Measuring Carotid Stenosis on Contrast-Enhanced Magnetic Resonance Angiography
Diagnostic Performance and Reproducibility of 3 Different Methods

Jean Marie K.S. U-King-Im, MRCS; Rikin A. Trivedi, MRCP, MRCS; Justin J. Cross, MRCP, FRCR; Nicholas J.P. Higgins, MRCP, FRCR; William Hollingworth, PhD; Martin Graves, MSc; Ilse Joubert; Peter J. Kirkpatrick, MSc, FRCS; Nagui M. Antoun, FRCP, FRCR; Jonathan H. Gillard, BSc, MD, FRCR

Background and Purpose—The aim of this study was to compare diagnostic performance and reproducibility of 3 different methods of quantifying stenosis on contrast-enhanced magnetic resonance angiography (CEMRA), with intra-arterial digital subtraction angiography (DSA) as the reference standard.

Methods—167 symptomatic patients scheduled for DSA, after screening Doppler ultrasound, were prospectively recruited to undergo CEMRA. Severity of stenosis was measured according to the North American Symptomatic Trial Collaborators (NASCET), European Carotid Surgery Trial (ECST), and the common carotid (CC) methods. Measurements for each method were made for 284 vessels (142 included patients) on both CEMRA and DSA in a blinded and randomized manner by 3 independent attending neuroradiologists.

Results—Significant differences in prevalence of severe stenosis were seen with the 3 methods on both DSA and CEMRA, with ECST yielding the least and NASCET the most cases of severe stenosis. Overall, all 3 methods performed similarly well in terms of intermodality correlation and agreement. No significant differences in interobserver agreement were found on either modality. With CEMRA, however, we found a significantly lower sensitivity for detection of severe stenosis with ECST (79.8%) compared with NASCET (93.0%), with DSA as reference standard.

Conclusions—Uniformity of carotid stenosis measurement methods is desirable because patient management may otherwise differ substantially. All 3 methods are adequate for use with DSA. With CEMRA, however, this study supports use of the NASCET method because of improved sensitivity for detecting severe stenosis. (Stroke. 2004;35:2083-2088.)

Key Words: angiography ■ carotid stenosis ■ magnetic resonance imaging
previous studies have, however, used the NASCET method for quantifying stenosis. The ECST and CC methods have therefore not been previously validated for use with CEMRA. Although DSA is a purely anatomical technique based on intra-arterial injection, CEMRA is a completely different modality, with different spatial resolution, which relies on physiological blood flow and intravenous injection of contrast medium. This may have an influence on the reader’s perception and measurement accuracy made at different locations in the carotid arterial tree. The main objective of this study was, therefore, to directly compare the NASCET, ECST, and CC measurement methods on CEMRA, with DSA as the reference standard.

Materials and Methods

Patients
Between August 2000 and January 2003, consecutive symptomatic patients suspected of having ICA stenosis on the basis of screening Doppler ultrasound showing at least 50% stenosis in 1 or more arteries and who were scheduled for DSA were prospectively recruited at a single academic institution to undergo CEMRA. Patients with contraindications for magnetic resonance imaging (MRI) such as claustrophobia or metal were excluded. The study was approved by the Local Ethics Research Committee and all patients gave informed written consent.

Diagnostic Tests
Full details of the imaging protocol have been described elsewhere. Briefly, DSA was performed by an attending neuroradiologist, with bilateral selective CCA catheterization, with images of each bifurcation acquired on 4 projections (antero-posterior, lateral, and bilateral 45° obliques) with a resolution of 0.32×0.32 mm.

CEMRA was performed on a 1.5-T machine (Signa CV/I; GE Medical Systems). A bolus-timed 3-dimensional breath-hold fast spoiled gradient echo acquisition with an elliptic-centric phase-encode ordering scheme was performed. Zero-filling interpolation was used, resulting in a final interpolated voxel size of 0.45×0.45×0.40 mm. Each carotid bifurcation was recorded onto film as 4 maximum intensity projections identical to the ones used for DSA.

Image Analysis
Three experienced neuroradiologists, blinded to clinical history and other diagnostic tests, independently reviewed each MRA and conventional angiogram. Patients’ identifiers on films were masked. The DSA images were arranged in a randomized order and reported over a period of 4 months. The CEMRA images were then subsequently reported over the following 4 months, after re-randomization.

