Reproducibility of High-Resolution MRI for the Identification and the Quantification of Carotid Atherosclerotic Plaque Components

Consequences for Prognosis Studies and Therapeutic Trials

Emmanuel Touzé, MD, PhD; Jean-François Toussaint, MD, PhD; Joël Coste, MD, PhD; Emmanuelle Schmitt, MD; Fabrice Bonneville, MD; Pierre Vandemarcq, MD; Jean-Yves Gauvrit, MD; Françoise Douvrin, MD; Jean-François Meder, MD, PhD; Jean-Louis Mas, MD; Catherine Oppenheim, MD, PhD; for the HIRISC study group

Background and Purpose—Although MRI is increasingly proposed to investigate composition of carotid atherosclerosis, its reproducibility has rarely been addressed. We assessed the reproducibility of MRI for the identification and quantification of carotid atherosclerotic plaque components.

Methods—Using published criteria, 2 readers independently analyzed the carotid MRI (1.5-T MR units with a 4-channel phased-array surface coil, Machnet) of 85 consecutive patients with symptomatic (40% to 69% according to NASCET method) or asymptomatic (60% or greater) carotid artery stenosis enrolled in an ongoing prognostic study. One reader reevaluated all images. Fibrous cap was also secondarily identified independently on T2-weighted and time-of-flight (TOF) images.

Results—Intraobserver agreement was substantial for the identification of calcifications (kappa [κ]=0.70; 95% CI: 0.54 to 0.86) and lipid-rich/necrotic core (LR/NC) (κ=0.69; 0.31 to 0.86), almost perfect for hemorrhages (κ=0.82; 0.68 to 0.96), and moderate (κ=0.58; 0.27 to 0.88) and fair (κ=0.33; 0.09 to 0.56) for fibrous cap identification on T2-weighted and TOF images, respectively. Interobserver agreement was substantial for the identification of calcifications (κ=0.74; 0.59 to 0.89) and hemorrhages (κ=0.62; 0.43 to 0.81), and moderate for LR/NC (κ=0.58; 0.20 to 0.95). Agreement was fair for fibrous cap identification on both T2-weighted (κ=0.28; −0.03 to 0.59) and on TOF images (κ=0.26; 0.04 to 0.48). Agreement between T2 and TOF images for fibrous cap identification was slight (κ=0.16; 0.01 to 0.31). Intra- and interobserver reproducibility for quantitative area measurements of vessel, lumen, plaque, LR/NC, and fibrous components was high with intraclass correlation coefficients ranging from 0.73 to 0.99. However, for the LR/NC, the interval delimited by the Bland-Altman graphs was wide in comparison to the mean.

Conclusions—Vessel and plaque quantification is reproducible. Reproducibility of MRI for identifying and quantifying carotid plaque components is overall acceptable, but there is still significant variability that should be taken into account in the design of prognosis studies and clinical trials. Reproducibility for fibrous cap identification needs to be improved. (Stroke. 2007;38:1812-1819.)

Key Words: atherosclerosis ■ carotid stenosis ■ MRI ■ reproducibility ■ stroke

In vivo high-resolution MRI has emerged as a tool capable of measuring the volume of atherosclerotic carotid plaques1,2 and characterizing their components, including hemorrhages, calcifications, lipid component (ie, lipid-rich/necrotic core [LR/NC]),3,4 and fibrous tissue.5-7 By measuring atherosclerosis burden and plaque instability, MRI could potentially identify high-risk patients requiring more intensive treatments,8 monitor the effect of treatments,9 and help with understanding atherosclerosis pathophysiology. However, reproducibility of MRI for characterizing plaque components and measuring plaque volumes has rarely been addressed. Intra- and interobserver reproducibility of MRI for...
quantitative measurements of the wall, lumen, and different plaque components has been evaluated in histological correlation studies but reproducibility of qualitative in vivo MR measurements has never been assessed. These previous histological studies relied on small samples of selected patients with severe carotid atherosclerosis scheduled for endarterectomy. Furthermore, in the analyses, all arterial locations were pooled together, which could lead to inaccurate estimations. Finally, there is still some controversy about the best way to identify the fibrous cap. Some authors define it as a hypointense area on native 3-dimensional time-of-flight (TOF) images, whereas others regard it as a high signal area adjacent to the lumen on T2-weighted sequences.

