Reproducibility of Functional Transcranial Doppler Sonography in Determining Hemispheric Language Lateralization

S. Knecht, MD; M. Deppe, PhD; E.-B. Ringelstein, MD; M. Wirtz, PhD; H. Lohmann; B. Dräger; T. Huber; H. Henningsen, MD

Background and Purpose—Since functional transcranial Doppler ultrasonography (fTCD) allows convenient and fully automated quantification of language lateralization, it seems ideal for longitudinal studies of perfusion changes during deterioration as well as recovery of language functions. However, during serial examinations, the technical, stochastic, and physiological variabilities of cerebral blood flow velocities (CBFV) have to be considered. Therefore, before fTCD is accepted as a tool for evaluation of changes in lateralization in the diseased state, its reliability in healthy subjects needs to be determined.

Methods—We performed fTCD during a word generation task based on a previously validated technique with automated calculation of the averaged CBFV differences in the middle cerebral arteries providing an index of lateralization (LI).

Results—(1) The accuracy of the LI as assessed by the confidence interval was better than 1% of the mean hemispheric difference. (2) On repeated examination, LIs obtained from 10 subjects showed a high test-retest reproducibility (Pearson product moment correlation coefficient $r=0.95$, $P<0.0001$). (3) On 10 repeated assessments of LI in the same subject, no practice effects were detected.

Conclusions—Functional TCD is a suitable and very robust tool for the longitudinal quantitative measurement of cerebral language lateralization. (Stroke. 1998;29:1155-1159.)

Key Words: functional imaging ▪ language ▪ reliability ▪ reproducibility ▪ ultrasonography, Doppler

Recent developments in fTCD have made it possible to noninvasively and quantitatively measure lateralization and time course of brain activation. The technique of fTCD is based on the linkage of cerebral activation and perfusion, a principle also underlying fMRI and oxygen-15 positron emission tomography. Perfusion changes result in corresponding blood flow velocity modulations in the supporting basal intracranial arteries that can be continuously measured by fTCD.

Analysis of cerebral functional lateralization by fTCD as performed in the present study constitutes a fully automated, objective procedure. The quantitative measures obtained by fTCD are not biased by defining variable statistical thresholds, as is often the case in the analysis of fMRI data. Given the availability and easy applicability of Doppler sonography equipment, fTCD offers a possibility for the noninvasive assessment of lateralization of language and other higher brain functions. Such information would be valuable for the planning of neurosurgical procedures or to conveniently estimate the variability of lateralization of different cognitive functions in the general population. fTCD could also be of interest for the quantitative and longitudinal measurement of partial shifts of cerebral activation during the course of learning procedures or recovery from neurological deficits. For example, during recovery from aphasia, recruitment of areas contralateral to the lesioned hemisphere has been observed. Involvement of the unaffected hemisphere has also been demonstrated in other types of cerebral plasticity.

However, for fTCD to reliably detect potential shifts of language lateralization in patients in the course of their disease or recovery, two preconditions need to be fulfilled: (1) functional TCD must be sensitive enough to detect even small changes in hemispheric perfusion differences; and (2) in normal subjects, fTCD must yield a measure of lateralization that is highly reproducible over time and is not subject to random variability, learning, or habituation.

In the present study we investigated the accuracy and reproducibility of fTCD in healthy subjects using a word generation paradigm previously validated by direct comparison with the standard assessment for language lateralization, the transfemoral intracarotid amobarbital injection.

Subjects and Methods

Subjects
Four female and seven male subjects (aged 19 to 32 years) participated in the study after giving informed consent. All were healthy...
Protocol
Assessment of hemispheric language dominance was performed by a word generation task previously validated by direct comparison with surgical resection in patients with medically intractable epilepsy and who were evaluated for resective surgery.12 Subjects were seated in a reclining chair in front of a blank computer screen. Five seconds after a cueing tone, a letter was presented for 2.5 seconds. The language task consisted of silently finding as many words as possible starting with the displayed letter. After a second auditory signal 15 seconds after presentation of the letter, subjects had to articulate the words they had found to control cooperation in the task. All words had to be reported within a 5-second time period. After a relaxation period of 30 seconds, another letter in random order was presented in the same way. "Q," "X," and "Y" were excluded because very few words in German start with these letters. Every letter was displayed only once.

