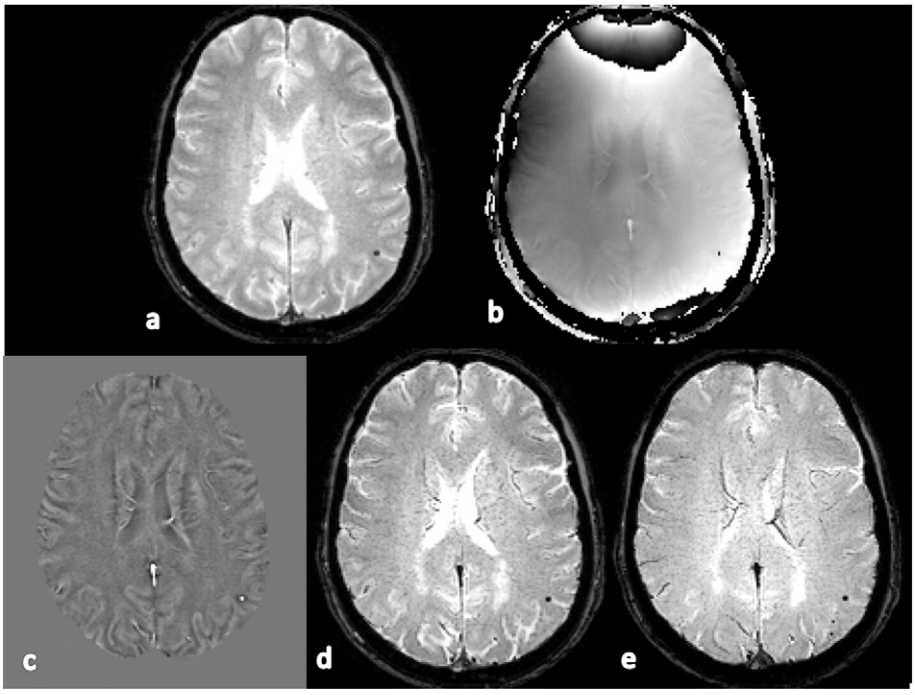

Fig. 1. Magnitude image (a) raw phase image (b), IFM (c), SWI (d), SWI-mIP, minimum intensity projection on three consecutive slices (e) are displayed.

in the consensus. If it was identified by only one observer as "possible" then it was discarded from the reference. If it was identified by only one observer as "definite", then it was considered as a "possible CMB".

Because of the well-known inter-rater variability when rating CMBs, the reliability of the consensus reference needed to be evaluated, to ensure that the consensus rating could be considered as a meaningful reference. Rating experiments were then analyzed on a lesion-based point of view by comparing the decision of each rater for each lesion in the reference. Two types of identification were analyzed: 1. all detected CMBs (either definite or possible); 2. CMBs detected as definite only. Furthermore, in order to assess clinical usefulness, ratings and reproducibility analyses were carried out on a subject-type point of view, by classifying patients in three groups: no CMB (G1), few CMBs (G2, less than 10 CMBs) and numerous CMBs (G3, more than 10 CMBs). Finally, rating durations were also analyzed, in order to evaluate clinical feasibility of IFM-based rating with respect to more standard T2* magnitude and SWI-mIP.

3. Results

The specificity (defined as $TN/(TN+FP)$) cannot be determined since *TN* (true negatives) is not a meaningful measure in CMBs detection. Here, the performance of each rater is characterized through the false positives (*FP*) rate (FPR), false negatives (*FN*) and true positives (*TP*). Dice coefficients ($\frac{2TP}{2TP+FN+FP}$), which combines *TP*, *FP* and *FN*, were also calculated for each rater.

3.1. Reference

In order to evaluate the reliability of the consensus reference, detected CMBs were compared between expert raters. Results are given in Table 2, through the number of CMBs (all or definite CMBs) detected by each of the two expert raters, the number of CMBs detected by both raters (overlap) or by a single rater (discrepancy) and the number of CMBs on the consensus computed by the scoring rules described in Table 1.

Cohen's kappa coefficient was computed to compare between raters the number of CMBs detected for each subject. When considering all lesions, the kappa value was 0.5 (p-value = 0.01) and for definite lesions only, the kappa value was 0.54 (p-value = 0.02).

Overall, the agreement on CMBs detection was satisfactory. Apart from one subject (subject 3 with 13 discrepant CMBs), the discrepancy between both raters was negligible (median for "no CMB": 0, "few CMBs": 0 and "numerous CMBs": 2.5). Examples of between-raters discrepancies are illustrated in Fig. 2.

3.2. Rating results: lesion-based point of view

Blind ratings were analyzed with respect to the consensus reference, in order to evaluate the performance of each image type independently for lesion detection. Because of the large variability in lesion number between subjects, TP, FP and FN were computed for each rater over all the CMBs detected on all subjects. Note that "all" refers to the union of "definite" and "possible" CMBs. The overall count of CMBs detected by each rater is given in Table 3 for the three image types (T2* magnitude, SWI-mIP and IFM). Total numbers of CMBs, TP, FN and FP are detailed for each rater.

Table 1
Reference consensus building up, scoring rules.

		Exp.1		
		no	p	d
Exp.2	no	no	no	p
	p	no	p	d
	d	p	d	d

"no" for discarded lesion, "p" for "possible CMB" and "d" for "definite CMB".

Table 2
Reference building-up: CMBs detected by expert raters and consensus result.

	Expert.1		Expert.2		Overlap/discrepancy			Consensus	
Subject	d	All	d	All	d/d	d/p	d/no	d	All
2	0	2	0	0	0	0	0	0	0
12	0	1	0	0	0	0	0	0	0
13	0	1	0	0	0	0	0	0	0
15	0	4	0	0	0	0	0	0	0
6	0	0	1	1	0	0	1	0	1
5	0	2	1	1	0	1	0	1	1
10	2	5	2	2	2	0	0	2	2
1	4	7	1	2	1	1	2	2	4
8	3	6	4	4	3	1	0	4	4
9	12	12	12	16	11	0	2	11	13
11	15	15	15	15	14	0	2	14	16
7	16	17	12	16	11	3	3	14	17
3	17	18	15	17	9	1	13	10	23
4	19	24	16	24	12	6	5	18	24
14	30	30	12	36	12	18	0	30	30
Total	118	144	91	134	75	31	28	106	135

"d" refers to "definite" CMBs and "all" to the sum of definite and possible CMBs. "d/d" is the number of CMBs detected as definite by both raters. "d/p" is the number of CMBs detected as definite by one rater and possible by the other, "d/no" is the number of CMBs detected as "definite" by one observer while not detected by the other.