We assessed the intra- and interobserver reproducibility of high-resolution MRI for the identification and quantification of carotid atherosclerotic plaque components in 85 consecutive patients enrolled in an ongoing prospective prognostic study.

Methods

Population
High-Resolution magnetic resonance Imaging in atherosclerotic Stenosis of the Carotid artery (HIRISC) is an ongoing multicenter prospective study assessing the prognostic value of carotid plaque vulnerability, as defined on MRI, for the prediction of cerebrovascular events (participants are listed in the Appendix). Patients were eligible for the study if 1) they had symptomatic (40% to 69% according to the NASCET method) or asymptomatic (60% or greater NASCET) stenosis of the internal carotid artery bifurcation; 2) they were not scheduled for endarterectomy within the next 6 months; and 3) they did not have any other major cause of stroke. The decision not to operate was jointly reached with the neurologists and the patient after having the benefits and risks of carotid surgery explained. Patients were considered symptomatic if they have had a recent (12 or less months) transient ischemic attack or a nondisabling (Rankin score 3 or less) ischemic stroke in the territory of the qualifying carotid stenosis. The degree of stenosis was obtained from carotid duplex ultrasound and at least one of the following techniques: MR angiography, CT angiography, or digitized catheter angiography. The study was approved by the local ethics committee and all patients signed an informed consent.

MRI Protocol
All patients were imaged in 1.5-T MR units using the same 4-channel phased-array surface coil (Machnet). Before starting the study, MR protocol and acquisition parameters were standardized across platforms; a fast gradient echo pulse sequence was used in axial, sagittal, and coronal planes as a localizer. The median sagittal image was used to plan a 2-dimensional TOF gradient echo sequence. Twenty to 30 slices with a thickness of 4 mm were set to cover the neck area using the phased array coil. The z-axis coordinates of the qualifying carotid bifurcation on the 2-dimensional TOF images were used to position the 4 following pulse sequences: 3-dimensional TOF, T1-weighted (T1w), proton density-weighted (PDw), and T2-weighted (T2w). The field of view (130×130 mm) was identical for all 4 sequences. T1w, T2w, and PDw images were obtained with double-inversion recovery (ie, black blood) fast spin echo sequences with electrocardiographic gating during free breathing using 8 axial sections (3-mm thick, 0.3-mm gap) centered on the qualifying carotid stenosis. PDw and T2w parameters were: repetition time: 2 RR intervals; effective echo time: 16 to 20 ms for PDw and 50 ms for T2w; acquisition matrix: 256×256 with zero filling (acquired in plane resolution 508×508 μm interpolated to 254×254 μm by zero-filling in K-space); signal averaged: 2. Fat suppression. T1w parameters were: repetition time: 1 RR interval; echo time: 9 to 10 ms; acquisition matrix: 352×256, 512 with zero filling (acquired in plane resolution 451×508 μm, interpolated to 254×254 μm by zero-filling in K-space); signal averaged: 3. The 3-dimensional TOF sequence used a gradient echo pulse sequence with repetition time: 30 ms; echo time: 6.9 ms; flip angle: 20°; acquisition matrix: 288×224, 512 zero filling (acquired in plane resolution 451×580 μm interpolated to 254×254 μm by zero-filling in K-space); number of signal averaged: 2; 20 slices of 2.2-mm thickness, one slab. Total acquisition time was approximately 25 minutes. Images were provided in a DICOM format for analysis.

Image Review
After a period of joint working to standardize the methods of assessment and recording, 2 independent readers blind to baseline and follow-up clinical data examined all MR images of the qualifying carotid artery using a standardized form and published criteria (see subsequently). Three months after the first reading session, one observer reevaluated all images, which were presented in a different order. For each location, the 4 MR sequences (PDw, T2w, T1w, TOF) were reviewed together. Image quality was assessed by each observer using a 4-point scale (excellent, good, average, poor). This latter category was being excluded from our analyses based on the signal quality and on the motion artifacts.

