



Hemispheric language lateralization in presurgical patients with temporal lobe epilepsy: Improving the retest reliability of functional transcranial Doppler sonography

Nadine Conradi^{a,*}, Anke Hermsen^a, Kristina Krause^b, Iris Gorny^b, Adam Strzelczyk^a, Susanne Knake^b, Felix Rosenow^a

^a Epilepsy Center Frankfurt Rhine-Main, Department of Neurology, Goethe University, Frankfurt am Main, Germany

^b Epilepsy Center Hessen, Department of Neurology, Philipps-University Marburg, Marburg, Germany

ARTICLE INFO

Article history:

Received 9 April 2018

Revised 9 August 2018

Accepted 12 August 2018

Available online 11 September 2018

Keywords:

Functional transcranial Doppler sonography

Hemispheric language lateralization

Reproducibility

Retest reliability

Qualitative data analysis

Epilepsy

ABSTRACT

Since its introduction, functional transcranial Doppler sonography (fTCD) has been extensively applied in research and clinical settings and has now become part of the routine presurgical work-up of patients with epilepsy. Because of its importance in planning neurosurgical interventions and predicting possible cognitive risks, the reproducibility of fTCD in determining hemispheric language lateralization (HLL) has to be ensured. In the present study, fTCD was performed twice in 33 initially lateralized patients with temporal lobe epilepsy (TLE) as part of their presurgical work-up, using a standard word generation paradigm.

Initially, the standard analysis, including only the *statistical* examination of fTCD data, was applied, and a rather poor retest reliability of $r = 0.41$ was obtained ($p = 0.017$). Because of doubts concerning appropriate task performance in some patients, subsequently, a two-step data analysis was introduced, including an additional *qualitative* evaluation of fTCD data regarding (1) instruction-compliant task performance, (2) sufficient quality of the baseline phase, and (3) adequate increase in cerebral blood flow velocity (CBFV) during the activation phase. Attributable to a more valid interpretation of fTCD data after the application of the qualitative step, the reproducibility of HLL significantly improved ($p = 0.007$) to a high retest reliability of $r = 0.84$ ($p < 0.000$).

In clinical settings, psychological and situational factors seem to strongly influence the reproducibility of fTCD determining HLL. Accordingly, we highly recommend the complementation of the standard *statistical* examination of fTCD data by an additional *qualitative* evaluation (two-step data analysis), as this extra security is particularly desirable because of its direct implications for the further evaluation of neurosurgical interventions.

This article is part of the Special Issue "Individualized Epilepsy Management: Medicines, Surgery and Beyond".

© 2018 Elsevier Inc. All rights reserved.

1. Introduction

Since its introduction, functional transcranial Doppler sonography (fTCD) has been extensively applied in research and clinical care [1], as it provides a noninvasive and cost-effective technique to determine hemispheric language lateralization (HLL). Because of its easy application and high tolerability, fTCD can be used repetitively, including in children and patients with cognitive impairment [2–4]. However, fTCD cannot be applied in about 5–10% of patients who lack sufficient temporal bone windows [5]. This difficulty can be overcome by giving an intravenous echo-enhancing agent (SonoVue), which allows the determination of HLL in 7 of 10 patients with poor temporal bone windows [6].

In patients with temporal lobe epilepsy (TLE), it is of special interest to determine HLL; firstly, because of the relatively high proportion of patients with TLE with atypical HLL due to hemispheric shifting of language functions, presumably due to left temporal lobe dysfunction-induced plasticity [7,8]; secondly, because HLL, together with other factors (such as side of planned surgery, magnetic resonance imaging (MRI) findings, and neuropsychological baseline performance), plays an important role in the prediction of postsurgical cognitive impairment in patients with TLE [9,10].

As fTCD is increasingly established as part of the routine presurgical work-up of these patients [23], it is essential to assure the reproducibility of HLL determined by fTCD.