Evaluation of Statistical Differences Between Examinations by t Test

<table>
<thead>
<tr>
<th>Subject</th>
<th>Exam Interval</th>
<th>N</th>
<th>LI</th>
<th>sLI</th>
<th>P</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>20</td>
<td>-0.66</td>
<td>0.31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2 mo</td>
<td>20</td>
<td>-1.05</td>
<td>0.62</td>
<td>0.5802</td>
<td>NS</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>20</td>
<td>1.54</td>
<td>0.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>12 mo</td>
<td>22</td>
<td>2.54</td>
<td>0.54</td>
<td>0.1689</td>
<td>NS</td>
</tr>
<tr>
<td>4</td>
<td>2 mo</td>
<td>20</td>
<td>4.26</td>
<td>0.72</td>
<td>0.6283</td>
<td>NS</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>22</td>
<td>3.00</td>
<td>0.37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>14 mo</td>
<td>22</td>
<td>2.65</td>
<td>0.41</td>
<td>0.5313</td>
<td>NS</td>
</tr>
<tr>
<td>6</td>
<td>1 wk</td>
<td>23</td>
<td>3.01</td>
<td>0.46</td>
<td>0.8778</td>
<td>NS</td>
</tr>
<tr>
<td>7</td>
<td>1 wk</td>
<td>23</td>
<td>3.26</td>
<td>0.45</td>
<td>0.1299</td>
<td>NS</td>
</tr>
<tr>
<td>8</td>
<td>1 h</td>
<td>21</td>
<td>1.59</td>
<td>0.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>20</td>
<td>0.60</td>
<td>0.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1 wk</td>
<td>23</td>
<td>2.20</td>
<td>0.62</td>
<td>0.5183</td>
<td>NS</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>20</td>
<td>-0.60</td>
<td>0.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>14 mo</td>
<td>19</td>
<td>4.49</td>
<td>0.67</td>
<td>0.0564</td>
<td>NS</td>
</tr>
</tbody>
</table>

See “Subjects and Methods” for definitions.

Figure 1. Schematic of averaging procedure. The top panel shows the way in which relative event-related CBFV changes in both MCAs during individual repetitions of the task (1 through N) are collected and averaged. The bottom panel depicts the subtraction of averaged CBFV changes in the right MCA, providing a measure of the mean interhemispheric CBFV difference (dVr) over the course of the task, with the corresponding standard deviations at each point in time (gray shading). This mean interhemispheric CBFV difference curve is time locked to fixed events during the task, ie, presentation of the cue, the letter, silent word generation in response to a given letter, and onset of speaking. Calculation of the LI during word generation based on Td and sLI during Tmax is explained in detail in “Subjects and Methods.”

Functional TCD
TCD assessment was performed in a manner previously described in detail.13 Briefly, a commercially available dual transcranial Doppler ultrasonography device (Multi-Dop T, DWL) was used. The MCAs were insonated at a depth of 50 mm with two 2-MHz transducer probes attached to a headband and placed at the temporal skull windows bilaterally (Figure 1). Details of the insonation technique, particularly the correct identification of the MCA, have been published elsewhere.25 The spectral envelope curves of the Doppler signal were recorded with a rate of 28 sample points per second and stored for off-line processing with the TCD-analysis computer program Average.13 Data analysis, ie, reduction of spontaneous signal fluctuations, filtering, and artifact reduction, was performed as previously described in detail.13 Epochs containing CBFV values outside the range of 30% to 200% of the mean flow velocity were rejected from further processing. The remaining data were integrated over the corresponding cardiac cycles, were segmented into epochs that related to the cueing tone, and were then averaged. The epochs were set to begin 15 seconds before and to end 35 seconds after the cueing tone. The mean velocity in the 15-second precueing interval (Vprecue) was taken as the baseline value. The relative CBFV changes (dV) during cerebral activation were calculated by the formula \( dV(t) = 100\% \frac{V(t) - V_{\text{precue}}}{V_{\text{precue}}} \) (formula 1), where \( V(t) \) is the CBFV over time.

Laterality Index
Figure 1 gives a schematic overview of the procedure by which the LI was obtained. Mathematically the TCD LI was specified by the formula:
(2) \[ LI = \frac{1}{N} \sum_{i=1}^{N} LI^{(i)} \] with \[ LI^{(i)} = \frac{1}{T_{int}} \int_{t_{1}^{i} + 0.5 T_{int}}^{t_{2}^{i} + 0.5 T_{int}} \Delta V^{(i)}(t) dt \]

where

(3) \[ \Delta V^{(i)}(t) = dV^{(i)}(t)_{\text{left}} - dV^{(i)}(t)_{\text{right}} \]

is the side-to-side difference between the relative velocity changes of the left and right MCAs of epoch number \( i = 1, \ldots, N \) (\( N \) = number of epochs). \( t_{1}^{i} \) and \( t_{2}^{i} \) represent the latency of the absolute maximum of \( \Delta V(t) \) during an interval of 8 to 18 seconds after cueing, ie, during verbal processing. For integration a time period of \( T_{int} = 2 \) seconds was chosen.

