

Impact of white-matter mask selection on DTI histogram-based metrics as potential biomarkers in cerebral small vessel disease

Ana R Fouto (✉ anafouto@tecnico.ulisboa.pt)

Institute for Systems and Robotics - Lisboa and Department of Bioengineering, Instituto Superior Técnico, Universidade de Lisboa <https://orcid.org/0000-0002-0362-9571>

Rita G. Nunes

Institute for Systems and Robotics - Lisboa and Department of Bioengineering, Instituto Superior Técnico, Universidade de Lisboa

Joana Pinto

Institute for Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford, UK

Luísa Alves

Neurology Department, Hospital Egas Moniz, Centro Hospitalar de Lisboa Ocidental, Lisbon, Portugal

Sofia Calado

Neurology Department, Hospital Egas Moniz, Centro Hospitalar de Lisboa Ocidental, Lisbon, Portugal

Carina Gonçalves

Neurology Department, Hospital Egas Moniz, Centro Hospitalar de Lisboa Ocidental, Lisbon, Portugal

Margarida Rebolo

Faculdade de Medicina, Universidade de Lisboa, Lisbon, Portugal

Miguel Viana-Baptista

Neurology Department, Hospital Egas Moniz, Centro Hospitalar de Lisboa Ocidental, Lisbon, Portugal

Pedro Vilela

Imaging Department, Hospital da Luz, Lisbon, Portugal

Patrícia Figueiredo

Institute for Systems and Robotics - Lisboa and Department of Bioengineering, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

Research Article

Keywords: cerebral small vessel disease, cognitive dysfunction, magnetic resonance imaging, diffusion-tensor imaging, white-matter.

Posted Date: July 19th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-703117/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Purpose Histogram-based metrics extracted from diffusion-tensor imaging (DTI) have been suggested as potential biomarkers for cerebral small vessel disease (SVD), but methods and results have varied across studies. This work aims to assess the impact of mask selection for extracting histogram-based metrics of fractional anisotropy (FA) and mean diffusivity (MD) on their sensitivity as SVD biomarkers.

Methods DTI data were collected from 17 SVD patients and 12 healthy controls. For each participant, FA and MD maps were estimated; from these, histograms were computed on two alternative whole-brain white-matter masks: normal-appearing white-matter (NAWM) and mean FA tract skeleton (TBSS). Histogram-based metrics (median, peak height, peak width, peak value) were extracted from the FA and MD maps. These were compared between patients and controls, and correlated with the patients' cognitive scores (executive function and processing speed).

Results White matter mask selection significantly impacted FA and MD histogram metrics and affected their ability to discriminate between groups. Moreover, we observed that the mask can influence the correlations with cognitive measures. Nevertheless, the MD peak height and MD peak width metrics remained significantly correlated with executive function, regardless of the mask.

Conclusion Our results corroborate previous reports and further support the value of DTI histogram-based metrics as SVD biomarkers. However, they also highlight the importance of the processing methodology, in particular the choice of white matter mask, as hence the urgent need to mitigate the lack of standardized MRI data-processing pipelines.

1. Introduction

Cerebral small vessel disease (SVD) is associated with progressive cognitive impairment and is one of the leading causes of dementia among the elderly. SVD is associated with the pathological mechanisms affecting the small vessels of the brain (i.e. 5 μ m-2mm in diameter) [1]. Several magnetic resonance imaging (MRI) modalities have been used to identify neuroimaging features that may provide potential disease biomarkers, useful for early diagnosis, monitoring disease progression, and evaluation of new therapies [2]. Biomarkers based on conventional structural MRI have been extensively studied [3–6]. Unfortunately, most of these features are only detectable at advanced stages of the disease and only poorly correlate with cognitive performance [2], which is typically dominated by deficits in executive function and processing speed [7].