The validity of the method has been investigated in several comparative studies with functional MRI (fMRI) or Wada test used as gold standard, and promising results have been obtained [11–14]. In contrast, the methods' retest reliability has rarely been investigated, and only two

* Corresponding author at: Epilepsy Center Frankfurt Rhine-Main, Department of Neurology, Goethe University, Schleusenweg 2-16, 60528 Frankfurt am Main, Germany.
E-mail address: Nadine.Conradi@kgu.de (N. Conradi).

studies have assessed the reproducibility of fTCD in determining HLL by using a word generation paradigm [13–15]. Because of rather small sample sizes and study populations solely consisting of healthy participants, further research in this field was needed.

In clinical routine, fTCD is repeated in patients with atypical or no clear lateralization, in order to verify or improve results of the first measurement and to avoid the necessity of using additional diagnostic techniques such as fMRI or Wada test. However, the repetition of fTCD measurements only makes sense if the methods' retest reliability is sufficient in general. Thus, in this study, fTCD was applied twice in *initially lateralized* patients, i.e., in patients with already clear lateralization results in the first fTCD, in order to investigate the consistency of results in a larger, well-defined sample of presurgical patients with TLE.

As previously reported [13–16], several factors possibly influencing the retest reliability of fTCD exist: firstly, technical issues, such as inconsistency in the fTCD procedure or large variability of the probe positions between the two measurements, have to be considered and avoided if possible. Secondly, physiological aspects, such as fluctuations in cerebral blood flow velocity (CBFV) due to arousal-related modulations of the heart rate, the recent use of nicotine as well as hormonal changes during the menstrual cycle of female participants [17], should be taken into account if this information is available. Thirdly, psychological and situational factors, such as the patients' ability to concentrate on the task, as well as aspects of motivation and excitement during test situations, have to be observed as potential sources of variability between measurements.

In order to control specifically for psychological and situational aspects of fTCD measurements, we introduced a two-step data analysis, including the standard *statistical* examination, complemented by an additional *qualitative* evaluation of fTCD data.

As fTCD determining HLL was part of the routine presurgical workup, this study is also part of continuous quality assurance, as recommended by the guidelines of the working group for presurgical epilepsy diagnosis and operative epilepsy treatment [18].

2. Material and methods

2.1. Patients

In the current study, fTCD was performed twice in 33 patients (16 men, 17 women; mean age: 36.10 years, standard deviation (SD) = 11.35) with TLE (61% left-sided, 30% right-sided, 9% bihemispheric) as part of presurgical work-up. All patients were initially lateralized, i.e.,

already showed clear lateralization results in the first fTCD. The second measurement was applied 1 to 94 days after the initial examination (median = 7 days), and data were analyzed retrospectively. As determined by the Edinburgh Handedness Inventory (EHI) [19], 61% of patients were consistently right-handed (EHI ≥ 50), 15% were consistently left-handed (EHI ≤ -50), and 24% were ambidextrous.

Initially, fTCD data of all patients were statistically evaluated using the well-established standard analysis (see Fig. 1a). Subsequently, a two-step analysis was introduced, and all patients underwent an additional qualitative examination of fTCD data (see Fig. 1b).

As fTCD data of the patients with TLE were collected as part of their routine presurgical workup during video EEG monitoring and afterwards, data were anonymized and analyzed retrospectively, no ethics committee approval was obtained in the first place.

2.2. fTCD

The determination of HLL by fTCD was performed in a standardized fashion, as previously described [7,20]. Briefly, by fixing two 2-MHz ultrasonic probes on both sides of the head with the help of a flexible hat, CBFV in both middle cerebral arteries (MCAs) was measured simultaneously and continuously through temporal bone windows by a standard transcranial Doppler sonography (TCD) device (Pioneer, EME, Medilab, Würzburg, Germany) at a depth between 50 and 60 mm.

A cued word generation paradigm consisting of three phases (each initiated by a cueing tone) was used. Firstly, patients were instructed to silently generate words for 20 s starting with a letter displayed on a computer screen for 5 s (word generation period = activation phase). Secondly, patients were asked to name the generated words loudly for 5 s (speech period = compliance control phase). Thirdly, a relaxing time of 30 s followed (resting period = baseline phase). The whole measurement consisted of 20 epochs (each including all three phases) and took about 20 min in total. Letters were presented randomly.