IFM yields a slight improvement of the overall ratings with respect to T2* whereas SWI-mIP yields systematic higher numbers of false positives. TP were lower for SWI than for T2* and IFM (78 vs 88–89). Both FN and FP were higher for SWI (FN: 57 vs 48–46, FP: 34 vs 16–9). Furthermore, ranges for TP, FN and FP were narrower for IFM than T2* (TP: 32 vs 54, FN: 32 vs 54, FP: 73 vs 100). Dice coefficients, given for each rater in Table 3, were slightly higher for IFM than T2* and SWI-mIP (0.74 vs 0.69–0.61).

When comparing performances between raters, we note that two raters tend to underestimate the number of CMBs (CRA and Exp.Junior1) and one tends to largely overestimate it (Exp.Junior2). Interestingly, this difference is lessened by the use of IFM with respect to the other two image types (total count range: 99 for IFM vs 150 for T2* and 198 for SWI-mIP).

In order to better understand the detection pattern between raters, ratings for definite CMBs were also analyzed. Detailed results are given in Table 4.

Overall, the results are in accordance with Table 3. More specifically, IFM yields a higher number of definite CMBs than the other two image types (81 vs 72 and 66 for T2* and SWI-mIP, respectively) and smaller number of FN (25 vs 34 and 41) and FP (4 vs 9 and 19). Value ranges follow the same trend (TP: 23 vs 39 and 16, FN: 23 vs 39 and 46, FP: 14 vs 20 and 70). Dice values for definite CMBs were also higher for IFM than T2* and SWI-mIP (0.83 vs 0.77–0.66). IFM thus seems to mitigate inter-rater variability for definite lesions (range for total number: 36 vs 55 and 108).

3.3. Rating results: subject-type point of view

In order to evaluate the three image types with respect to their clinical usefulness, rating results were evaluated for each subgroup ("no CMB", "few CMBs", "numerous CMBs") (Table 5).

The overall trend is confirmed for the subgroups, FPR being decreased by IFM with respect to T2* and SWI-mIP ((0 vs 2 and 2 for G1), (1 vs 4 and 5 G2) and (7 vs 13 and 27 for G3)), even if the performance is more balanced between methods for "no CMB" and "few CMBs" groups.

To investigate clinical relevance, image types were compared regarding to their ability to correctly classify the subjects in the three groups of interest. The number of correctly classified subjects for each rater and each image type is given in Table 6.

