The Accuracy of Transcranial Doppler in the Diagnosis of Middle Cerebral Artery Stenosis

Jose C. Navarro\textsuperscript{a} Annabelle Y. Lao\textsuperscript{a, b} Vijay K. Sharma\textsuperscript{b, c} Georgios Tsivgoulis\textsuperscript{b, d} Andrei V. Alexandrov\textsuperscript{b}

\textsuperscript{a}Department of Neurology and Psychiatry, University of Santo Tomas Hospital, Manila, Philippines; \textsuperscript{b}Department of Neurology, Barrow Neurological Institute, St. Joseph Hospital, Phoenix, Ariz., USA; \textsuperscript{c}Division of Neurology, National University Hospital, Singapore, Singapore; \textsuperscript{d}Department of Neurology, University of Athens School of Medicine, Eginition Hospital, Athens, Greece

\textbf{Abstract}

\textbf{Background and Purpose:} It was the aim of this study to systematically review available literature on the accuracy of transcranial Doppler (TCD) compared with angiography for the diagnosis of \(\geq 50\%\) middle cerebral artery stenosis in patients with transient ischemic attack or ischemic stroke.

\textbf{Methods:} We performed a systematic review that included original articles published on TCD accuracy from 1982 until the end of December 2005 using angiography as the gold standard. The following measures of diagnostic accuracy were obtained from each primary study: sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV). Weighted mean averages were then calculated from individual results for different velocity cutoffs.

\textbf{Results:} Six papers met our selection criteria. Using laboratory-specific variable mean flow velocity cutoffs, self-reported best accuracy results yield a mean weighted average sensitivity of 92\%, specificity of 92\%, PPV of 88\% and NPV of 98\% for 80 cm/s cutoff. For 100 cm/s cutoff, the sensitivities were 100\%, specificity 97\%, PPV 88\% and NPV 100\%.

\textbf{Conclusions:} Although limited to few reports, this analysis demonstrates fair TCD performance against angiography. Since increasing velocity cutoffs do not yield decreasing sensitivity and increasing specificity, further studies are required to determine optimal velocity values and possibly other criteria such as velocity ratios to develop a screening test with balanced performance parameters.

\textbf{Introduction}

Ultrasound detection of intracranial stenosis helps to identify patients at high risk of recurrent stroke [1–5]. Detection of an arterial stenosis with spectral Doppler is based on focal velocity elevations at the point of maximum arterial narrowing [6, 7]. It is assumed that higher velocities can predict a greater chance of more severe disease on angiography. This approach is well studied with the cervical internal carotid artery disease [8, 9] but has received less scrutiny for intracranial disease.

Numerous articles have been published dealing with the accuracy of transcranial Doppler (TCD) in the diagnosis of a stenosis of the middle cerebral artery (MCA) [10–40]. In these reports, different diagnostic criteria and...
velocity thresholds for the MCA stenosis yielded variable sensitivity and specificity parameters. Furthermore, a recent multicenter clinical study of TCD and magnetic resonance angiography (MRA) to identify ≥50% intracranial stenosis showed positive predictive values (PPVs) of noninvasive testing below the expected PPVs of 55% for TCD and 66% for MRA [41]. To date, no consensus is available regarding a reliable velocity threshold for MCA stenosis.

Therefore, the aim of this paper was to systematically review available literature on the subject of accuracy of TCD compared with angiography for the diagnosis of MCA stenosis. We thought to obtain test performance parameters that could be helpful to attain reliable velocity criteria for MCA stenosis.

**Subjects and Methods**

**Study Selection Criteria**

We first screened all original articles published since the first TCD paper in 1982 [20] until the end of December 2005 that reported a comparison of TCD with any cerebral angiography as the reference imaging standard. To be eligible, the report had to include TCD comparison with angiography as the gold standard, specified percent stenosis measurement and to provide details of patient populations studied. Patient populations eligible for the analysis were adults (18 years or older) of both genders with a history of stroke or transient ischemic attack. The eligible articles had to report TCD sensitivity, specificity, PPV and negative predictive value (NPV) or the likelihood ratios, or present the raw data suitable for recalculation of these values.

**Retrieval of the Relevant Literature**

Electronic searches were carried out using Medline, Embase, Cinahl and Cochrane Library from 1982 to 2005 applying the following key words: MCA stenosis, MCA infarction, transcranial Doppler, cerebral angiography, diagnostic test, gold standard, sensitivity, specificity, positive predictive value, negative predictive value, receiver operator characteristic (ROC) curve and likelihood ratio. Only articles written in English were included in this study. Two authors (V.K.S. and A.Y.L.) screened all retrieved articles. We first summarized all published different diagnostic criteria in original studies that also report the criteria performance against any angiographic studies (Table 1). This information was used to derive the mean flow velocity (MFV) thresholds for subsequent systematic review. We then reviewed all papers for eligibility, and the reasons why certain papers did not meet the inclusion criteria are listed in Table 2. Although TCD also screens for intracranial stenosis, only published articles on TCD will be included in this review for homogeneity.