By synchronizing the stimulation software with the TCD device, markers were recorded at the beginning of each epoch. Later, the spectral envelope curves of the Doppler signal were analyzed offline with the automated software AVERAGE [21]. After an artefact rejection and heart cycle integration, the Doppler signal was segmented into the 20 epochs. The mean CBFV changes during the activation phase relative to the baseline phase were calculated and averaged across the 20 epochs. By analyzing the difference in relative CBFV changes between the left and right MCAs during the activation phase, the laterality

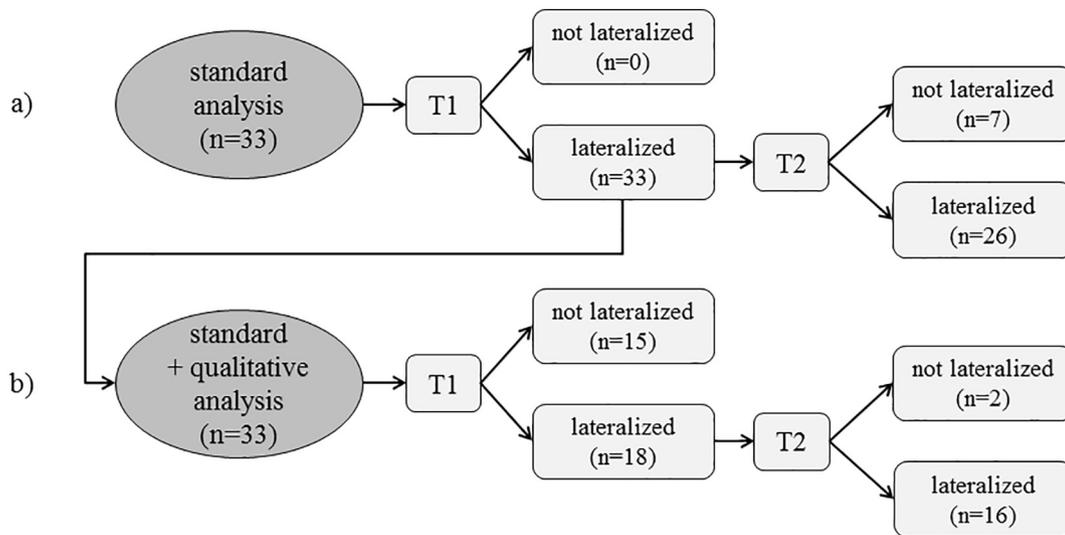


Fig. 1. Flowchart of patients taken for retest reliability analysis. a) fTCD data of all patients (n = 33) were analyzed statistically, and retest reliability between the first (T1) and second (T2) measurements was computed for patients with clear lateralization at T1 (n = 33). b) All patients (n = 33) underwent an additional qualitative analysis of fTCD data, and retest reliability between T1 and T2 was computed for patients with clear lateralization at T1 (n = 18).

index (LI) of the word generation period and its standard error of the mean (SEM) were computed. For a more detailed review, see the work of Deppe et al. [22].

2.3. Standard analysis

According to the standard statistical analysis of fTCD data, 95% confidence intervals for the LIs were computed. Left HLL was expected if the LI and its confidence interval were clearly positive ($LI \pm 1.96 * SEM \geq 0.2$). Correspondingly, if the LI and its confidence interval were clearly negative ($LI \pm 1.96 * SEM \leq -0.2$), right HLL was concluded. In all other cases, no lateralization was assumed.

2.4. Qualitative analysis

For the first time, an additional qualitative evaluation of fTCD data was applied by supplementary rating of the quality of measurements regarding three categories.