Tissue Components (qualitative assessment)
All signal intensities were compared with the adjacent sternocleidomastoid muscle. Calcifications were defined as areas of hypointensity on all 4 sequences. Recent and fresh intraplaque hemorrhages were considered together as hyperintensities on T1w and TOF images. LR/NC and fibrous components share the same signal on PDw images, ie, a signal iso- or slightly hyperintense to the sternocleidomastoid muscle. PDw and T2w images were compared so as to discriminate between LR/NC and fibrous component as follows: LR/NC was identified as an area in which the signal dropped on T2w when compared with PDw images, whereas fibrous component corresponded to a relatively high signal area on both sequences (Figures 1 and 2). Calcifications, hemorrhages, and LR/NC were considered present if they were observed on one slice at least. Fibrous cap was defined as a high signal area adjacent to the lumen on T2w images. On TOF images, fibrous cap was identified as a hypointense band adjacent to the lumen on the plaque surface. Readers had to indicate the presence of a fibrous cap on each sequence independently of the finding on the other sequence. The reproducibility of MRI for the identification of the fibrous cap on TOF and TOF images was assessed on one randomly selected location per patient. In the absence of LR/NC, fibrous cap cannot be distinguished from the fibrous component on T2w images. Therefore, locations where no LR/NC was identified by both readers were excluded from this analysis.

Quantitative Measurements
Area measurements of vessel, lumen, plaque, LR/NC, and hemorrhages were obtained for each location by manually tracing the boundaries of each component (Figures 1 and 2). Calcification areas were not quantified because of their usual small size and for the various susceptibility artifacts they induce. Vessel and lumen areas were calculated on PDw images and plaque and LR/NC areas on T2w images. Fibrous component area was calculated by subtracting LR/NC and hemorrhage from plaque area. Mean areas per artery were calculated by dividing the sum of areas from each cross-sectional location by the number of cross-sectional locations analyzed. Because the absolute values of the LR/NC area do not provide any information on the relative size of the lipid core compared with the vessel (which is a potential marker of plaque instability), we also calculated the ratio of LR/NC to wall (ie, vessel minus lumen) areas for each patient and classified it into 4 categories: less than 10%, 10% to 19%, 20% to 29%, and 30% or greater. These categories covered the range of ratios observed in our study sample (0% to 45%).

Statistical Analyses
We calculated the sample size required to detect kappa values significantly superior to 0.40 with a power of 0.80 assuming that the prevalences of the calcifications and hemorrhages would range
between 30% and 50% and kappa values between 0.70 and 0.80. According to Donner and Eliasziw,19 50 to 80 patients were required. For each patient, the carotid artery qualifying for the study was analyzed. Agreements were assessed by calculating kappa ($\kappa$) for nominal (dichotomous) data and weighted (quadratic weighting) $\kappa$ for statistic ranked ordinal data,20 and intraclass correlation coefficients with a 2-way random effect for continuous variables.21 For all agreement parameters, 95% CIs were calculated. According to Landis and Koch,22 values of $\kappa$ between 0.8 and 1 indicate almost perfect agreement, 0.6 to 0.8 substantial agreement, 0.4 to 0.6 moderate agreement, 0.2 to 0.4 fair agreement, 0.0 to 0.2 slight agreement, and 0.0 to 0.0 poor agreement. Values of intraclass correlation coefficient more than 0.80 were considered excellent agreement.21 Finally, we plotted the mean differences between 2 measurements against the mean of 2 measurements according to the method described by Bland and Altman.23 This method allows judging whether agreement of quantitative measurements varies across the range of values, quantifying the bias between observers, and calculating the precision of the estimated limits of agreement.

Results

MRI examinations of the first 90 patients enrolled in the HIgh-Resolution magnetic resonance Imaging in atherosclerotic Stenosis of the Carotid artery study were selected for this reproducibility analysis. Five (6%) patients were excluded because of poor image quality (n=4) or incomplete MR examination (n=1). Therefore, 85 patients (mean age 69.4±9.9 years; 71% of men) were finally included. Fifty-four patients were asymptomatic, 12 had a transient ischemic attack (including one amaurosis fugax), and 19 a minor stroke. The median number of locations per patient evaluated by both readers (ie, with significant plaque allowing areas measurements) was 3 (one to 5) with a total of 245 locations. The number of patients with images the quality of which was considered average by both readers was small: PDw: 0, TOF: one, T2w: 2, and T1w: 8. Only one patient had 2 average quality sequences according to both readers (excluding this patient did not change our results).