Assessment of Accuracy

Since the LI represents the mean of \( N \) interhemispheric stochastic perfusion differences from repeated runs involving word generation, its confidence interval can be deduced, under the assumption of normal distribution, from the standard error

(5) \[ \sigma_{LI} = \sigma_{LI}/\sqrt{N} \]

where

(6) \[ \sigma_{LI} = \sqrt{\frac{\sum_{i=1}^{N} (LI^{(i)} - \text{mean}(LI))^{2}}{N-1}} \]

corresponds to the standard deviation of LI. For example, the 95% confidence region (\( P < 0.05 \)) of LI can be estimated by \( LI \pm 1.96 \sigma_{LI} \). The frequency distribution of 165 laterality indices \( LI^{(i)} \) obtained from 10 repetitions in the same subject was gaussian shaped, supporting the assumption of normal distribution. The results of the Kolmogorov-Smirnov test and Shapiro-Wilks’ W test for detection of deviations from normal distribution accorded with normal distribution. Therefore, \( s_{LI} \) was regarded as a statistical estimator for the immanent uncertainty of LI, ie, the predicted accuracy of the applied technique for estimation of language.

Measures of Reproducibility

Reproducibility of language lateralization was performed in several ways: First the LI was analyzed on the basis of two measurements in 10 healthy subjects. For test-retest reproducibility, the Pearson product moment correlation coefficient

(7) \[ r = \frac{\sum_{i=1}^{10} (LI_{i}^{(1)} - \text{mean}(LI_{i})) (LI_{i}^{(2)} - \text{mean}(LI_{i}))}{\sqrt{\sum_{i=1}^{10} (LI_{i}^{(1)} - \text{mean}(LI_{i}))^{2} \sum_{i=1}^{10} (LI_{i}^{(2)} - \text{mean}(LI_{i}))^{2}}} \]

was used.\(^{25}\) \( LI_{i}^{(l)} \) represents the LI of the subject with index \( k = 1, \ldots, 10 \) of the first (\( l = 1 \)) or the second (\( l = 2 \)) examination.

(8) \[ LI^{(l)} = \frac{1}{10} \sum_{k=1}^{10} LI_{k}^{(l)} \]

is the corresponding average laterality from the 10 subjects during the first or the second examination. The correlation coefficient \( r \) is based on the correlation between individual LIs from repeated examinations relative to the variances of the indices.

Additionally, all measurements were submitted to a Student’s \( t \) test to evaluate whether in any subject a statistically significant difference in the LI was present on repeated examination (Table). LIs from these repeated measurements in 10 different subjects were further evaluated by ANOVA to determine whether the statistically deduced confidence region \( s_{LI} \) for the laterality measure LI describes the true uncertainty in LI. In this case variations in LI during sequential measurements can be explained by the estimated stochastic fluctuations \( s_{LI} \). A two-way factorial design was used with the following factors: subject, examination number, and subject * examination number. (The symbol * indicates interactions between factors.) Second, LIs from 10 successive examinations in a single subject were evaluated by a one-way ANOVA (ANOVA with the factor: examination number) to determine whether the statistically deduced confidence region \( s_{LI} \) for the laterality measure LI also describes the true uncertainty in LI. In this case variations in LI of successive measurements should similarly be explainable by the estimated stochastic fluctuations \( s_{LI} \). Additionally, we tested for time-dependent trends in the lateralization, eg, due to learning or habituation. This was done by linear regression of the 10 subsequent LIs in this subject. The regression function, ie, the dependence of LI on the number of examinations, as well as its 95% confidence interval, was calculated.

Results

Accuracy

The accuracy \( s_{LI} \) for detecting relative blood flow velocity differences between the left and right MCA by fTCD varied between 0.3% and 1.0% of the mean for the 20 examinations on the 10 different subjects (Table). The confidence interval \( s_{LI} \) for the 10 repeatedly measured LIs in a single subject varied between 0.55% and 1.1% (mean, 0.79%). On average, the extent of interhemispheric CBFV differences during word generation was approximately 3% in magnitude, ranging from −3.6% (right-dominant subjects) to 4.4% (left-dominant subjects).

Reproducibility

Figure 2 demonstrates that the indices of relative language lateralization obtained from 10 subjects were highly reproducible when reassessed 1 hour to 14 months after their first examination (Pearson product moment correlation coefficient: \( r = 0.95, P < 0.0001 \)). On testing for statistical differences between LI on the first and the second examinations by ANOVA (\( P = 0.48 \)) or Student’s \( t \) test (Table), no statistically significant difference (ie, \( P > 0.05 \)) was found in any individual. Moreover, the correlation analysis of differences between the first and the second examinations consistently revealed no dependence of the time interval (\( \Delta T \)) between examinations and these differences (regression function \( \Delta LI = 0.9 \times 0.09 \times \Delta T \); correlation \( r = 0.11; P = 0.78 \)).