- (1) Instruction-compliant task performance: during the measurement, behavioural observation was used to assess whether the patient paid attention to the letter displayed on the computer screen and seemed concentrated on the task in the silent word generation period. Furthermore, subsequent to the examination, all patients were asked a) how good they could relax during the resting period and b) whether they found it difficult to come up with as many words as possible in the speech period. Furthermore, the amount of generated words was qualitatively compared with the patients' educational level and results of word fluency and processing speed tasks in the preceding neuropsychological assessment.
- (2) Sufficient quality of the baseline phase: firstly, during the measurement, the resting period was analyzed by observing whether the patient kept quiet, did not talk, and moved as little as possible during the relaxation time. Secondly, as during the baseline phase, no task-related activation in the averaged CBFV is expected; after the measurement, the CBFV curves of the baseline phase were visually inspected regarding unsuspected increases in CBFV or differences between the left and right MCAs.
- (3) Adequate increase in CBFV during the activation phase: after the measurement, the CBFV curves of the activation phase were

visually inspected regarding an increase in the averaged CBFV compared with the individual baseline phase, as its absence or a fast decrease in CBFV during the word generation period would indicate a reduced task-related activation.

If one or more of the qualitative categories were not met, i.e., patients did not show instruction-compliant task performance, a sufficient quality of the baseline phase, or an adequate increase in CBFV during the activation phase, language lateralization results were subsequently redefined as “not lateralizing”, regardless of the previous standard fTCD analysis.

2.5. Statistics

Statistical analyses were performed using IBM SPSS Statistics 22 (IBM Corporation, USA).

Group differences in HLL were tested using Wilcoxon-tests, and subgroups were compared with the help of independent samples *t*-tests and chi-square-tests. Pearson product moment correlation coefficients were calculated firstly to correlate the LIs of the word generation and speech periods and secondly to analyze the retest reliability of fTCD measurements. Fisher *r*-to-*z* transformation was used to compare the retest reliability coefficients between the standard and the two-step analyses of fTCD data.

3. Results

3.1. Standard analysis

When only regarding the well-established *statistical* examination of fTCD data, left HLL was found in 45% of all patients ($n = 33$) during the first measurement (T1). Right HLL was seen in 55% of patients. During the second measurement (T2), left HLL was found in 55%, right HLL was found in 24%, and no lateralization was found in 21% of patients (see Fig. 2a).

Hemispheric language lateralization was consistent across both measurements in 55% of patients (group of “nonswitchers”). However, HLL of 21% of patients switched from lateralized at T1 to not lateralized at T2, and surprisingly, in the remaining 24% of patients, HLL even switched sides across the two measurements (one from left to right,

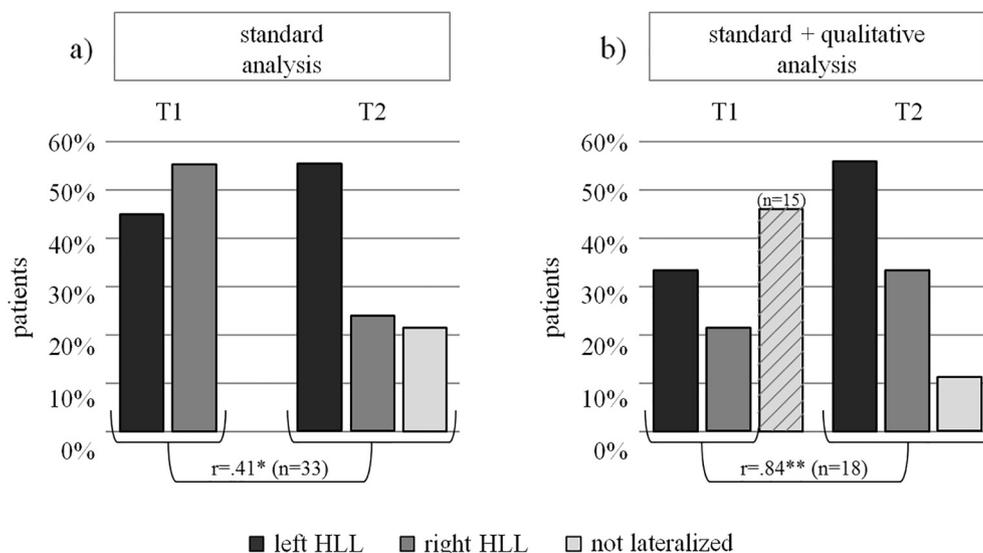


Fig. 2. Distribution of hemispheric language lateralization (HLL) a) after standard analysis and b) after standard + qualitative analysis. Retest reliability for each data analysis method was computed for patients with clear lateralization at T1 only, resulting in the exclusion of 15 patients after the qualitative analysis (crossed out bar). * $p < 0.05$; ** $p < 0.01$.