Figure 1. Moderate carotid stenosis as seen on contrast enhanced MR angiography (MRA) along with high-resolution axial images taken 1 cm distal to the bifurcation (arrow). Calcifications correspond to the area of hypointensity on T1w, TOF, PDw, and T2w sequences. A large LR/NC area, corresponding to the area in which signal dropped on T2w when compared with the PDw image, is delineated on the T2w image. The fibrous cap can be seen on TOF (hyposignal band between the smooth lumen boundary and the plaque) and on the T2w image (hyperintensity between the lumen and the LR/NC area).
there was a small $\kappa$ value for the identification of a fibrous cap on T2w images ($\kappa=0.28; -0.03$ to $0.59$). Of note, this $\kappa$ value slightly improved when images with hemorrhages (ie, identified by at least one observer) were excluded ($\kappa=0.35; 0.004$ to $0.71$) or when images with calcifications were excluded ($\kappa=0.44; 0.0$ to $0.89$), and reached $0.78$ ($0.36$ to $1.0$) when those with both hemorrhages and calcifications were excluded. On TOF images, fibrous cap was more seldom identified with a percentage of agreement of $69\%$ and a $\kappa$ value of $0.26$ ($0.04$ to $0.48$). In contrast to findings on T2w images, the exclusion of images with hemorrhage and calcifications did not change the interobserver agreement on TOF images ($\kappa=0.20; -0.16$ to $0.56$). Finally, given this low reproducibility, 2 months after the first readings, one observer examined a random location (ie, a location different from that used in the first analysis) from each patient on T2w images only. Then, 4 weeks later, he analyzed the same locations on TOF images only (presented in a different order). In this analysis, agreement between T2w and TOF images for the identification of a fibrous cap was slight ($\kappa=0.16; 0.01$ to $0.31$).

Interobserver reproducibility for quantitative areas measurements was high: intraclass correlation coefficient was $0.96$ ($0.95$ to $0.97$) for the vessel, $0.98$ ($0.98$ to $0.99$) for the lumen, $0.94$ ($0.92$ to $0.96$) for the plaque, $0.73$ ($0.66$ to $0.80$) for the LR/NC, and $0.84$ ($0.79$ to $0.88$) for the fibrous component. In plaques in which both readers agreed on the presence of hemorrhages, intraclass correlation coefficient was $0.60$ ($0.37$ to $0.81$) for hemorrhagic areas. Bland and Altman plots (Figure 3) showed that the absolute differences in measurements were small (ranging from $-3.0$ to $0.9 \text{ mm}^2$ depending on the component) and that there was no relation between differences and means. However, the interval of

Table 1. Intraobserver Reproducibility

<table>
<thead>
<tr>
<th>Identification of Plaque Components</th>
<th>Agreement</th>
<th>$\kappa$ (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calcifications</td>
<td>86%</td>
<td>0.70 (0.54 to 0.86)</td>
</tr>
<tr>
<td>Hemorrhages</td>
<td>93%</td>
<td>0.82 (0.68 to 0.96)</td>
</tr>
<tr>
<td>Lipids</td>
<td>92%</td>
<td>0.69 (0.40 to 0.98)</td>
</tr>
<tr>
<td>Fibrous cap (T2w)</td>
<td>91%</td>
<td>0.58 (0.27 to 0.88)</td>
</tr>
<tr>
<td>Fibrous cap (TOF)</td>
<td>72%</td>
<td>0.33 (0.09 to 0.56)</td>
</tr>
</tbody>
</table>

Quantitative Measurements Intraclass Coefficient Correlation (95% CI)

| Vessel                             | 0.98 (0.98 to 0.99) |
| Lumen                              | 0.99 (0.98 to 0.99) |
| Plaque                             | 0.90 (0.87 to 0.93) |
| Lipids                             | 0.72 (0.65 to 0.80) |
| Fibrous component                  | 0.77 (0.70 to 0.83) |
| Hemorrhages                        | 0.70 (0.52 to 0.85) |
agreement (mean of differences ± 2 SDs) for the LR/NC was wide (−9.8 to 8.2 mm²) compared with the mean which ranged from 0 to 30 mm² (mean, 10.1 mm²). Consequently, as shown in Table 3, the agreement for the classification of plaques according to the ratio LR/NC to wall areas was only substantial (weighted κ = 0.61; 0.48 to 0.74). Findings regarding LR/NC quantification were unchanged when patients with hemorrhage or calcifications were excluded.