Consequently, other MRI techniques targeting earlier pathological changes have been explored. Remarkably, diffusion-weighted imaging (DWI) methods have been found to be sensitive to white matter (WM) microstructural abnormalities in SVD patients, not only in areas already affected by lesions, but also in the normal-appearing white matter (NAWM) [8]. In particular, fractional anisotropy (FA) and mean diffusivity (MD) parameters derived from diffusion-tensor imaging (DTI) have been most commonly used. Furthermore, histogram analysis has been preferably employed in SVD studies to better describe

the diffuse nature of white matter pathology in this disease [9]. Histogram analysis enables the assessment of microstructural abnormalities distributed across the whole-brain, in contrast to region-of-interest (ROI) analyses which provide measures of localized changes in a given area or structure. Moreover, a previous study has shown that histogram measures are more reproducible [10]. However, the choice of the WM mask used to compute the histograms has differed substantially between studies. The majority of prior research has employed WM/NAWM masks or tract-based spatial statistics (TBSS) WM skeleton masks. While different results are expected when these distinct masks are used, the impact of these differences on the detection of pathology is not easily predictable. Most critically, this methodological variability may at least partially explain the inconsistencies between studies regarding the best predictor of cognitive impairment [7], given that the choice of WM mask can influence the correlations with cognitive measures.

For example, previous literature suggested the MD peak height computed over the WM mask as a sensitive biomarker to predict progressive decline in executive function in a sample of sporadic SVD (sSVD) patients [11, 12]. In contrast, Nitkunan et al. considered a mask including all visible lesions as well as voxels for which the probability of being grey matter (GM) or WM was higher than 0.5, and reported that the FA peak height was more strongly associated with executive function than other FA or MD histogram-based metrics [13]. Moreover, Croall et al. reported the strongest predictor of executive dysfunction to be the FA median when testing both WM and NAWM masks [14]. On the other hand, Baykara et al. [10] reported that the strongest predictor of processing speed was the MD peak width evaluated on a TBSS skeleton built to compare across different SVD patient groups [15]. Conversely, Lawrence et al. demonstrated that the MD peak height obtained over the NAWM voxels was strongly associated with processing speed impairment [16]. Nonetheless, in recent work, the MD peak width has been utilized as a reference measure to validate the predictive power of novel biomarkers due to its well-established value in detecting microscopic tissue changes in a broad range of pathological WM conditions [10, 17–20]; as a result, retrospective validation in existing DWI datasets may help to establish a consistent processing pipeline for future longitudinal studies

Our study aimed to assess the impact of using different masks for computing DTI histogram-based metrics on their sensitivity as SVD biomarkers. For this purpose: 1) we extracted different DTI histogram-based metrics using two distinct masks: NAWM *versus* TBSS; 2) we evaluated their sensitivity to discriminate between patients and healthy controls; and 3) we performed a correlation analysis between a selection of DTI metrics that have been most frequently considered in SVD studies (i.e, FA median, FA peak height, MD peak height, and MD peak width) and the cognitive domains more typically impaired in SVD patients (i.e. processing speed and executive function).

2. Methods

2.1. Study Population

The study was approved by the local Ethics Committee, and all subjects gave written informed consent in accordance with the Declaration of Helsinki.

A total of N = 17 patients were recruited including two different groups of SVD patients: Cerebral Autosomal Dominant Arteriopathy With Subcortical Infarcts and Leukoencephalopathy (CADASIL) patients (N = 6) and sporadic SVD (sSVD) patients (N = 11). A healthy control (HC) group was also recruited consisting of N = 12 age-matched healthy volunteers, with no known relevant medical history, namely SVD or dementia. The patients' inclusion criteria were: 1) independence in daily activities, as assessed by the Instrumental Activities of Daily Living (IADL) scale; absence of hemodynamically significant large vessel disease, as evaluated by Doppler ultrasound (carotid and vertebral); 2) for sSVD group: evidence of SVD in MRI (deep WMH lesions, with no other plausible explanation, with moderate and severe degrees, II/III or III/III, according to the Fazekas scale [21]; 3) for CADASIL group: symptomatic patients with evidence of WMH lesions and molecular diagnosis confirmed by mutation on the NOTCH3 gene. Exclusion criteria were: 1) contraindications for MRI; 2) for the patients: evidence of WMH lesions from other known aetiology based on clinical and laboratory findings; presence of concurrent chronic incapacitating disorders; stroke in the past three months; illiteracy; and compromised visual acuity. All patients were examined by an experienced neurologist in order to exclude other neurological or neuropsychiatric diseases and to perform neurological examinations. Demographic data and cerebrovascular risk factors (including history of previous stroke, hypertension, blood pressure, hypercholesterolaemia, diabetes and smoking) were recorded. The cohort demographics and clinical data are presented in Table S1 (supplementary material). At another session, cognitive functions were evaluated by an experienced neuropsychologist. Finally, all participants underwent MRI scanning, according to a standardized protocol.