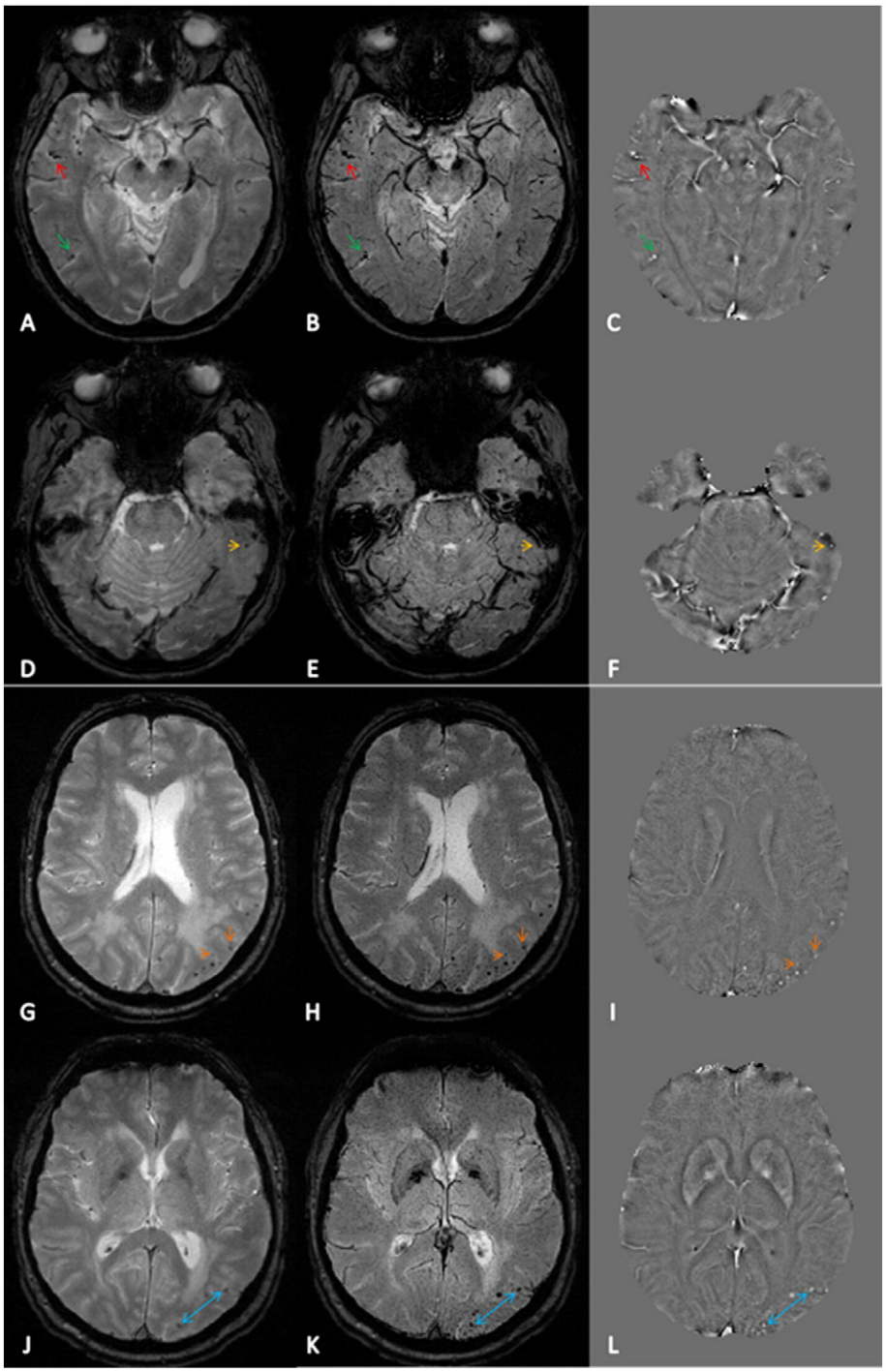

Fig. 2. Between-raters discrepancies during reference building-up; first column: magnitude images, second column: SWI-mIP images, third column: internal field map. First two rows: discrepancy cases from subject 3. Lesions shown here by red and green arrows are doubtful due to their shape that can be seen either as two adjacent round CMBs or as a relatively linear structure like a blood vessel. The CMB showed by the yellow arrow is very close to susceptibility artifact. Last two rows: discrepancy cases from subject 14. CMBs pointed by orange and blue arrows may have been ambiguous because of low contrast.

Overall T2* magnitude and IFM classify better than SWI-mIP (median value: G1: 3/4 and 4/4 vs 2.5/4, G2: 2/5 and 2/5 vs 2/5, G3: 5.5/6 and 5/6 vs 4.5/6). IFM yields more similar results between raters than T2* magnitude and SWI-mIP (ranges: G1: 2 vs 3 and 3, G2: 2 vs 2 and 1, G3: 1 vs 4 and 4).

Table 7 indicates recorded rating duration. Ratings from IFM took longer in most cases, but the relative difference decreased from participants with no CMBs to participants with numerous CMBs. In the worst case, it took only 1.2 times longer. Interestingly, for all image types, it took less time to do the rating for participants with few CMBs than those with no CMB.

4. Discussion

We presented here a comparison of CMBs detection performance when using three different kinds of images built from the same 2D GRE T2* weighted dataset: T2* magnitude, SWI-mIP and IFM. Blind ratings by six raters were evaluated with respect to a reference built from a

Table 3
Total number of detected CMBs and comparison with the reference: TP, FN, FP and Dice.

	Total number			∑ TP			∑ FN			∑ FP			Dice coef		
	T2*	SWI-mIP	IFM	T2*	SWI-mIP	IFM	T2*	SWI-mIP	IFM	T2*	SWI-mIP	IFM	T2*	SWI-mIP	IFM
CRA	63	70	82	51	42	73	84	93	62	12	28	9	0,52	0.41	0.67
Exp.Junior1	82	81	89	74	65	86	61	70	49	8	16	3	0,68	0.60	0.77
Exp.Senior1	132	113	107	99	78	95	36	57	40	33	35	12	0,74	0.63	0.79
Trained. Ing	100	115	101	83	82	92	52	53	43	17	33	9	0,71	0.66	0.78
Exp.Senior2	107	119	84	92	78	78	43	57	57	15	41	6	0,76	0.61	0.71
Exp.Junior2	213	268	181	105	99	105	30	36	30	108	169	76	0,6	0.49	0.66
Median	**104**	**114**	**95**	**88**	**78**	**89**	**48**	**57**	**46**	**16**	**34**	**9**	**0.69**	**0.61**	**0.74**
Min	63	70	82	51	42	73	30	36	30	8	16	3	0.52	0.41	0.66
Max	213	268	181	105	99	105	84	93	62	108	169	76	0.76	0.66	0.97

Median is highlighted in bold character.

Table 4
Total number of CMBs identified as "definite" and comparison with the reference.

	Total number			∑ TP			∑ FN			∑ FP			Dice Coef		
	T2*	SWI-mIP	IFM	T2*	SWI-mIP	IFM	T2*	SWI-mIP	IFM	T2*	SWI-mIP	IFM	T2*	SWI-mIP	IFM
CRA	56	54	66	47	38	62	59	68	44	9	16	4	0.58	0.48	0.72
Exp.Junior1	72	73	81	67	60	78	39	46	28	5	13	3	0.75	0.67	0.83
Exp.Senior1	87	74	88	81	66	84	25	40	22	6	8	4	0.84	0.73	0.87
Trained. Ing	76	87	91	67	65	85	39	41	21	9	22	6	0.74	0.67	0.86
Exp.Senior2	91	108	74	77	69	70	29	37	36	14	39	4	0.78	0.64	0.78
Exp.Junior2	111	162	102	86	84	85	20	22	21	25	78	17	0.79	0.63	0.82
Median	**82**	**81**	**85**	**72**	**66**	**81**	**34**	**41**	**25**	**9**	**19**	**4**	**0.77**	**0.66**	**0.83**
Min	56	54	66	47	38	62	20	22	21	5	8	3	0.58	0.48	0.72
Max	111	162	102	86	84	85	59	68	44	25	78	17	0.84	0.73	0.87

Median is highlighted in bold character.

reliable consensus between two expert raters. Overall, 2D-dedicated phase processing yielding IFM proved a very promising tool to improve CMB detection in clinical setting. In fact, it yielded increased sensitivity and decreased FPR compared to T2* magnitude and SWI-mIP images for lesion detection. Furthermore, IFM yielded less inter-rater variability when classifying patients with numerous lesions than both other methods, with only a slight increase in rating duration.

One of the main novelties of our work was to compare between ratings performed on images obtained from the same 2D acquisition. Previous studies (Vernooij et al., 2008) compared 2D T2* magnitude images and 3D SWI; the differences that was obtained in this study may in fact be mainly due to the differences between 2D and 3D acquisitions. Using a unique acquisition was motivated by the fact that 2D GRE multi-slice sequences are more common than 3D multi-echo GRE T2* in large multi-center clinical research studies. IFM thus proved an efficient mean to improve CMBs detection. The method we used here, 2DHF, removes 2D acquisition and/or reconstruction artifacts from phase images while keeping fine details with limited border effects.