For each study, the projection that demonstrated the tightest narrowing was selected. Four measurements were then made: (1) luminal diameter of the normal distal ICA beyond the bulb where the artery wall is parallel; (2) luminal diameter at the site of maximal narrowing; (3) luminal diameter of the estimated original width of the artery at the site of maximal narrowing; and, finally, (4) luminal diameter of the proximal disease-free CCA where the artery wall becomes parallel. The calculations of percentage stenosis according to NASCET, ECST, and CC methods are illustrated in Figure 1.

If no stenosis was present, the degree of stenosis was arbitrarily assigned as 0% for all the measurement methods. When an ICA demonstrated a near occlusion with collapse of the distal lumen, degree of stenosis was assigned as 95% for all methods. For CEMRA, the presence of a signal void was assumed to represent a severe stenosis for each measurement method.

Statistical Analysis
For the NASCET method, stenosis was classified as mild (0% to 49%), moderate (50% to 69%), severe (70% to 99%), or complete occlusion. Analysis of 1001 angiograms by Rothwell et al also showed an almost equivalent relationship between the ECST and CC methods. Therefore, for both ECST and CC methods, stenosis was classified as mild (0% to 64%), moderate (65% to 81%), severe (82% to 99%), or complete occlusion.

All data were entered in SPSS for Windows (V10.1; SPSS Inc) for analysis. For each measurement method, DSA was used as the reference standard. The final reading used was defined as the mean of the 3 independent readers for all 3 methods. Correlation between CEMRA and DSA was calculated using Pearson correlation coefficient and intraclass coefficient. Agreement between the 2 modalities was assessed using the Bland–Altman method and Cohen Kappa statistics. Sensitivity, specificity, and predictive values were calculated to characterize the ability of CEMRA to detect DSA-defined severe stenosis. Finally, interobserver agreement was determined for each measurement method using Kappa statistics.

Results

Patients
One hundred sixty-seven consecutive symptomatic patients (45 women; mean age, 70; range, 41 to 89 years old) were recruited. However, 25 patients were excluded from the final analysis because they failed to complete one of the procedures. Full details of patient characteristics, reasons for exclusion, and image quality have been published elsewhere. One hundred forty-two patients, who completed both modalities, remained, yielding 284 arteries for analysis.

Distribution of Disease
Table 1 shows agreement between CEMRA and DSA according to different stenosis categories for the NASCET, ECST, and CC methods, respectively. With DSA, use of NASCET, ECST, and CC methods would have resulted in 71 cases (NASCET), 57 cases (ECST), and 65 cases (CC) of severe stenosis respectively. With CEMRA, use of NASCET, ECST, and CC methods would have resulted in 91 cases (NASCET), 66 cases (ECST), and 81 cases (CC) of severe stenosis, respectively.
There was excellent and almost identical correlation between CEMRA and DSA irrespective of measurement method (Figure 2). Pearson correlation coefficient was 0.956 ($P<0.001$) for NASCET, 0.954 ($P<0.001$) for ECST, and 0.956 ($P<0.001$) for CC methods. Similarly, intraclass coefficient was 0.978 (95% CI, 0.971, 0.983) for NASCET, 0.976 (95% CI, 0.970, 0.981) for ECST, and 0.977 (95% CI, 0.971, 0.982) for CC methods.

**Agreement**

Bland–Altman plots between CEMRA and DSA for each measurement method are shown in Figure 3. With each method, there was a slight tendency for CEMRA to overestimate stenosis compared with DSA, with mean biases ranging from 0.6% to 2.4%. The limits of agreement, within which 95% of the differences between the 2 modalities are expected to lie, were very similar for all 3 methods, ranging from $-20.0\%$ to $20.3\%$ around the mean biases. Categorial agreement for NASCET (kappa=0.73; 95% CI, 0.67, 0.79), CC (kappa=0.72; 95% CI, 0.65, 0.78), and ECST (kappa=0.70; 95% CI, 0.63, 0.77) methods were also similar in magnitude.

**Diagnostic Performance**

Diagnostic performance of CEMRA for identifying DSA-defined severe stenosis for each measurement method is summarized in Table 2. For specificity, positive predictive values, and negative predictive values, there were no significant differences between the 3 methods. However, the use of ECST method on CEMRA was associated with a much lower sensitivity (78.9%) for detection of severe stenosis compared with NASCET (93.0%), a difference that was statistically significant ($P=0.015$, Fischer exact test). The CC method had an intermediate sensitivity of 87.7%.