Figure 3. Bland and Altman plots (interobserver reproducibility). This figure shows the mean differences between 2 measurements against the mean of 2 measurements (interobserver agreement). Areas are expressed in millimeters squared and correspond to the sum of areas from each cross-sectional location divided by the number of cross-sectional locations analyzed. Y-scales are the same except for lumen areas. On each graph, dotted lines represent the value of the mean bias ± 1.96 SDs (difference between the 2 observers' measurements). There was no variation of agreement across the range of values. The 95% CIs of mean biases were narrow compared with the mean values of measurements except for lipid component areas in which the 95% CI of the mean bias (−9.8 to 8.2) was relatively broad compared with the mean of all lipid areas (10.1 mm²).
TABLE 2. Interobserver Reproducibility of MRI in Identifying the Different Plaque Components

<table>
<thead>
<tr>
<th>Component</th>
<th>Observer 1</th>
<th>Observer 2</th>
<th>Agreement</th>
<th>( \kappa ) (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calcification</td>
<td>Absence</td>
<td>25</td>
<td>5</td>
<td>88%</td>
</tr>
<tr>
<td></td>
<td>Presence</td>
<td>5</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Hemorrhage</td>
<td>Absence</td>
<td>55</td>
<td>6</td>
<td>85%</td>
</tr>
<tr>
<td></td>
<td>Presence</td>
<td>7</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>LR/NC</td>
<td>Absence</td>
<td>3</td>
<td>1</td>
<td>95%</td>
</tr>
<tr>
<td></td>
<td>Presence</td>
<td>3</td>
<td>78</td>
<td></td>
</tr>
<tr>
<td>Fibrous cap†</td>
<td>T2w</td>
<td>Absence</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Presence</td>
<td>7</td>
<td>65</td>
<td></td>
</tr>
<tr>
<td>TOF</td>
<td>Absence</td>
<td>45</td>
<td>7</td>
<td>69%</td>
</tr>
<tr>
<td></td>
<td>Presence</td>
<td>17</td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

*Defined on one random location per patient (see “Methods”).

All results related to interobserver reproducibility for qualitative (Table 2) and quantitative (data not shown) measurements were similar when the second observer 1 readings were used or when patients with at least one average quality sequence (n=10) were excluded (data not shown). Our findings were also similar for symptomatic and asymptomatic patients (data not shown).

Discussion

We have shown that the interobserver reproducibility of high-resolution MRI for the in vivo identification of hemorrhages, calcifications, and LR/NC was moderate to substantial. The reproducibility was excellent for the quantification of vessel and plaque areas but not perfect for the quantification of hemorrhages and LR/NC. We have also shown that fibrous cap cannot be reliably identified. As expected, intraobserver reproducibility was better than interobserver reproducibility. Although there are published data on reproducibility of MRI for the quantification of different plaque components,1,10,11 to our knowledge, intra- and interobserver reproducibility of MRI for the identification of different plaque components and fibrous caps had never specifically been assessed so far.