False positives in CMB detection can be explained by CMBs mimics, which have similar shape and signal properties.

The most frequent source of false positives is linked with blood vessels when not fully embedded in the acquired slice. In fact, when blood vessels are really perpendicular to acquisition plane, they will appear as round hyposignal. In some cases, SWI-mIP may help to differentiate these vascular-related mimics as the projection makes it possible to highlight the 3D tubular shape, as illustrated in Fig. 3. Less frequent false positives originate from susceptibility artifacts and Partial Volume Effect (PVE). PVE-related mimics are more likely to occur adjacent to the petrous temporal bones, para-nasal sinuses, frontal bones, orbit and occipital bone (Werring, 2011; Gregoire, 2014; Greenberg et al., 2009). Experienced raters seem to better distinguish between PVE-related mimics and CMBs as they rely more on anatomical criteria.

Although the use of SWI with 3D-acquisitions is known to increase the contrast of CMBs, allowing to detect smaller CMBs, here SWI yielded increased FPR on 2D acquisitions. In fact, SWI may enhance flow voids from small blood vessels and hyposignal artifact, thus creating new mimics. Thus, the increased sensitivity with SWI is balanced by a higher FPR. Moreover, the blooming effect is emphasized by SWI compared to magnitude images and could merge close CMBs. Finally, high pass filtering technique has been proven sub-optimal for background field

Table 5
Rating results for all detected CMBs for each subject group (see Table 8 in Appendix 2 for more details).

		∑ TP			∑ FN			∑ FP		
		T2*	SWI-mIP	IFM	T2*	SWI-mIP	IFM	T2*	SWI-mIP	IFM
G1 (N = 4)	Median							**2**	**2**	**0**
	Min							0	0	0
	Max							10	5	4
G2 (N = 5)	Median	**4**	**3**	**4**	**9**	**9**	**8**	**4**	**5**	**1**
	Min	3	2	2	4	7	6	0	3	0
	Max	8	5	6	9	10	10	31	7	19
G3 (N = 6)	Median	**82**	**76**	**85**	**41**	**48**	**38**	**13**	**27**	**7**
	Min	48	38	70	22	27	23	5	12	2
	Max	101	96	100	75	85	53	67	160	53

Median is highlighted in bold character.

Table 6
Number of correctly classified patients using the three types of images.

	G1 (N = 4)			G2 (N = 5)			G3 (N = 6)			Total		
	T2*	SWI-mIP	IFM	T2*	SWI-mIP	IFM	T2*	SWI-mIP	IFM	T2*	SWI-mIP	IFM
CRA	3	1	2	2	3	2	2	1	5	7	5	9
Exp.Junior1	3	3	4	2	2	2	4	3	5	9	8	11
Exp.Senior1	1	2	4	3	2	4	6	5	6	10	9	14
Trained. Ing	4	2	4	4	3	2	5	5	6	13	10	12
Exp.Senior2	3	4	4	2	2	2	6	5	5	11	11	11
Exp.Junior2	1	3	2	2	2	3	6	4	5	9	9	10
Median	**3**	**2,5**	**4**	**2**	**2**	**2**	**5**	**4,5**	**5**	**9,5**	**9**	**11**
Min	1	1	2	2	2	2	2	1	5	7	5	9
Max	4	4	4	4	3	4	6	5	6	13	11	14

Median is highlighted in bold character.

Table 7
Mean recorded rating durations by image type and subject-type (in seconds).

	T2*	SWI-mIP	IFM	Mean
G1 (N = 4)	70	68	84	74
G2 (N = 5)	58	47	54	53
G3 (N = 6)	172	173	193	179
Mean	100	96	110	

removal, crucial for SWI (Kaaouana et al., 2015). This could yield an enhancement of some artifactual voxels and increased the FPR.

Undetected CMBs are predominantly small lesions with low contrast with respect to their background. The criterion based on how much of the CMB is embedded in the parenchyma varies between authors and may lead to miss CMBs very close to sulci. Examples of undetected CMBs are illustrated in Fig. 4.

The efficient calculation of IFM with the 2DHF method requires expressing boundary conditions to define the "internal" region of interest. These boundary conditions are derived from a mask of the region of interest, calculated using SPM (Statistical Parametric Mapping) software as described in (Kaaouana et al., 2015), and named the brain mask. However, the border of the brain mask co-localizes with areas of strong susceptibility gradients, thus leading to potentially strong border effects. To reduce these border effects, the brain mask was eroded. Theoretically, this may lead to miss little areas of the cortex, and thus cortical CMBs, as shown in the example in Fig. 5. The use of both IFM and T2* magnitude images may help solve this issue for CMB detection.

QSM technique was not considered in this study. In fact, the dipole inversion underlying QSM is intrinsically three dimensional and thus strong anisotropic resolution may lead to large error propagation into resulting QSM maps. Moreover, regularization parameter setting is a challenging issue for multi-center data. QSM requires further investigation and validation and clinicians may need more training to interpret resulting maps.

Similar to SWI, IFM enhances the contrast of paramagnetic structures (Fig. 2), and thus increases sensitivity with respect to T2*-magnitude images. However, compared to SWI, IFM further yields a magnetic signature of CMBs. In fact, these inclusions behave like small magnetic dipoles and thus create magnetic field patterns similar to those of dipoles; this dipolar magnetic field appears as a ring-like effect in the axial plane in IFM. The sign pattern depends on the lesion susceptibility class (paramagnetic or diamagnetic) and allows to discriminate between CMBs and cerebral micro-calcifications related mimics (Kaaouana et al., 2015). Overall, IFM thus improve sensitivity (with respect to magnitude image) and decrease FPR (with respect to SWI-mIP). This results in a better characterization of subjects between "few CMBs" and "numerous CMBs".