**Extraction and Display of Data**

Two authors (J.C.N. and A.V.A.) blinded to the authorship and study institutions independently reviewed eligible articles and scored them using the 8-point validity score for methodological quality (Table 3). For each fulfilled criterion, 1 point was given. If the criterion was not fulfilled or the answer was unclear, a score of 0 was given. Disagreements were resolved by discussion with other coauthors to arrive at a consensus assessment.

**Statistical Analysis**

The following measures of diagnostic accuracy were obtained from each primary study: sensitivity, specificity, PPV and NPV. First, we summarized self-reported highest test performance parameters in each selected study and calculated the mean average values of the 'best case scenario' TCD performance regardless of the laboratory-specific velocity cutoffs. We then identified sensitivity, specificity, PPV and NPV for different MCA MFV cutoffs.

<table>
<thead>
<tr>
<th>First author</th>
<th>Criteria</th>
<th>Validation imaging methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>de Bray et al. [10]</td>
<td>peak systolic frequency &gt; 3 kHz 20% segmental increase in peak systolic frequency</td>
<td>intra-arterial angiography</td>
</tr>
<tr>
<td>Ley-Pozo and Ringelstein [12]</td>
<td>MFV ≥ 80 cm/s</td>
<td>digital subtraction angiography</td>
</tr>
<tr>
<td>Carmelingo et al. [13]</td>
<td>MFV ≥ 80 cm/s</td>
<td>digital subtraction angiography</td>
</tr>
<tr>
<td>Rorick et al. [14]</td>
<td>peak systolic velocity &gt; 140 cm/s</td>
<td>magnetic resonance angiography</td>
</tr>
<tr>
<td>Gao et al. [15]</td>
<td>MFV &gt; 80 cm/s peak systolic velocity &gt; 140 cm/s</td>
<td>computed tomography angiography</td>
</tr>
<tr>
<td>Sumanwela et al. [16]</td>
<td>MFV ≥ 100 cm/s prestenotic to stenotic MCA velocity ratio of 1: ≥ 2</td>
<td>digital subtraction angiography</td>
</tr>
<tr>
<td>Felberg et al. [17]</td>
<td>MFV ≥ 80 cm/s</td>
<td>digital subtraction angiography</td>
</tr>
<tr>
<td>Bang et al. [18]</td>
<td>MFV ≥ 80 cm/s</td>
<td>digital subtraction angiography</td>
</tr>
<tr>
<td>Bang et al. [19]</td>
<td>MFV ≥ 80 cm/s</td>
<td>digital subtraction angiography</td>
</tr>
</tbody>
</table>

Peak systolic frequency (1 kHz = 30 cm/s) [45].
from available datasets. When studies reported different measures of diagnostic accuracy for the same MCA MFV cutoff, the weighted average sensitivity, specificity, PPV and NPV were then calculated from individual study results. The Statistical Package for Social Science (SPSS Inc., version 10.0 for Windows) was used for statistical analyses.

**Results**

The literature search originally identified 30 articles, 6 out of which (3 prospective and 3 retrospective) fulfilled our criteria for systematic review [12–14, 17–19]. TC2-62 EME [12, 13], TC22, TC2000 [14], Multigon and Multidop [17] and Transcan EME [18, 19] were used. All of these devices have a 2-MHz nonimaging Doppler probe and no angle correction. Their emission power generally does not affect velocity measurement.

The highest accuracy parameters for the different MFV cutoffs (80, 90 and 100 cm/s) used in the former studies are presented in table 4, together with the weighted average means for combined sensitivities, specificities, PPV and NPV for different MFV cutoffs.

When only prospective studies that evaluated patients in a consecutive fashion were analyzed, the weighted average mean accuracy parameters were higher as compared with the analysis of all the studies combined to-

**Table 2. Reasons for exclusion of retrieved articles from meta-analysis**

<table>
<thead>
<tr>
<th>Reason</th>
<th>Articles excluded</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use of transcranial duplex imaging technologies</td>
<td>6 articles [20–26]</td>
</tr>
<tr>
<td>No angiography done or no use of angiography as reference standard</td>
<td>14 articles [15, 16, 27–37]</td>
</tr>
<tr>
<td>Did not use 2 × 2 table of sensitivity and specificity, or results cannot be derived from the data</td>
<td>4 articles [38–41]</td>
</tr>
<tr>
<td>Total number rejected</td>
<td>25 articles</td>
</tr>
<tr>
<td>Total number eligible for the analysis</td>
<td>6 articles [12–14, 17–19]</td>
</tr>
</tbody>
</table>