2.2. Neuropsychological Evaluation

To evaluate cognitive functions, patients underwent a comprehensive battery of neuropsychological tests as close to the MRI scanning as possible (within 15 days). One patient was excluded from the statistical analysis due to unavailability to perform the neuropsychological evaluation. Global functioning was assessed by the Montreal Cognitive Assessment (MoCA) Test, and specific cognitive domains of interest (processing speed and executive function) were evaluated by Stroop test and Trail Making Test (TMT). We considered either specific scores of the appropriate cognitive tests or composite scores, as shown in Table S2 (supplementary material). For each test, we generated a cognitive z-score; we converted the Stroop Test percentile scores into z-scores considering the available normative data for the studied population [22]. In the case of the TMT, we used a publicly available database to calculate the z-scores directly [23]. Then, composite scores were generated based on previous reports to represent: (1) *executive function*: average of the z-scores of Stroop Test interference and TMT subtracting part A from part B (TMTB-A) [24, 25]; and (2) *processing speed*: average of the z-scores of TMT part A and part B [10].

2.3. MRI Acquisition

Whole-brain images were acquired on a 3T Siemens Verio scanner using a 12-channel radio-frequency receive head coil including: (i) DWI-Echo-Planar Imaging (EPI), with TR/TE = 4800/107 ms, 25 contiguous

slices with $1.7 \times 1.7 \times 5.2 \text{ mm}^3$ resolution, 3 repetitions of diffusion-sensitizing gradients along 20 directions with $b = 1000 \text{ s/mm}^2$ and $b = 0 \text{ s/mm}^2$ image, acquisition time (TA) = 5:18 min.; (ii) Magnetization-Prepared Rapid Gradient-Echo (MPRAGE)- T_1 -weighted imaging (T1WI), with TR/TE/TI = 2250/2.26/900 ms, flip angle = 9° , and 144 contiguous slices with 1 mm^3 isotropic resolution, TA = 9:56 min.; (iii) Fluid-attenuated inversion recovery (FLAIR)- T_2 -weighted imaging (T2WI), with TR/TE/TI = 8500/97/2500ms, and 45 contiguous slices with $0.9 \times 0.7 \times 3 \text{ mm}^3$ resolution, TA = 4:17 min.; and (iv) Susceptibility-weighted imaging (SWI) using a gradient-echo sequence, with TR/TE = 28/20; flip angle = 15° , 96 slices with $0.7 \times 0.6 \times 1.4 \text{ mm}^3$ resolution and TA = 4:49min.

2.4. MRI Analysis

The main steps of the processing pipeline (see details below) followed in this work for the extraction of both conventional structural imaging metrics and DTI histogram-based metrics are depicted in Fig. 1.

2.4.1. Conventional structural imaging

For comparison with DTI derived metrics, the following structural measures were obtained from conventional structural imaging:

Normalised brain volume (NBV), which is expected to provide a measure of brain atrophy, was automatically estimated from MPRAGE-T1WI with FSL's SIENAX [26], using the white-matter hyperintensities (WMH) lesion masks as an input (details below) to minimize GM misclassification. The total brain volume was considered as the sum of GM and WM volumes and was normalised by the intracranial volume to obtain NBV. Additionally, brain tissue masks (GM, WM and cerebrospinal fluid (CSF)) were obtained by segmentation of the MPRAGE image with FSL's FAST [27]. Segmentation of subcortical structures was also performed with FSL's FIRST [28], and their spurious contributions were removed to obtain a corrected WM mask.

WMH lesions, cerebral microbleeds and lacunes were identified by an experienced neuroradiologist. WMH lesions were manually segmented considering regions of increased signal on FLAIR-T2WI (using *fs/view* from version 5.0.9 of FSL). **Normalised WMH lesion load (NWMHLL)** was then estimated as the percentage of WMH lesion volume relative to the whole-brain NBV. We transformed the lesion masks to MPRAGE space by performing linear registrations between the two types of images using FSL's FLIRT [29, 30]; linear interpolation was used in the transformation; in order to obtain a binary mask, the registered image was thresholded at 0.5 and binarized. The **number of cerebral microbleeds (nCMB)** was obtained according to the Microbleed Anatomical Rating Scale (MARS) criterion [31] based on SWI. The **number of lacunes (nLac)** was obtained by quantifying the areas of tissue loss with cavitation on MPRAGE-T1WI with 0.3–1.5 cm of diameter [32].