In none of 10 repeated fTCD examinations in the same subject was a statistically significant variation of the
The TCD measurement of the CBFV is dependent on the angle of insonation. Changes of this angle from 0° to 30° by variability of probe position or arterial anatomy can render differences in the calculated, absolute CBFV in the magnitude of up to 15% between examinations or sides. Also, in a narrowed arterial segment incidentally insonated during the test, the absolute velocity increase in blood flow due to cerebral activation would be greater than in a regular segment. This is why flow velocities used for statistical analysis were normalized. fTCD is a tool for the evaluation of perfusion modulations associated with cerebral activation. It therefore need not exploit absolute but relative changes of CBFV. Flow velocities at rest were set as zero baseline, and CBFV changes during the activated state were expressed as values in percentages relative to this baseline. The use of relative CBFV values eliminated the variability associated with changes in insonation angle or vessel diameter.

Functional CBFV data display a stochastic variability because the averaged CBFV curves represent a limited sample from a signal that is affected by a number of cardiovascular fluctuations such as the heart beat, modulations related to breathing, and other low-frequency modulations such as the B-waves and Mayer waves. These cardiovascular fluctuations are of a considerably greater amplitude (approximately 50% of the mean signal) than the changes related to neuropsychological activation (approximately 3% during word generation). However, by heart cycle integration and subsequent averaging, variability of data can effectively be reduced. An additional procedure to further reduce variability between individual recordings was the evaluation of the relative side-to-side difference in CBFV increase, ie, by subtraction of the relative CBFV increase in the right from the left MCA. By so doing, effects from systemic cardiovascular fluctuations on CBFV were completely eliminated. As demonstrated above, the LI obtained by these mathematical procedures had a low standard error, ie, an accuracy better than 1% of the mean CBFV. With respect to longitudinal studies, we therefore expect to reliably detect shifts in language-related hemispheric perfusion exceeding 1% of the mean CBFV.

Physiological changes over time are another major source of variability between fTCD examinations. Some of the cardiovascular modulations probably constitute a nonspecific autonomic response to the subjects’ confrontation with the task. The heart beat, for example, shows a biphasic modulation that is paralleled by biphasic CBFV changes. Autonomic responses are subject to habituation and are thus likely to change from one examination to the next. The cerebral hemodynamic response is also influenced by diurnal fluctuations or by substances such as nicotine. Since we have not controlled for these aspects, they are likely to add to the variability of CBFV. However, since these factors affect the entire cardiovascular system, they again manifest as bilateral, synchronous changes in CBFV and can also be eliminated by calculating the relative side-to-side difference in CBFV increase.

Practice effects could be another source of physiological variability between examinations. Repeating the same language task several times could reduce cognitive demands or improve performance. Either effect could result in a systematic change of CBFV increases during successive fTCD
evaluations. Conceivably, this may even manifest predominately unihemispherically. However, in neither the 10 subjects reexamined a second time nor in the single subject reexamined 10 times was there any decrease or increase of the relative side-to-side CBFV increase during word generation. We do not know whether the repetition of the language task is in fact associated with a change in cognitive demand or effort. If it were, one would have to speculate that the stability of the language LI could be due to a ceiling effect of cerebral perfusion and blood flow velocity increases. Moderate cognitive effort would thus result in a maximal blood flow increase within a delimited activated area. Changes in effort between moderate and maximal strength would subsequently not change this local effect because regional arterial dilatation and blood flow velocity are already at their maximal level. Global changes of perfusion associated with changes in effort would be canceled out because of the side-to-side subtraction algorithm used in the analysis. Alterations in the language LI would occur only if the size or the interhemispheric distribution of the regions with an increased perfusion changed.

In conclusion, perfusional lateralization related to word generation—despite technical, statistical, and physiological variability—can reproducibly be quantified by fTCD. Since the reproducibility of this method falls within the limits of its predictable accuracy, fTCD in longitudinal studies can be expected to detect shifts in language lateralizations with an accuracy of 1% of the mean perfusion in a cerebral vascular territory. On the basis of these results, future serial studies with this simple technique will be able to establish the extent to which interhemispheric functional redistribution plays a role in the recovery of function.

Acknowledgments

This study was supported by the Deutsche Forschungsgemeinschaft Kn 2854/1, the Bennigsen-Foerder grant from the Ministry of Science and Research of Nordrhein-Westfalen, and a grant by Innovative Medizin Forschung, Germany.

References

Reproducibility of Functional Transcranial Doppler Sonography in Determining Hemispheric Language Lateralization

*Stroke*. 1998;29:1155-1159
doi: 10.1161/01.STR.29.6.1155

*Stroke* is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 1998 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/29/6/1155

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/