In our study, intra- and interobserver agreements were not perfect for most of the components. However, the reproducibility found here was very similar to that observed in a recent histological carotid plaque study24 and was at least as good as that of different carotid imaging techniques.25–27 Our findings of a good reproducibility of MRI in quantifying vessel, lumen, and plaque areas are in keeping with those of previous smaller studies, which included patients scheduled for carotid endarterectomy.1,11 Taken together, these results mean that MRI can be used quite confidently in clinical studies addressing whether the risk of stroke is correlated to atherosclerosis burden and in studies addressing the effect of drugs on the size of atherosclerosis plaques. However, the quantification of the LR/NC was not perfect. Although intraclass correlation coefficient (0.73) and \( \kappa \) coefficient for the lipids/wall ratio (0.61) suggested a relatively good agreement, Bland and Altman plot showed a relatively wide interval of agreement compared with the mean. This indicated that there could be discrepancies between 2 readers. This result highlights the technical limitations of the intraobserver reproducibility coefficient in addressing the agreement between measurements of continuous variables28 and the need to carefully analyze Bland and Altman plots.23 Similarly, we found a slightly lower interobserver reproducibility for quantifying hemorrhages than that found in a previous study.1 The usual small size of hemorrhagic areas along with their frequent blurred edges on MRI and the heterogeneity of their signal are likely to explain the result.24,29

We found that fibrous cap was more often identified on T2w than on TOF images, but reproducibility was overall poor for both methods and there was no agreement between both methods. On T2w images, the \( \kappa \) value, although low, was better in plaques without calcifications and hemorrhages. These difficulties in distinguishing fibrous cap from LR/NC in plaques with hemorrhages and calcifications have been previously noticed.6 An overlap between the signal intensity of hemorrhages, which varies according to their age and the signal of the other components, could explain these difficulties. Given the relative high prevalence of these components, this potential explanation has no implications for the practice but could direct future research. Another plausible explanation could be that calcifications are sometimes adjacent to the lumen and disrupt the signal. Considering these difficulties in identifying fibrous cap reliably, we decided not to carry out further analyzes on the fibrous cap status. The discrimination
between LR/NC and fibrous cap on T2w images is based on differences in spectroscopic properties between these both components and has been well documented in both in vitro and in vivo studies.\textsuperscript{3,6,13–16} Conversely, the diagnostic value of TOF sequences has been less documented and,\textsuperscript{12} so far, findings have not been replicated. In a recent study addressing whether plaque vulnerability defined on MRI predicted cerebrovascular events, fibrous cap was defined after having several readers reach a consensus.\textsuperscript{8} Nevertheless, although T2w fibrous cap was identified more frequently on T2w than on TOF sequences, our results highlight that much remains to be done to improve the ability of MRI to identify fibrous cap confidently. The use of gadolinium contrast agent has been shown to improve the distinction between fibrous cap and LR/NC and could help with more accurate measurements of the fibrous cap.\textsuperscript{30} However, to our knowledge, reproducibility of postcontrast assessments has not been studied yet.

There are several potential shortcomings to our study. First, we included patients who were not scheduled for endarterectomy and it might therefore be argued that our population did not cover the whole spectrum of the disease severity, which could question the generalizability of our results. However, patients included in this study are those for whom the benefit of endarterectomy is still controversial and for whom much is to be expected from noninvasive techniques. As a consequence, no pathologic confirmation data were available in our study and we may have found good agreement despite that both readers misinterpreted images. However, we carefully used previously published criteria.\textsuperscript{1,4,16,18} Moreover, agreement studies do not require a reference standard,\textsuperscript{31} and methods that are not reproducible are useless in clinical practice although they are accurate. In addition, there are some data on MRI reproducibility in patients with more severe carotid stenosis, which are consistent with our results.\textsuperscript{1} Second, testing interobserver reproducibility with 2 readers and intraobserver reproducibility with one observer only may also question the generalizability of our results. However, testing intraobserver reproducibility is mainly useful to make sure that interobserver reproducibility is interpretable. In addition, there is no reason to believe that our results would have been very different with other readers provided they use the same strict criteria and have a training period. Third, in contrast to some studies, we did not use an automatic imaging analyzing tool for quantitative measurements.\textsuperscript{1} However, such tools have not thoroughly been validated for the detection and the measurement of all plaque components.\textsuperscript{32} Fourth, like some previous studies, we used a 3-mm slice thickness,\textsuperscript{3,6,9} but others used a 2-mm slice thickness.\textsuperscript{5,11} This greater thickness may have increased partial volume effects and potentially accounted for part of the variability observed. However, for a given acquisition time and magnetic field, decreasing the slice thickness by one third would have resulted in a loss of signal and contrast. In addition, a recent randomized study assessing effects of statins on carotid plaques used a 3-mm slice thickness.\textsuperscript{9} Finally, there are several well-described limitations to the use of the $\kappa$ statistic that could influence the interpretation of our results.\textsuperscript{31} Very high (or low) prevalence results in high levels of expected agreement and consequently, the $\kappa$ value is often low despite nearly perfect agreement as we observed for the LR/NC and the fibrous cap on T2w images. Observer bias, in the sense of a systematic difference between 2 observers in the image interpretation, is a form of disagreement with important practical implications, but it is not separately identified by the $\kappa$ statistic. Despite the use of standardized and strict criteria defined during the period of joint working, such a bias could be suspected for the first readings on TOF (fibrous cap identification). However, there was no such bias during the second readings, and because observer bias tends to overestimate the $\kappa$ value, the true value of the first readings $\kappa$ value would have been even smaller ($\kappa=0.16; -0.08$ to $0.40$) after adjustment using the bias index.\textsuperscript{33}