2.4.2. Diffusion-tensor imaging

Firstly, DWI data was corrected for eddy-current distortions and motion with FSL's *eddy* tool [33] using the outliers replacement option (*repol* [34]). To obtain FA and MD maps from the corrected DWI data, we

performed tensor-fitting using FSL's *dtifit* tool [35, 36]. Then, we evaluated two alternative WM masks for deriving the histogram-based measures of FA and MD: TBSS *versus* NAWM.

To generate the TBSS mask, we applied the TBSS pipeline from FSL [15] to both the FA and MD maps by first performing erosion of FA maps to remove outliers contributions from tensor fitting. Then, FA maps were non-linearly aligned to an FA template (FMRIB58_FA_1mm) in the MNI space [37] using linear interpolation (ANTs tools, <http://stnava.github.io/ANTs/>). Subsequently, from the FA maps, we derived a mask representing the most common WM tracts across the subjects group mean WM skeleton, which was thresholded at 0.3. This threshold was chosen after evaluating two potential values: 0.2 (FSL's default) and 0.3 [10, 38]. MD values were also projected onto the previously derived mean WM skeleton by applying an adapted version of the FSL's *tbss_non_FA* script (as previously indicated, non-linear registration was also performed using ANTs tools.).

To define the NAWM mask, we subtracted the WMH mask registered into the MPRAGE space from the total corrected WM mask. Then, the NAWM mask was transformed into the DWI native space (where the FA and MD maps are originally computed). For this purpose, an affine transformation between the DWI ($b = 0$ image) and MPRAGE space was estimated using FSL's FLIRT; and subsequently used to transform the NAWM mask (*MPRAGE space*), using linear interpolation.

Normalized histograms (divided by the total number of voxels and bin width) of both FA and MD maps were computed in R (www.r-project.org). FA range was defined between 0 and 1. For the MD range, we considered as the upper limit the maximum value of the 99th percentile across subjects, intending to exclude CSF voxel contributions (using this criterion, the MD upper limit was $0.0025 \text{ mm}^2\text{s}^{-1}$). The optimal number of bins was estimated using the Freedman-Diaconis rule [39], for each mask. To enable comparisons with the literature, the metrics more frequently extracted from both the FA and MD histograms were selected based on previous reports published between 2008–2018 [8]: median, peak height, peak width (difference between the 5th and 95th *percentiles*) and peak value.

2.5. Statistical Analysis

Statistical analysis was performed in R (www.r-project.org). All metrics were first transformed to z-scores. Clinical data and neuroimaging features extracted from conventional structural MRI were compared between groups using independent T-tests, Mann-Whitney tests, or chi-square tests, as appropriate. For the analysis of the patients' cognitive profile, we performed a Wilcoxon signed-rank test between each neuropsychological score and the reference performance expected for the healthy population (z-score = 0). The impact of each mask on the ability of the FA and MD histogram-based metrics to discriminate between patients and controls was assessed by performing a 2-way repeated-measures ANOVA with factors: between-subjects factor Group (SVD and HC); and within-subjects factor Mask (TBSS and NAWM). When interactions between Group and Method and/or Mask were identified as being statistically significant, a post-hoc analysis was performed. Finally, a partial Spearman correlation analysis was employed to estimate the power to predict either executive function or processing speed (corrected for the effect of age and sex), by using each of the four selected metrics (FA median, FA peak height, MD peak

height, and MD peak width), generated using each mask. Bonferroni correction was used to adjust the p-values in order to correct for multiple comparisons, and significant effects were considered for corrected $p < 0.05$.

3. Results

3.1. Neuropsychological evaluation

The cognitive profile across the group of patients is presented in Fig. 2. Performance was significantly impaired only for TMTA and TMTB cognitive scores. According to the respective composite scores, only processing speed was significantly impaired in this patient group, compared to the expected performance for a healthy population ($p < 0.01$). Nevertheless, executive function performance also showed considerable variability across patients.