Sensitivity and FPR of CMB detection appeared to vary between raters. Three types of rater behaviors were noted: two raters identified less CMBs on all image types and thus tended to under-rate (CRA and Exp.Junior1); one rater largely over-rated (Exp.Junior2); the last three raters had a similar tendency for medium rating (Trained.Ing and the two Exp.Seniors). For the two "under-raters", IFM was more sensitive and more specific than T2* magnitude and SWI-mIP. By contrast, for

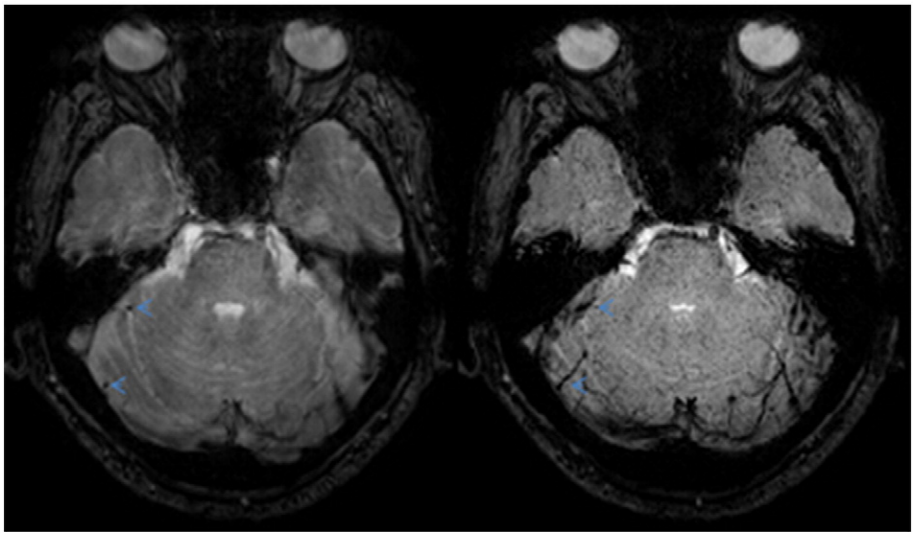

Fig. 3. Example of FPs detected on T2*-magnitude image; these two hypointensities, pointed by blue arrows, were recognized as vascular-related mimics on SWI-mIP image because of their tubular shape.

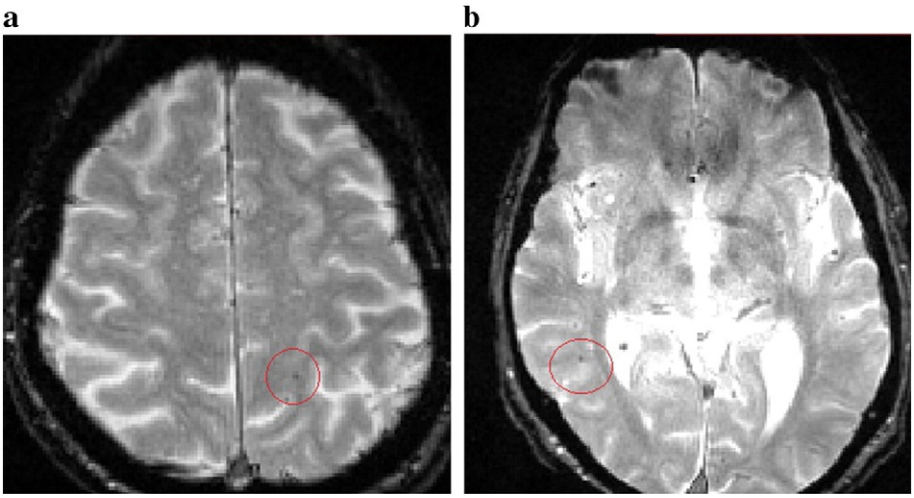

Fig. 4. Example of undetected CMBs due to its low contrast (a) and/or its distance to sulci (b).

the "over-rater", sensitivity was very high for all image types, but most FPs were identified on SWI-mIP and FPR was twice better for IFM than T2* magnitude and SWI-mIP. For the most experienced raters, all image types were comparable regarding sensitivity but IFM shows lower FPR.

IFM proved to yield only a slight increase in rating duration, even though it was a new image type for all raters. Rating duration may be reduced by making a better use of the magnetic signature in the visualization tool.

Although studies on CMBs are becoming more common, some ambiguity remains on their rating leading in low inter-rater reproducibility. In fact, the main reasons of between-raters discrepancy are low contrast, distance with respect to sulci and lesion size. Better integrating new neuroimaging tools is likely to lead to considerable improvements with respect to these issues.

For reference building-up, consensus was obtained with a specific scoring rule, in order to take into account the intrinsic variability of CMB detection (definite and possible CMBs). In fact, although the two observers reassessed the CMBs by considering all image types simultaneously, agreement was not perfect and a few cases of non-negligible disagreement were noticed. Even though relying on a third observer or consensus meetings may have been more standard approaches, both may suffer from subjective bias, whereas our approach considered both observer equally. Two visual scales have been proposed, MARS (The Microbleed Anatomical Rating Scale) (Gregoire et al., 2009) and BOMBS (Brain Observer Microbleed Scale) (Cordonnier et al., 2009). These scales were designed for clinical practice on T2* GRE weighted images and aimed at characterizing the number of CMBs and their distribution in the brain, as the clinical relevance of CMBs has been shown to be related with their localization. Here, a specific GUI was used to record the coordinates of each identified lesion allowing more accurate analysis of agreement or discrepancy cases. Moreover, rating results are embedded within an image with voxels labeled as "definite CMBs" and "possible CMBs" that can be combined either with a registered atlas or with a segmentation method to create an automatic report corresponding to the rating scales.