seven from right to left) (group of “switchers”). This led to a rather poor retest reliability of $r = 0.41$ (95% confidence interval = 0.08–0.66; $p = 0.017$).

No significant differences between “nonswitchers” and “switchers” regarding sex, age at onset, or side of epilepsy could be revealed (all p -values > 0.05). As motor-associated artefacts may occur during loud word generation (speech period), this period is not commonly used for language lateralization itself. However, because of its establishment as compliance control, in this special case the speech period was examined, and only “nonswitchers” exhibited a correlation among silent and loud word generation ($r = 0.65$, $p = 0.007$). As “switchers” did not show such a correlation, doubts regarding appropriate task performance in that subgroup were raised.

3.2. Standard + qualitative analysis

All patients ($n = 33$) underwent an additional qualitative evaluation of fTCD data (see Table 1). Thereby, the unusually high proportion of patients with statistically defined right HLL during T1 decreased (from 55% to 21%), and in total 46% of initially lateralized patients ($n = 15$) were now qualitatively reevaluated as showing no clear lateralization at T1 (see Fig. 2b). Thus, as the focus of this study was the investigation of the reproducibility of HLL, these fifteen patients were excluded from retest reliability analysis, leading to the final sample ($n = 18$) reported hereunder.

Of the final sample of patients who were statistically and qualitatively lateralized at T1, 89% showed consistent HLL across both

Table 1
Overview of all 33 patients after standard and qualitative analyses of the first fTCD measurement.

Patient	Sex	Age	Standard analysis		Qualitative analysis		
			LI	SEM	Category 1 met	Category 2 met	Category 3 met
1 ^a	M	33	-10.6	1.41	No	No	Yes
2 ^{a,b}	F	28	-6.75	0.85	Yes	No	No
3	F	28	-6.37	0.83	Yes	Yes	Yes
4	M	33	-5.64	0.63	Yes	Yes	Yes
5 ^{a,b}	F	46	-5.04	0.94	Yes	No	No
6 ^{a,b}	M	56	-3.88	0.59	Yes	No	No
7	M	48	-3.67	0.68	Yes	Yes	Yes
8 ^{a,b}	M	45	-2.80	0.35	No	No	Yes
9	F	45	-2.70	0.76	Yes	Yes	Yes
10 ^{a,b}	F	26	-2.68	0.61	No	No	Yes
11	F	34	-2.66	0.71	Yes	Yes	Yes
12 ^{a,b}	F	60	-2.47	1.08	Yes	No	No
13 ^a	F	33	-2.18	0.64	Yes	No	No
14 ^{a,b}	F	29	-1.87	0.68	No	Yes	No
15	M	34	-1.70	0.61	Yes	Yes	Yes
16 ^a	F	38	-1.55	0.34	No	Yes	No
17	F	45	-1.26	0.44	Yes	Yes	Yes
18 ^a	F	22	-1.13	0.47	No	Yes	No
19 ^a	M	38	1.06	0.42	No	No	No
20 ^a	M	23	1.32	0.49	No	Yes	No
21	M	31	1.68	0.64	Yes	Yes	Yes
22	M	31	1.74	0.77	Yes	Yes	Yes
23	M	27	1.84	0.59	Yes	Yes	Yes
24	M	51	2.01	0.87	Yes	Yes	Yes
25	F	36	2.28	0.52	Yes	Yes	Yes
26	F	17	2.57	0.59	Yes	Yes	Yes
27 ^a	F	17	2.85	1.20	Yes	No	Yes
28	F	35	2.91	0.78	Yes	Yes	Yes
29 ^{a,b}	M	27	3.98	0.73	Yes	No	No
30	M	33	4.55	0.61	Yes	Yes	Yes
31	M	57	5.23	0.54	Yes	Yes	Yes
32	M	41	6.53	1.50	Yes	Yes	Yes
33	F	37	6.59	0.74	Yes	Yes	Yes

F = female, M = male.