3.2. Conventional structural imaging metrics

Figure 2 shows an illustrative example of WMH lesions, cerebral microbleeds and lacunes. Table 1 summarises the structural imaging metrics extracted from MPRAGE-T1WI, FLAIR-T2WI and SWI. NBV was marginally higher in SVD patients compared to healthy controls ($p = 0.034$). However, no significant differences between patient groups were found ($p = 0.098$). As expected, patients showed a higher NWMHLL compared to controls ($p < 0.001$). Regarding nCMB and nLac, and comparing with previous reports, this cohort showed a lower lesion burden. Specifically, nCMB presented an average value of 1.8 [range:0–20] and nLac of 0.5 [range:0–4]. For instance, Croall et al. reported an average of 4.41 ± 4.73 lacunes (N = 102) [14]. Similarly, Zeestraten et al. reported an average of 4.2 lacunes [range:0–27] and 5.5 microbleeds [range:0-144] (N = 99) [12]. For this reason, we did not include these measures in further analyses.

Table 1

– **Structural imaging metrics extracted from MPRAGE-T1WI, FLAIR-T2WI and SWI.** Significant differences between groups are shown in bold ($p < 0.05$). Numbers represent mean \pm standard deviation (or range). NBV = Normalised brain volume; NWMHLL = Normalised white matter hyperintensities lesion load; nCMB = Number of cerebral microbleeds; nLac = Number of lacunes.

Metrics	SVD (N = 17)		Healthy controls (N = 12)	Group comparison p-value
	CADASIL (N = 6)	sSVD (N = 11)		
NBV (cm ³)	1506 \pm 72		1451 \pm 67	p = 0.034
	1551 \pm 50	1481 \pm 72		p = 0.098*
NWMHLL (%)	2.8 \pm 2.3		0.05 \pm 0.07	p < 0.001
	4.3 \pm 3.2	2.0 \pm 1.1		p = 0.149*
nCMB [range]	1.8 [0–20]		-	-
	3.5 [0–20]	0.9 [0–4]		
nLac [range]	0.5 [0–4]		-	-
	1.5 [0–4]	0		
*Comparison between CADASIL and sSVD groups.				

3.3. DTI histogram metrics

TBSS and NAWM masks obtained are exemplified in Fig. 4 for one representative subject. The FA and MD histograms extracted from each type of mask, together with the respective histogram-based metrics (median, peak height, peak width, and peak value), are illustrated in Fig. 5. Histograms of all subjects are depicted in Figures S1 and S2 (supplementary material).

Boxplots showing the distributions across groups of all FA and MD histogram metrics, obtained using each mask are displayed in Fig. 6. Results suggest that changes in FA histogram are generally characterized by reduced FA in patients compared to healthy controls as evidenced by lower median and peak values. On the contrary, MD histograms generally showed higher median and peak values in patients. Boxplots showing the distributions across groups of all FA and MD histogram metrics, obtained using each mask, are displayed in Fig. 6. Results suggest that changes in FA histograms are generally characterized by reduced FA in patients compared to healthy controls as evidenced by lower median and peak values. On the contrary, MD histograms generally showed higher median and peak values in patients.

A two-way mixed factor ANOVA was used to examine the effect of the factor Mask, as well as the subject Group, on the DTI histogram metrics. Table 2 summarizes the major results of this analysis. Overall, there was a significant main effect of Mask for all FA metrics ($p < 0.001$) and MD peak height ($p < 0.001$). The

post-hoc tests showed that, irrespective of group, the two masks yielded significantly different results; furthermore, as seen from the boxplots (Fig. 6), FA histogram-derived metrics are significantly higher for TBSS relative to NAWM. Significant main effects of Group were found only for the following metrics: FA median ($p = 0.008$); FA peak height ($p = 0.041$); FA peak value ($p = 0.004$); MD peak height ($p = 0.048$); and MD peak width ($p = 0.024$). Significant interactions were found between Mask and Group for: FA peak height ($p = 0.027$); MD median ($p = 0.035$); and MD peak width ($p = 0.047$). Post-hoc analysis revealed that: (1) when using either mask, we observed no significant differences in FA peak height and MD median between groups; however, significant differences were found for MD peak width with TBSS ($p = 0.037$) but not NAWM ($p = 0.690$).