In this study, we have shown that IFM appears as an interesting add-on to T2*-magnitude image for the detection of CMBs. As expected, it allows discriminating mimics from real CMBs, visible "ring like" effects making it more specific in deep white and grey matter. On the contrary, T2*-magnitude image only and SWI-mIP seemed to increase false positives detection. IFM offers a simple and practical solution to assess the

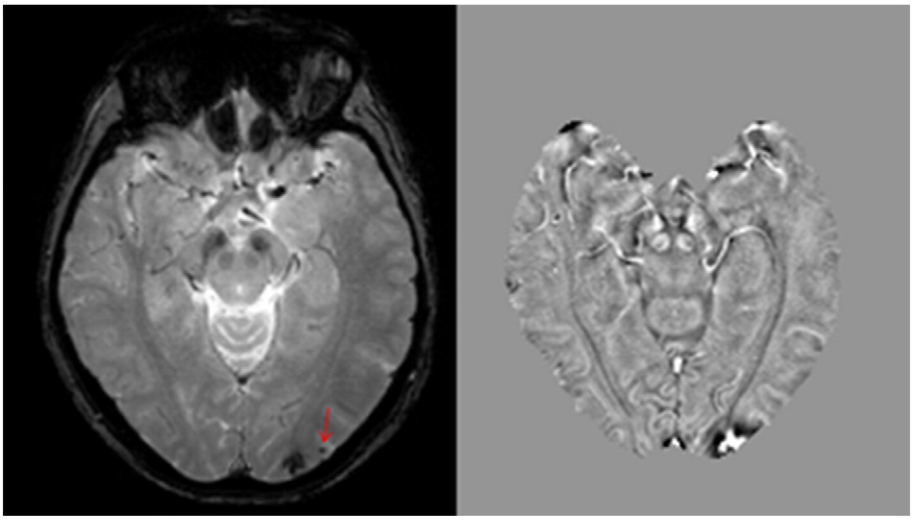

Fig. 5. CMB on the outer cortical part not visible on IFM; T2*-magnitude image shows two lesions (a spread lesion in the occipital lobe and a CMB (red arrow)). The hemorrhage was still seen in the IFM while the CMB disappeared due to masking operation.

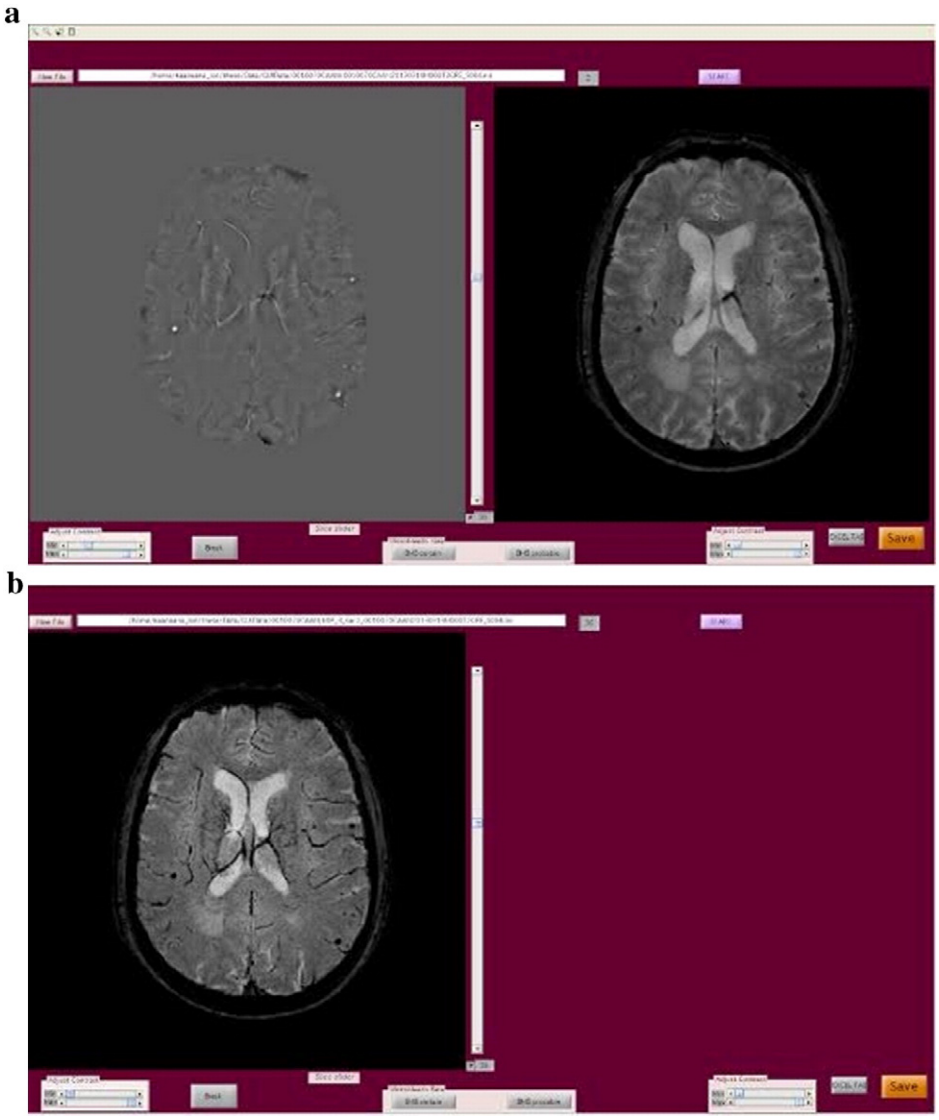

Fig. 6. (a) Evaluation GUI when IFM (on the left) is displayed (Magnitude image is on the right) and (b) when SWI-mIP is displayed.

presence, number and distribution of CMBs on standard clinical multicenter dataset. Further clinical studies on more subjects would help better assess the advantages of each type of images with respect to clinical usefulness, together with histological studies in order to infer the physical meaning of the magnetic signature. Finally, a comparison on 3D isotropic datasets would allow a better assessment of advantages of IFM with respect to SWI when these datasets are available.

Acknowledgments

The research leading to these results has received funding from the program "Investissements d'avenir" ANR-10-IAIHU-06. The MEMENTO study is undertaken through the sponsorship of "Bordeaux CHU" and the financial support of "Fondation Plan Alzheimer". The funding sources had no role in the design and conduct of the study; collection,

Table 8
Rating results for all detected CMBs for each subject group (details of Table 5).