^a These patients ($n = 15$) were subsequently redefined as “not lateralizing” because one or more of the qualitative categories (Category 1 = instruction-compliant task performance, Category 2 = sufficient quality of the baseline phase, Category 3 = adequate increase in CBFV during the activation phase) were not met.

^b Group of “switchers” = HLL switched sides across the two measurements.

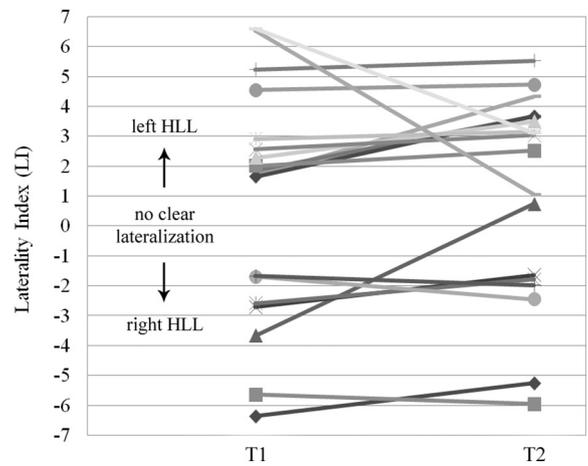


Fig. 3. Consistency of LIs between the first (T1) and second (T2) measurements of the final sample ($n = 18$) of patients who were statistically and qualitatively lateralized at T1.

measurements. This led to a highly significant retest reliability of $r = 0.84$ (95% confidence interval = 0.61–0.94; $p < 0.000$; see Figs. 2b and 3).

The reproducibility of HLL was significantly improved ($z = -2.48$, $p = 0.007$) by the use of the two-step (standard + qualitative) data analysis, as also highlighted by the higher consistency across T1 and T2 regarding the ratio of the left and right HLLs after the two-step data analysis compared with the standard data analysis (see Fig. 2). Interestingly, at T2, a significantly higher consistency regarding HLL ($p = 0.005$) between the standard and two-step data analyses could be obtained, presumably attributable to a lower amount of qualitative redefinitions due to a general significant improvement in the quality of T2 compared with T1 ($p = 0.003$).

4. Discussion

In contrast to previous studies assessing the reproducibility of fTCD determining HLL [13–15], in our sample, a rather poor retest reliability between the first and second measurements could be obtained, when only using the well-established statistical examination of fTCD data. Furthermore, even given the known relatively high number of patients with TLE with atypical HLL [7,8], in our sample, the proportion of statistically defined patients with right HLL during the first measurement appeared to be extraordinarily high.

As technical aspects, such as fTCD procedure and probe positioning, were kept consistent between measurements, especially *psychological and situational factors* [16] might be an explanation for our unusual findings. In this study, fTCD measurements were performed as part of presurgical work-up and thus were often implemented following a long neuropsychological assessment taking 2–3 h. Together with other factors, such as withdrawal of antiepileptic medications and the recent increase in seizure frequency during video-electroencephalography (video-EEG) monitoring, these aspects presumably impaired the patients' ability to concentrate on the task and probably influenced their motivation during fTCD measurements.

This potential source of variability not only limits the comparability to other reproducibility studies, as *there* often healthy participants in controlled conditions were examined. More importantly, this observation encouraged us to search for additional means of fTCD data evaluation, in order to specifically control for these psychological and situational aspects, which seem to be of extraordinary importance in the clinical setting. Underlined by our findings, we conclude that the standard *statistical* examination of fTCD data should be complemented by an additional *qualitative* evaluation of the measurements.