**Table 3. Validity score for diagnostic test in selected papers**

<table>
<thead>
<tr>
<th>Reference</th>
<th>Study design</th>
<th>Patient selection</th>
<th>Data collection</th>
<th>Observer blinding</th>
<th>Details on both tests</th>
<th>Verification of test</th>
<th>Data reporting</th>
<th>Details of study population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ley-Pozo and Ringelstein [12]</td>
<td>1 = comparative, 0 = noncomparative</td>
<td>1 = consecutive, 0 = nonconsecutive</td>
<td>1 = prospective, 0 = retrospective</td>
<td>1 = blind, 0 = nonblind</td>
<td>1 = sufficient, 0 = insufficient</td>
<td>1 = complete, 0 = partial</td>
<td>1 = sufficient, 0 = insufficient</td>
<td></td>
</tr>
<tr>
<td>Carmelingo et al. [13]</td>
<td>1 = comparative, 0 = noncomparative</td>
<td>1 = consecutive, 0 = nonconsecutive</td>
<td>1 = prospective, 0 = retrospective</td>
<td>1 = blind, 0 = nonblind</td>
<td>1 = sufficient, 0 = insufficient</td>
<td>1 = complete, 0 = partial</td>
<td>1 = sufficient, 0 = insufficient</td>
<td></td>
</tr>
<tr>
<td>Rorick et al. [14]</td>
<td>1 = comparative, 0 = noncomparative</td>
<td>1 = consecutive, 0 = nonconsecutive</td>
<td>1 = prospective, 0 = retrospective</td>
<td>1 = blind, 0 = nonblind</td>
<td>1 = sufficient, 0 = insufficient</td>
<td>1 = complete, 0 = partial</td>
<td>1 = sufficient, 0 = insufficient</td>
<td></td>
</tr>
<tr>
<td>Felberg et al. [17]</td>
<td>1 = comparative, 0 = noncomparative</td>
<td>1 = consecutive, 0 = nonconsecutive</td>
<td>1 = prospective, 0 = retrospective</td>
<td>1 = blind, 0 = nonblind</td>
<td>1 = sufficient, 0 = insufficient</td>
<td>1 = complete, 0 = partial</td>
<td>1 = sufficient, 0 = insufficient</td>
<td></td>
</tr>
<tr>
<td>Bang et al. [18]</td>
<td>1 = comparative, 0 = noncomparative</td>
<td>1 = consecutive, 0 = nonconsecutive</td>
<td>1 = prospective, 0 = retrospective</td>
<td>1 = blind, 0 = nonblind</td>
<td>1 = sufficient, 0 = insufficient</td>
<td>1 = complete, 0 = partial</td>
<td>1 = sufficient, 0 = insufficient</td>
<td></td>
</tr>
<tr>
<td>Bang et al. [19]</td>
<td>1 = comparative, 0 = noncomparative</td>
<td>1 = consecutive, 0 = nonconsecutive</td>
<td>1 = prospective, 0 = retrospective</td>
<td>1 = blind, 0 = nonblind</td>
<td>1 = sufficient, 0 = insufficient</td>
<td>1 = complete, 0 = partial</td>
<td>1 = sufficient, 0 = insufficient</td>
<td></td>
</tr>
</tbody>
</table>

**Table 4. Sensitivity, specificity, PPV and NPV of different cutoff points of MCA MFV for the prediction of MCA stenosis in prospective and retrospective studies**

<table>
<thead>
<tr>
<th>MCA MFV cutoff</th>
<th>Studies</th>
<th>MCA studied</th>
<th>Sensitivity %</th>
<th>Specificity %</th>
<th>PPV %</th>
<th>NPV %</th>
</tr>
</thead>
<tbody>
<tr>
<td>80 cm/s</td>
<td>[12–14, 17–19]</td>
<td>616</td>
<td>82.9</td>
<td>85.4</td>
<td>75.8</td>
<td>90.2</td>
</tr>
<tr>
<td>90 cm/s</td>
<td>[14]</td>
<td>99</td>
<td>75.0</td>
<td>91.0</td>
<td>56.0</td>
<td>96.0</td>
</tr>
<tr>
<td>100 cm/s</td>
<td>[17]</td>
<td>113</td>
<td>100</td>
<td>97</td>
<td>88.0</td>
<td>100</td>
</tr>
<tr>
<td>120 cm/s</td>
<td>[7]</td>
<td>113</td>
<td>68.7</td>
<td>100</td>
<td>100</td>
<td>95.0</td>
</tr>
</tbody>
</table>

Data are weighted average means, except MCA studied.
gether (sensitivity 91.8 vs. 82.9%, specificity 92.2 vs. 85.4%, PPV 88 vs. 75.8%, and NPV 98.4 vs. 90.2%) as shown in table 5. The former analyses indicate that the validity of TCD seems to improve when evaluated in a prospective setting.