	G1 (N = 4)			G2 (N = 5)									G3 (N = 6)								
	∑ FP			∑ TP			∑ FN			∑ FP			∑ TP			∑ FN			∑ FP		
	T2*	SWI-mIP	IFM	T2*	SWI-mIP	IFM	T2*	SWI-mIP	IFM	T2*	SWI-mIP	IFM	T2*	SWI-mIP	IFM	T2*	SWI-mIP	IFM	T2*	SWI-mIP	IFM
CRA	2	5	3	3	4	3	9	8	9	5	7	1	48	38	70	75	85	53	5	16	5
Exp.Junior1	1	1	0	4	3	3	8	9	9	1	3	0	70	62	83	53	61	40	6	12	3
Exp.Senior1	3	2	0	3	3	6	9	9	6	2	6	1	96	75	89	27	48	34	28	27	11
Trained.Ing	0	2	0	8	5	5	4	7	7	6	4	1	75	77	87	48	46	36	11	27	8
Exp.Senior2	1	0	0	3	2	2	9	10	10	0	4	4	89	76	76	34	47	47	14	37	2
Exp.Junior2	10	2	4	4	3	5	8	9	7	31	7	19	101	96	96	22	27	23	67	160	53
Median	2	2	0	4	3	4	9	9	8	4	5	1	82	76	76	41	48	38	13	27	7
Min	0	0	0	3	2	2	4	7	6	0	3	0	48	38	38	22	27	23	5	12	2
Max	10	5	4	8	5	6	9	10	10	31	7	19	101	96	96	75	85	53	67	160	53

management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Appendix 1. Pilot experiment

A pilot experiment was first carried-out in order to evaluate the experiment settings: number of subjects, image types and the specifically designed Graphical User Interface (GUI) built with GUIDE in MATLAB. This was conducted by an experienced neuroradiologist who selected the image types to be considered. Possible image types included: T2* magnitude, IFM and SWI minimum intensity projection (mIP) on n slices (n = [2,3,4]). 15 subjects proved feasible and T2* magnitude, SWI-mIP on three slices (slab thickness of 12 mm) and IFM were kept for the comparison experiments and the GUI was finalized as in Fig. 6. This evaluation GUI allows for displaying the overall 45 images, randomized on subjects and image types in order not to bias the comparison.

Appendix 2

Table 8 presents the rating results for all detected CMBs for each subject group (details of Table 5).

References

Charidimou, A., Shakeshaft, C., Werring, D.J., 2012. Cerebral microbleeds on magnetic resonance imaging and anticoagulant-associated intracerebral hemorrhage risk. Front. Neurol. 3. http://dx.doi.org/10.3389/fneur.2012.00133 September.

Chene, G., Ousset, P.-J., Pasquier, F., Dartigues, J.-F., Gabelle, A., Hugon, J., Blanc, F., et al., 2014. Memento: a national cohort on determinants and biomarkers of Alzheimer's disease and associated disorders. Alzheimers Dement.: J. Alzheimers Assoc. 10 (4): P814. http://dx.doi.org/10.1016/j.jalz.2014.05.1600.

Cheng, A.-L., Batool, S., McCreary, C.R., Lauzon, M.L., Frayne, R., Goyal, M., Smith, E.E., 2013. Susceptibility-weighted imaging is more reliable than T2*-weighted gradient-recalled echo MRI for detecting microbleeds. Stroke;J. Cereb. Circ. 44 (10): 2782–2786. http://dx.doi.org/10.1161/STROKEAHA.113.002267.

Conijn, M.M.A., Geerlings, M.I., Biessels, G.-J., Takahara, T., Witkamp, T.D., Zwanenburg, J.J.M., Luijten, P.R., Hendrikse, J., 2011. Cerebral microbleeds on MR imaging: comparison between 1.5 and 7T. Am. J. Neuroradiol. 32 (6):1043–1049. http://dx.doi.org/10.3174/ajnr.A2450.

Cordonnier, 2011. Brain microbleeds: more evidence, but still a clinical dilemma. Curr. Opin. Neurol. 24 (1):69–74. http://dx.doi.org/10.1097/WCO.0b013e328341f8c0.

Cordonnier, C., van der Flier, W.M., Sluimer, J.D., Leys, D., Barkhof, F., Scheltens, P., 2006. Prevalence and severity of microbleeds in a memory clinic setting. Neurology 66 (9):1356–1360. http://dx.doi.org/10.1212/01.wnl.0000210535.20297.ae.

Cordonnier, C., Potter, G.M., Jackson, C.A., Doubal, F., Keir, S., Sudlow, C.L.M., Wardlaw, J.M., Salman, R.A.-S., 2009. Improving interrater agreement about brain microbleeds: development of the Brain Observer MicroBleed Scale (BOMBS). Stroke; J. Cereb. Circ. 40 (1):94–99. http://dx.doi.org/10.1161/STROKEAHA.108.526996.

de Bresser, J., Brundel, M., Conijn, M.M., van Dillen, J.J., Geerlings, M.I., Viergever, M.A., Luijten, P.R., Biessels, G.J., 2013. Visual cerebral microbleed detection on 7T MR imaging: reliability and effects of image processing. AJNR Am. J. Neuroradiol. 34 (6): E61–E64. http://dx.doi.org/10.3174/ajnr.A2960.

Goos, J.D.C., van der Flier, W.M., Knol, D.L., Pouwels, P.J.W., Scheltens, P., Barkhof, F., Wattjes, M.P., 2011. Clinical relevance of improved microbleed detection by susceptibility-weighted magnetic resonance imaging. Stroke; J. Cereb Circ. 42 (7):1894–1900. http://dx.doi.org/10.1161/STROKEAHA.110.599837.

Greenberg, S.M., Vernooij, M.W., Cordonnier, C., Viswanathan, A., Salman, R.A.-S., Warach, S., Launer, L.J., Van Buchem, M.A., Breteler, M.M., 2009. Cerebral microbleeds: a guide to detection and interpretation. Lancet Neurol. 8 (2):165–174. http://dx.doi.org/10.1016/S1474-4422(09)70013-4.

Gregoire, S.M.F., 2014. Cerebral Microbleeds as a Marker of Small Vessel Disease: New Insights From Neuro-imaging and Clinical Studies in Stroke Patients." Doctoral. UCL (University College London) (http://discovery.ucl.ac.uk/1437813/).