Our results have several practical implications for future research. First, MRI is reproducible for the quantification of vessel and plaque areas and can be used confidently in future studies, which require measuring plaque size. Second, there is still significant interobserver variability for the identification and quantification of different plaque components. This variability should be taken into account in the design of prognostic studies and therapeutic trials, because it can attenuate the correlations between MRI findings and clinical outcomes and consequently increase the required sample sizes.\textsuperscript{20} We would also recommend that all future studies publish the reproducibility of their own assessment unless the final decision is reached by consensus. Third, much remains to be done to improve the ability of MRI to identify fibrous cap reliably and to better understand the influence of hemorrhages and calcifications on the analysis of the other components. These important questions could potentially be answered by carrying out larger well-designed histological–radiological correlation studies\textsuperscript{34} using contrast agents and/or new MRI sequences,\textsuperscript{2,6,35} developing automatic images analyzing tools,\textsuperscript{36,37} and using higher magnetic fields.\textsuperscript{38,39}

### Appendix

#### List of Investigators


#### Acknowledgments

We thank the manufacturer engineers who helped with determining the MR acquisition protocols: Thierry Munier and Guillaume Calmon (General Electric Healthcare Care), Stéphane Breil (Philips & Siemens), Gwenaël Herigault (Philips). We also thank Sandra Madamaly and all other staff and collaborators of the Unité de Recherche Clinique, Hôpital Cochin, Paris. Study concept and design: Oppenheim, Touzé, Mas, Meder, and Coste. Acquisition of data: Touzé, Toussaint, Oppenheim, Schmitt, Bonneville, Vandemarcq, and Douvrin. Analysis and interpretation of data: Touzé, Oppenheim, and Toussaint. Drafting of the manuscript: Touzé, Toussaint, and Oppenheim. Critical revision of the manuscript for important intellectual content: Mas, Meder, and...
Coste. Statistical analysis: Touzé and Coste. Obtained funding: Oppenheim, Touzé, Mas, and Meder.

Sources of Funding
Supported by a grant from the Programme Hospitalier de Recherche Clinique of the French Ministry of Health (No. AOR 02 009) and a grant from the Institut de l’Athérombose (Sanofi-Aventis & Bristol-Myers-Squibb Pharmaceuticals).

Disclosures
None.

References
Reproducibility of High-Resolution MRI for the Identification and the Quantification of Carotid Atherosclerotic Plaque Components: Consequences for Prognosis Studies and Therapeutic Trials

Emmanuel Touzé, Jean-François Toussaint, Joël Coste, Emmanuelle Schmitt, Fabrice Bonneville, Pierre Vandermarcq, Jean-Yves Gauvrit, Françoise Douvrin, Jean-François Meder, Jean-Louis Mas and Catherine Oppenheim

for the HIgh-Resolution magnetic resonance Imaging in atherosclerotic Stenosis of the Carotid artery (HIRISC) study group

Stroke. 2007;38:1812-1819; originally published online April 26, 2007;
doi: 10.1161/STROKEAHA.106.479139

Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2007 American Heart Association, Inc. All rights reserved.
Print ISSN: 0039-2499. Online ISSN: 1524-4628

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/38/6/1812