Table 2
 – Statistical analysis of FA and MD histogram-based metrics: 2-way mixed factor repeated measures ANOVA with within-subject factor Mask and between-subject factor Group. Bonferroni-corrected p values are presented, and significant effects are indicated in bold ($p < 0.05$). FA = Fractional anisotropy; MD = Mean diffusivity.

Map	Metrics	Mask	Mask*Group	Group
FA	Median	< 0.001	0.874	0.008
	Peak Height	< 0.001	0.027	0.041
	Peak Width	< 0.001	0.397	0.844
	Peak Value	< 0.001	0.713	0.004
MD	Median	0.528	0.035	0.067
	Peak Height	< 0.001	0.183	0.048
	Peak Width	0.175	0.047	0.024
	Peak Value	0.801	0.157	0.135

3.4. Correlation of DWI histogram metrics with cognitive scores

Figure 7 presents the correlation coefficients obtained between each selected DTI metric (FA median, FA peak height, MD peak height, and MD peak width) and performance in each cognitive domain (executive function and processing speed). Illustrative examples of scatter plots showing the relationship between two DTI metrics (MD peak width and height) and cognitive performance are depicted in Figures S3 and S4 (supplementary material). For both masks, the MD peak width displayed a significant negative correlation with executive function (TBSS: $p = 0.003$; NAWM: $p = 0.007$). In addition, the MD peak height also showed a statistically significant positive correlation with executive function using both masks (TBSS: $p = 0.030$; NAWM: $p = 0.040$). Overall, there were no significant associations between FA peak

height and the cognitive variables. Finally, FA median showed a significant positive correlation with executive function using NAWM mask ($p = 0.020$)

4. Discussion

This study evaluated the effect of using different whole-brain WM masks on the potential of DTI histogram-based metrics as biomarkers in a group of patients with cerebral SVD. We found that the choice of mask significantly impacted the FA and MD histogram metrics considered. Most critically, in some cases, they affected the ability of the metrics to discriminate between patients and controls. Moreover, we found that the mask also affected the relationship between DTI metrics and cognitive performance in tests of executive function and processing speed.

4.1. Impact of mask selection on DTI metrics

Our study shows that the WM mask used to generate the FA and MD histograms has a significant impact on the values of the respective metrics, in some cases influencing their ability to discriminate between groups. In fact, prior studies have varied substantially regarding the mask used to compute FA and MD histograms, which have included the TBSS skeleton, WM, NAWM and WMH. The choice of the mask is critical as it determines the whole-brain region that is probed for pathological microstructural changes underlying the disease. Previous studies reported weak correlations between DTI parameters within lesion areas and cognitive impairment [24, 38], suggesting that WMH regions are not representative of early brain impairment. Moreover, WMH regions also typically include relatively few voxels rendering them less statistically valid than larger regions. For these reasons, we decided to restrict our analysis to TBSS and NAWM masks, which exclude WMH regions. By using a NAWM mask, we aimed to detect microstructural changes that are not yet manifested on T1WI or T2WI. However, NAWM masks are very likely to be contaminated by partial volume effects after transformation between structural and diffusion imaging spaces. In contrast, TBSS is an automatic technique that overcomes this issue by restricting the analysis to the centre of major WM tracts, the mean FA skeleton identified based on the highest FA values [10]. All FA metrics were more sensitive to mask selection than MD metrics (significant main effect of mask). This is expected given that MD values are substantially more uniform throughout the brain (with very minor differences between GM and WM and hence less impact from partial volume effects) than FA values. On the other hand, our results indicate that the ability of the metrics FA median, FA peak value, FA peak height, MD peak width and MD peak height to discriminate between patients and controls appears to be robust to the mask employed (significant main effect of Group). Nevertheless, FA peak height and MD peak width also showed a significant interaction between Group and Mask. Interestingly, post-hoc analysis revealed that MD peak width was able to differentiate between groups, but only when using the TBSS mask and the NAWM mask. These findings suggest that some of these metrics may be preferentially used as biomarkers in cerebral SVD; however, their association with cognitive variables may also influence their relevance as biomarkers.