Firstly, this supplementary step of data analysis helps to detect and exclude fTCD measurements of insufficient quality and by that provides

a more conservative and even more secure means of data evaluation. Since fTCD plays a significant role in the routine presurgical work-up of patients with epilepsy regarding the planning of neurosurgical interventions and the prediction of possible cognitive risks, this extra security is especially desirable in the clinical setting.

Secondly, this additional qualitative assessment supports the identification of possible opportunities to intervene before or even *during* the fTCD examination in order to achieve better quality of measurements in the first place. For example, even in the limited time of a video-EEG monitoring, it might be better not to perform fTCD following a long neuropsychological assessment. Moreover, the importance of instruction-compliant task performance could be emphasized even more, and the patients' motivation to give their best effort might be checked more often during measurements. Furthermore, the implementation of a short break in the middle of the 20 epochs might be considered in order to increase the patients' ability to concentrate on the task. The latter should be part of further investigations.

Until now, in clinical routine, qualitative aspects of measurements have already been included in the evaluation of fTCD examinations but so far mostly in a very unsystematic way and usually by means of subjective "clinical impressions" of clinicians performing fTCD. In order to alert other medical professionals to look beyond solely statistical analyses and rely on a combination of quantitative and qualitative information, this "clinical impression" needed to be standardized and underpinned by empirical data. In this study, we aimed to a) formalize this qualitative approach by defining clear criteria (three qualitative categories) and strictly applying them to data analysis and b) investigate how results change when using this additional qualitative analysis in order to support this approach by empirical findings.

We hope our findings contribute to the understanding that quantitative and qualitative analyses not inevitably have to be considered as polar opposites and that the integration of both analyses can improve the evaluation of fTCD data in clinical routine.

5. Conclusion

Our results underline that especially in clinical settings, psychological and situational factors seem to influence the reproducibility of fTCD determining HLL. While this has already been the "clinical impression" of many clinicians when examining cognitive data, the approach of using qualitative information for fTCD data analysis has been standardized and proven empirically for the first time in this study. Accordingly, we highly recommend the complementation of the standard *statistical* examination of fTCD data by an additional *qualitative* evaluation and want to emphasize the importance of critically scrutinizing the circumstances in which fTCD measurements are performed in clinical routine in order to achieve better quality of measurements in general.

As confirmed by our findings, the implementation of this two-step data analysis significantly improves the method's retest reliability, indicating that repetitions of fTCD measurements are generally reliable and thus useful for verifying or improving the results of the first measurement. Because of its direct implications for the further evaluation of neurosurgical interventions, the extra security of the two-step data analysis is particularly desirable.

Conflict of interest

The authors declare that they have no competing interest.

Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. We thank Rebecca Scharf, Anja Haag, and Silke Leonhard for their contribution to data collection.