Gregoire, S.M., Chaudhary, U.J., Brown, M.M., Yousry, T.A., Kallis, C., Jager, H.R., Werring, D.J., 2009. The Microbleed Anatomical Rating Scale (MARS): Reliability of a Tool to Map Brain Microbleeds. 73 (21):pp. 1759–1766. http://dx.doi.org/10.1212/WNL.0b013e3181c34a7d.

Guo, L.F., Geng, J., Qiu, M.H., Mao, C.H., Liu, C., Cui, L., 2013. Quantification of phase values of cerebral microbleeds in hypertensive patients using ESWAN MRI. Clin. Neuroradiol. 23 (3):197–205. http://dx.doi.org/10.1007/s00062-012-0196-4.

Haacke, E.M., Reichenbach, J.R., 2011. Susceptibility Weighted Imaging in MRI: Basic Concepts and Clinical Applications. first ed. Wiley-Blackwell (an imprint of John Wiley & Sons Ltd).

Haacke, E.M., Mittal, S., Wu, Z., Neelavalli, J., Cheng, Y.-C.n., 2009. Susceptibility-weighted Imaging: Technical Aspects and Clinical Applications, Part 1. 30 (1):pp. 19–30. http://dx.doi.org/10.3174/ajnr.A1400.

Hilal, S., Saini, M., Tan, C.S., Catindig, J.A., Koay, W.I., Niessen, W.J., Vrooman, H.A., et al., 2014. Cerebral microbleeds and cognition: the epidemiology of dementia in Singapore study. Alzheimer Dis. Assoc. Disord. 28 (2):106–112. http://dx.doi.org/10.1097/WAD.0000000000000015.

Kaaouana, T., de Rochefort, L., Samaille, T., Thiery, N., Dufouil, C., Delmaire, C., Dormont, D., Chupin, M., 2015. 2D harmonic filtering of MR phase images in multicenter clinical setting: toward a magnetic signature of cerebral microbleeds. NeuroImage 104 (January):287–300. http://dx.doi.org/10.1016/j.neuroimage.2014.08.024.

Klohs, J., Deistung, A., Schweser, F., Grandjean, J., Dominietto, M., Waschkies, C., Nitsch, R.M., Knuesel, I., Reichenbach, J.R., Rudin, M., 2011. Detection of cerebral microbleeds with quantitative susceptibility mapping in the ArcAbeta mouse model of cerebral amyloidosis. J. Cereb. Blood Flow Metab. 31 (12):2282–2292. http://dx.doi.org/10.1038/jcbfm.2011.118.

Lee, S.H., Bae, H.J., Kwon, S.J., Kim, H., Kim, Y.H., Yoon, B.W., Roh, J.K., 2004. Cerebral microbleeds are regionally associated with intracerebral hemorrhage. Neurology 62 (1), 72–76.

Liu, T., Surapaneni, K., Lou, M., Cheng, L., Spincemaille, P., Wang, Y., 2012. Cerebral microbleeds: burden assessment by using quantitative susceptibility mapping. Radiology 262 (1):269–278. http://dx.doi.org/10.1148/radiol.11110251.

McAuley, G., Schrag, M., Sipos, P., Sun, S.-W., Obenaus, A., Jaladhar, N., Mark Haacke, E., Holshouser, B., Madacsi, R., Kirsch, W., 2010. Quantification of punctate iron sources using magnetic resonance phase. Magn. Reson. Med. 63 (1):106–115. http://dx.doi.org/10.1002/mrm.22185.

Naka, H., Nomura, E., Wakabayashi, S., Kajikawa, H., Kohriyama, T., Mimori, Y., Nakamura, S., Matsumoto, M., 2004. Frequency of asymptomatic microbleeds on T2*-weighted MR images of patients with recurrent stroke: association with combination of stroke subtypes and leukoaraiosis. AJNR Am. J. Neuroradiol. 25 (5), 714–719.

Nandigam, R.N.K., Viswanathan, A., Delgado, P., Skehan, M.E., Smith, E.E., Rosand, J., Greenberg, S.M., Dickerson, B.C., 2009. MR imaging detection of cerebral microbleeds: effect of susceptibility-weighted imaging, section thickness, and field strength. AJNR Am. J. Neuroradiol. 30 (2):338–343. http://dx.doi.org/10.3174/ajnr.A1355.

Park, J.-H., Seo, S.W., Kim, C., Kim, G.H., Noh, H.J., Kim, S.T., Kwak, K.-C., et al., 2013. Pathogenesis of cerebral microbleeds: in vivo imaging of amyloid and subcortical ischemic small vessel disease in 226 individuals with cognitive impairment. Ann. Neurol. 73 (5):584–593. http://dx.doi.org/10.1002/ana.23845.

Reichenbach, J.R., Venkatesan, R., Schillinger, D.J., Kido, D.K., Haacke, E.M., 1997. Small vessels in the human brain: MR venography with deoxyhemoglobin as an intrinsic contrast agent. Radiology 204 (1):272–277. http://dx.doi.org/10.1148/radiology.204.1.9205259.

Song, M.-H., Napel, S., Pelc, N.J., Glover, G.H., 1995. Phase unwrapping of MR phase images using Poisson equation. IEEE Trans. Image Process. 4 (5):667–676. http://dx.doi.org/10.1109/83.382500.

Tsushima, Y., Aoki, J., Endo, K., 2003. Brain microhemorrhages detected on T2*-weighted gradient-echo MR images. AJNR Am. J. Neuroradiol. 24 (1), 88–96.

Vernooij, M.W., Arfan Ikram, M., Wielopolski, P.A., Krestin, G.P., Breteler, M.M.B., van der Lugt, A., 2008. Cerebral microbleeds: accelerated 3D T2*-weighted GRE MR imaging versus conventional 2D T2*-weighted GRE MR imaging for detection. Radiology 248 (1):272–277. http://dx.doi.org/10.1148/radiol.2481071158.

Werring, D.J., 2011. Cerebral Microbleeds: Pathophysiology to Clinical Practice. Cambridge University Press.

Yamada, N., Imakita, S., Sakuma, T., Takamiya, M., 1996. Intracranial calcification on gradient-echo phase image: depiction of diamagnetic susceptibility. Radiology 198 (1), 171–178.