Among the different velocity criteria, 100 cm/s have shown the most balanced accuracy parameters with a sensitivity of 100%, specificity of 97%, PPV of 88% and NPV of 100%.

**Discussion**

Our study showed variable accuracy parameters with increasing MFV for a significant (≥50%) MCA stenosis. Average TCD performance parameters were found to be below the best results reported in individual studies. Our analysis also showed that only one MFV cutoff was evaluated in several studies. Using only the highest quality papers, it was noted that the accuracy parameters increased, which are more reflective of the previously published data regarding TCD accuracy.

Findings in the multicenter Stroke Outcomes and Neuroimaging of Intracranial Atherosclerosis and Warfarin Aspirin Symptomatic Intracranial Disease (SONIA-WASID) study showed a PPV of only 55% [41], which was lower than the target value expected for the SONIA trials and reported in other studies. Our study showed that the expected average TCD performance based on previously published series can have a PPV of 76%. Factors responsible for this incongruence are likely to be (but not limited to) the following. First, most of the subjects enrolled in the SONIA study were Caucasians who have less preponderance of intracranial disease. Second, a verification bias could play a role since angiography was not done in patients with normal TCD results. Finally, a single MFV cutoff may not take into account systemic hemodynamics or collaterals that can increase the number of false-positive TCD results.

Suboptimal TCD performance against angiography is unsettling, and besides being attributable to operator dependency of TCD, it also raises questions as to the validity of a single or dual projection digital subtraction angiography (DSA) in identifying mild and moderate intracranial disease. Our findings and those of the SONIA study may also mean that either the gold standard of invasive angiography poorly reflects the nature of intracranial disease, or current TCD velocity criteria are too simplistic and should be further refined against multiplanar DSA.

The obvious limitation of this meta-analysis is too few eligible reports that include a comparison with DSA, as well as relatively limited numbers of patients in each report. Given the publication bias towards the best results, our findings should be interpreted with caution; yet, they show deficiencies in current data on TCD performance and stress the need of further validation studies of various velocity parameters and ratios.

Among these other parameters, the peak systolic velocity has been validated using MRA in only one study [15]. Interestingly, elevated peak systolic velocity (>140 cm/s) was an independent prognostic factor in predicting the risk of recurrent cerebral ischemic events [42]. In an- other study, Arenillas et al. [3] used MFV to predict the risk of recurrent ischemic events. The values for the MFV cutoffs were mild (80–120 cm/s), moderate (120–140 cm/s) and severe (>140 cm/s). Patients with MFV ≥120 cm/s had more clinical events in this study. In the future, it would be important to determine the severity of intracranial disease these velocities represent on angiography. One way to compensate for velocity changes due to hemodynamic factors is to deploy velocity ratios in a similar way as proposed in carotid testing [9] and grading cerebral vasospasm [43, 44]. So far, only one paper has deployed the affected:nonaffected MCA ratio of ≥2 to support velocity findings [17]. Previously suggested asymmetry indexes may also prove helpful in future studies [11, 33].

From published reports, it is not yet possible to conclude that higher velocity thresholds yield lower sensitiv-

<table>
<thead>
<tr>
<th>MCA MFV cutoff</th>
<th>Studies</th>
<th>MCA studied</th>
<th>Sensitivity %</th>
<th>Specificity %</th>
<th>PPV %</th>
<th>NPV %</th>
</tr>
</thead>
<tbody>
<tr>
<td>80 cm/s</td>
<td>[12–14]</td>
<td>315</td>
<td>91.8</td>
<td>92.2</td>
<td>88.8</td>
<td>98.4</td>
</tr>
<tr>
<td>100 cm/s</td>
<td>[14]</td>
<td>113</td>
<td>100</td>
<td>97</td>
<td>88.0</td>
<td>100</td>
</tr>
</tbody>
</table>

Data are weighted average means, except MCA studied.
ity and higher specificity as they do in the cervical carotid testing [8]. Second, the most commonly reported cutoff of 80 cm/s may yield fair results compared with best reported series. Further studies are required to determine the velocity threshold and other criteria (i.e. ratios) that could help develop a screening test with balanced performance parameters appropriate to patient populations.

Acknowledgements

A.Y.L., MD, received a scholarship grant from the Tan Yan Kee foundation, Philippines. V.K.S., MD, received a financial grant and fellowship from the National Healthcare Group and National University Hospital, Singapore. G.T. received a financial grant and fellowship from the Neurology Department of Eginition Hospital, University of Athens, Greece. The authors would like to acknowledge Michael D. Hill, MD, for his guidance and statistical input.

References