**Interobserver Agreement**

For both DSA and CEMRA, there was no significant difference among observers for the different measurement methods. For DSA, mean kappa values for different pairs of readers was 0.84 (0.83, 0.84, and 0.86) for NASCET, 0.83 (0.81, 0.84, and 0.85) for ECST, and 0.84 (0.83, 0.84, and 0.85) for CC methods, respectively. For CEMRA, mean kappa values for different pairs of readers was 0.84 (0.83, 0.83, and 0.85) for NASCET, 0.80 (0.78, 0.79, and 0.82) for ECST, and 0.79 (0.77, 0.79, and 0.82) for CC methods, respectively. These differences for CEMRA were not statistically significant ($\chi^2$ test, NASCET versus ECST, $P=0.298$, NASCET versus CC, $P=0.198$).

**Discussion**

It is well-understood that both the ECST and CC methods result in a higher percentage stenosis compared with...
NASCET, given that the denominator used in NASCET, the distal ICA, is generally smaller than the proximal ICA or CCA. To account for this, Rothwell et al demonstrated the mathematical equivalence of the 3 methods through analysis of a large number of angiograms and mathematical regression techniques. We were therefore surprised that, despite this equivalence, all 3 methods still resulted in significant differences in the prevalence of severe stenosis. For instance, with DSA, only 57 cases would have been classified as severe with the ECST method, compared with 65 severe cases with the CC method and 71 severe cases with the NASCET method. This tendency of the ECST method to "under-report" severity of stenosis compared with NASCET was even more pronounced with the use of CEMRA. Thus, as many as 25 cases (66 severe cases ECST versus 91 severe cases NASCET) would have been potentially managed differently on CEMRA; a statistically significant difference (P<0.005, χ² test). Our study therefore highlights that while mathematical conversions are useful as an approximate guide, in practice, significant errors of categorization are likely, especially taking into account inherent interobserver variability in measuring carotid stenosis. Rothwell et al reported a rate of misclassification of up to 15% in their original series.6

Unfortunately, in the absence of a gold standard method, it is practically impossible to determine which measurement method is most anatomically correct. Although planimetric measurements of stenosis on excised surgical specimens might theoretically represent a reference standard against which to validate angiographic measurement methods, problems such as surgical damage or specimen shrinkage may limit accuracy.13

The reproducibility of the 3 methods on DSA has been previously studied by several authors. Rothwell et al found a consistently greater reproducibility for the CC method compared with both NASCET and ECST.4 This has been the basis for some centers choosing to use the CC method.5 Vanninen et al found a marginally higher interobserver agreement for the ECST method compared with NASCET.14 In our study, on both DSA and CEMRA, there were no differences between the 3 methods which all had excellent reproducibility, a finding consistent with other studies.14,15

With regards to intermodality correlation and agreement, all 3 methods overall performed similarly. Moreover, using the Bland–Altman method we found similar limits of agreement of ~±20% between CEMRA and DSA. This suggests that, irrespective of measurement method, sizable disagreements may occur between CEMRA and DSA, findings consistent with previous literature.16,17 However, a substantial part of this disagreement is actually because of the fact that the measurement methods themselves are prone to significant variability, even with experienced readers.9 CEMRA tended to overestimate severity of stenosis; this was marginally more evident with NASCET, as evidenced by the mean biases shown in Figure 3. As illustrated by the spread of the disagreements (Figure 3), differences between CEMRA and DSA were more pronounced with milder degrees of stenosis but this observation was consistent across all 3 measurement methods.

This is the first study to our knowledge to explore whether diagnostic performance of CEMRA varies according to the measurement method used. All 3 methods also had very similar performance characteristics with regard to specificity and predictive values. However, a finding of clinical and statistical significance was that ECST was associated with a much lower sensitivity (78.9%) compared with both NASCET (93.0%) and CC (87.7%). This may be related to the difficulty of visually estimating the bulbous nature of the unseen normal proximal ICA on CEMRA, with its intrinsically lower spatial resolution. Despite the fact that

Figure 2. Scatterplots illustrating correlation between CEMRA and DSA according to each measurement method. A, NASCET. B, ECST. C, CC methods. Values used for each modality are the mean of the 3 independent readers. Signal voids on CEMRA were excluded because exact numerical values could not be assigned.
highly impressive submillimeter resolution can now be achieved with CEMRA, the true in-plane resolution (0.80×0.90 mm) is still 2- to 3-times less than that of DSA (0.32×0.32 mm).