References

- [1] Duschek S, Schandry R. Functional transcranial Doppler sonography as a tool in psychophysiological research. *Psychophysiology* 2003;40(3):436–54.
- [2] Bishop DV, Badcock NA, Holt G. Assessment of cerebral lateralization in children using functional transcranial Doppler ultrasound (fTCD). *J Vis Exp* 2010;43.
- [3] Haag A, Moeller N, Knake S, Hermsen A, Oertel WH, Rosenow F, et al. Language lateralization in children using functional transcranial Doppler sonography. *Dev Med Child Neurol* 2010;52(4):331–6.
- [4] Lohmann H, Dräger B, Müller-Ehrenberg S, Deppe M, Knecht S. Language lateralization in young children assessed by functional transcranial Doppler sonography. *NeuroImage* 2005;24(3):780–90.
- [5] Wijnhoud AD, Franckena M, Van Der Lugt A, Koudstaal PJ. Inadequate acoustical temporal bone window in patients with a transient ischemic attack or minor stroke: role of skull thickness and bone density. *Ultrasound Med Biol* 2008;34(6):923–9.
- [6] House PM, Brückner KE, Lohmann HH. Presurgical functional transcranial Doppler sonography (fTCD) with intravenous echo enhancing agent SonoVue enables determination of language lateralization in epilepsy patients with poor temporal bone windows. *Epilepsia* 2011;52(3):636–9.
- [7] Knake S, Haag A, Pilgramm G, Dittmer C, Reis J, Aßmann H, et al. Language dominance in mesial temporal lobe epilepsy: a functional transcranial Doppler sonography study of brain plasticity. *Epilepsy Behav* 2006;9(2):345–8.
- [8] Springer JA, Binder JR, Hammeke TA, Swanson SJ, Frost JA, Bellgowan PS, et al. Language dominance in neurologically normal and epilepsy subjects: a functional MRI study. *Brain* 1999;122(11):2033–46.
- [9] Helmstaedter C. Neuropsychological aspects of epilepsy surgery. *Epilepsy Behav* 2004;5:45–55.
- [10] Stroup E, Langfitt J, Berg M, McDermott M, Pilcher WMDP, Como P. Predicting verbal memory decline following anterior temporal lobectomy (ATL). *Neurology* 2003;60(8):1266–73.
- [11] Deppe M, Knecht S, Papke K, Lohmann H, Fleischer H, Heindel W, et al. Assessment of hemispheric language lateralization: a comparison between fMRI and fTCD. *J Cereb Blood Flow Metab* 2000;20(2):263–8.
- [12] Knake S, Haag A, Hamer HM, Dittmer C, Bien S, Oertel WH, et al. Language lateralization in patients with temporal lobe epilepsy: a comparison of functional transcranial Doppler sonography and the Wada test. *NeuroImage* 2003;19(3):1228–32.
- [13] Knecht S, Deppe M, Ebner A, Henningsen H, Huber T, Jokeit H, et al. Noninvasive determination of language lateralization by functional transcranial Doppler sonography: a comparison with the Wada test. *Stroke* 1998;29(1):82–6.
- [14] Knecht S, Deppe M, Ringelstein EB, Wirtz M, Lohmann H, Dräger B, et al. Reproducibility of functional transcranial Doppler sonography in determining hemispheric language lateralization. *Stroke* 1998;29(6):1155–9.
- [15] Stroobant N, Vingerhoets G. Test–retest reliability of functional transcranial Doppler ultrasonography. *Ultrasound Med Biol* 2001;27(4):509–14.
- [16] Stroobant N, Vingerhoets G. Transcranial Doppler ultrasonography monitoring of cerebral hemodynamics during performance of cognitive tasks: a review. *Neuropsychol Rev* 2000;10(4):213–31.
- [17] Helmstaedter C, Jockwitz C, Witt JA. Menstrual cycle corrupts reliable and valid assessment of language dominance: consequences for presurgical evaluation of patients with epilepsy. *Seizure* 2015;28:26–31.
- [18] Rosenow F, Bast T, Czech T, Feucht M, Hans VH, Helmstaedter C, et al. Revised version of quality guidelines for presurgical epilepsy evaluation and surgical epilepsy therapy issued by the Austrian, German, and Swiss working group on presurgical epilepsy diagnosis and operative epilepsy treatment. *Epilepsia* 2016;57(8):1215–20.
- [19] Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 1971;9(1):97–113.
- [20] Knecht S, Henningsen H, Deppe M, Huber T. Successive activation of both hemispheres during cued word generation. *Neuroreport: An International Journal for the Rapid Communication of Research in Neuroscience* 1996;7(3):820–4.
- [21] Deppe M, Knecht S, Henningsen H, Ringelstein EB. AVERAGE: a Windows® program for automated analysis of event related cerebral blood flow. *J Neurosci Methods* 1997;75(2):147–54.
- [22] Deppe M, Ringelstein EB, Knecht S. The investigation of functional brain lateralization by transcranial Doppler sonography. *NeuroImage* 2004;21(3):1124–46.
- [23] Vogt VL, Äikiä M, Del Barrio A, Boon P, Borbély C, Bran E, et al. Current standards of neuropsychological assessment in epilepsy surgery centers across Europe. *Epilepsia* 2017;58(3):343–55.