A limitation of our study is that our readers, despite being experienced neuroradiologists working in a tertiary referral center with a large annual volume of carotid imaging studies, primarily use the NASCET method in routine clinical practice. However, we carefully researched and minimized potential sources of bias before making our measurements, and this is reflected by our high measures of reproducibility for all 3 methods. Results from other groups are, however, needed to confirm the generalizability of our results. Moreover, we used the artery rather than the patient as observational unit. The assumption of independence of each unit is, however, statistically valid because each artery was reported independently from the contralateral side and that stenosis severity between arteries from the same patient is not correlated.

Figure 3. Bland–Altman plots of agreement between CEMRA and DSA for each measurement method. A, NASCET. B, ECST. C, CC methods. Difference in percent stenosis between CEMRA and DSA is plotted against mean of CEMRA and DSA readings. Middle line represents mean bias between CEMRA and DSA. Dotted lines represent upper and lower limits of agreement, within which 95% of the differences between CEMRA and DSA are expected to lie. Signal voids on CEMRA were excluded as exact numerical values could not be assigned.
TABLE 2. Diagnostic Accuracy of CEMRA for Detection of Severe Stenosis According to Methods of Measurement

<table>
<thead>
<tr>
<th>Method</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Positive Predictive Value</th>
<th>Negative Predictive Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>NASCET</td>
<td>93.0% (87.1, 98.9)</td>
<td>88.3% (84.0, 92.6)</td>
<td>72.5% (65.0, 80.0)</td>
<td>97.4% (95.1, 99.6)</td>
</tr>
<tr>
<td>ECST</td>
<td>78.9% (68.4, 89.5)</td>
<td>90.7% (87.0, 94.5)</td>
<td>68.2% (56.9, 79.4)</td>
<td>94.5% (91.5, 97.5)</td>
</tr>
<tr>
<td>CC</td>
<td>87.7% (79.7, 95.7)</td>
<td>89.0% (84.9, 93.1)</td>
<td>70.4% (60.4, 80.3)</td>
<td>96.1% (93.4, 98.7)</td>
</tr>
</tbody>
</table>

95% CIs are quoted within parentheses.

Rothwell et al, in their pooled analysis of individual patient data, recently called for the NASCET method to be adopted as the standard method because of the potential for confusion. Our data show that uniformity of measurement methods is indeed desirable because the use of different methods also potentially results in significant differences in the prevalence of surgically relevant severe stenosis, even allowing for different thresholds for the definition of severe disease. In regard to which method should be adopted, we have shown that overall, no method was clearly superior to any other and all 3 methods are acceptably valid with DSA. However, because the prognostic value of the CC method has been shown indirectly through regression analysis rather than based on primary trial data, we agree with Eliasziw et al and question the value of using this method for decision-making. The CC method in our study did not offer any advantages in terms of reproducibility for DSA and CEMRA. We believe that given the strong trend toward noninvasive techniques, the lower sensitivity for detection of severe stenosis with the ECST method with CEMRA provides small but significant support for the NASCET method to be adopted as gold standard, especially for centers that use CEMRA. This is the most logical alternative given that the most compelling body of evidence for guiding clinical practice for carotid disease, the pooled analysis of individual patient data from the major carotid endarterectomy trials, has been analyzed according to the NASCET method. However, individual centers that routinely use CC or ECST need to be familiar with the NASCET method and associated potential sources of error, and to validate these measurement methods locally in terms of accuracy and interobserver variability before introducing any changes. Similarly, our study highlights that other noninvasive modalities such as computed tomography angiography need to be formally validated according to the measurement method used and that the assumption that all measurement methods are equal if different thresholds are adopted may not necessarily hold true across different imaging modalities.

References

Measuring Carotid Stenosis on Contrast-Enhanced Magnetic Resonance Angiography: Diagnostic Performance and Reproducibility of 3 Different Methods
Jean Marie K.S. U-King-Im, Rikin A. Trivedi, Justin J. Cross, Nicholas J.P. Higgins, William Hollingworth, Martin Graves, Ilse Joubert, Peter J. Kirkpatrick, Nagui M. Antoun and Jonathan H. Gillard

Stroke. 2004;35:2083-2088; originally published online July 8, 2004;
doi: 10.1161/01.STR.0000136722.30008.b1
Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2004 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/35/9/2083

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Stroke can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Stroke is online at:
http://stroke.ahajournals.org//subscriptions/