4.2. Value of selected DTI metrics for predicting cognitive impairment

We found a significant correlation between executive function and the metrics MD peak height and MD peak width, regardless of the mask used. On the other hand, the correlation between executive function and FA Median only survived multiple comparison correction when using the NAWM mask. Although not significant after correction, a correlation was also found between processing speed (as measured by the respective composite score, TMTA + TMTB) and MD peak width. This result is consistent with the influential study by Baykara et al. [10], where MD peak width was consistently correlated with processing speed across multiple samples of patients. As in our study, the processing speed function of the CADASIL patients was assessed with the same composite score (TMTA + TMTB). However, unlike CADASIL, their sSVD sample was evaluated with a combination of 1-letter subtask of the Paper-Pencil Memory Scanning Test and Letter-Digit Substitution Task tests. Moreover, the authors reported that the association was weaker for the sSVD group. The discrepancy between tests used to assess processing speed, may partly explain the relatively weak association found in our sample, which is mostly composed of sSVD patients. Importantly, our methodology (0.3 skeleton threshold optimized to avoid partial volume effects; and linear interpolation, TBSS mask) bears a close resemblance to the one employed by Baykara et al [10]; the only difference is that they performed an additional step to remove regions close to the skeleton likely to be contaminated with CSF, by using a custom-made mask. We omitted this step as it would introduce operator variability. In summary, our results strongly support the MD peak width as a biomarker of SVD, as in Baykara, by further demonstrating a robust association with executive function.

Similarly to the MD peak width, MD peak height also showed significant positive association with executive function regardless of the mask. The MD peak height is the only DTI metric that has previously been reported to be correlated both with executive function and processing speed (from the same longitudinal study: Lawrence et al. [16]; Zeestraten et al. [11, 12]). An equivalent processing approach was also initially implemented to baseline data of their longitudinal study: NAWM mask by Lawrence et al. [16]. Nonetheless, our results regarding MD peak height were not correlated with processing speed which might be once more related to differences in the cognitive measures used in the two studies. The most recent reports from their group demonstrated that including both NAWM and WMH regions in the WM mask produces larger changes over time in the DTI parameters, compared to including only NAWM; the observed changes were strongly correlated with cognitive decline of executive function. However, the prominent contamination by CSF of non-TBSS based masks should be considered carefully [40].

Regarding FA peak height, we found no correlation with cognitive measures, in contrast with a previous study [13]. One explanation might be the different cognitive measures used as representative of executive function, which partly differ from ours: digit span backwards (from the Wechsler Memory scale III), verbal fluency (from the Delis-Kaplan Executive Function system) and TMTB. In addition, our sample is smaller and less homogenous than the one assessed by the above-mentioned study. Indeed, including CADASIL patients might have introduced a higher variability in DTI parameters as well as in cognitive decline, and thereby influence the respective correlations. On the other hand, this discrepancy is also likely to reflect the impact of using different processing strategies. In fact, in contrast with our methodology, their study included GM regions while we only included WM regions (TBSS vs NAWM). It remains unclear, however, how the inclusion of more diffusion-isotropic areas (such as GM) influences the correlation with cognitive

measures. Thus, our results indicate a lower stability of FA peak height as a biomarker, not only due to the lack of sensitivity to distinguish between patients and controls but also due to the absence of correlations with cognition. Furthermore, a prior longitudinal study found that reductions in WM microstructural integrity did not result in substantial FA peak height changes over a three-year period [41]; which also suggests the lack of sensitivity of this metric to patients' cognitive decline.

Overall, our results revealed a greater sensitivity and stability of MD compared to FA measures in terms of their potential to correlate with cognitive performance of SVD patients, which is in line with previous reports [11, 12, 16].

4.3. Limitations

Some limitations of this study should be addressed in the future. Firstly, a clinical protocol (with $1.7 \times 1.7 \times 5.2 \text{ mm}^3$ resolution) was used to acquire the DTI data, which might not be ideal for the analysis to be performed. The acquisition protocol could be further optimized by using isotropic voxels and sampling a higher number of gradient directions instead of multiple repeats of the same set of 20 directions. A further improvement would be the additional acquisition of a non-DWI scan with opposing phase-encode direction, in order to allow the correction of geometric distortions due to B0 inhomogeneities using FSL's TOPUP tool [42]. However, some of these optimizations might imply longer acquisition protocols, unless more sophisticated sampling schemes are used [43]. On the other hand, our results indicate that clinical MRI protocols might provide sufficient information for the extraction of DTI biomarkers of cerebral SVD. Another important limitation is the small number of subjects in our sample, which constrained the statistical power of the analysis. Moreover, the limited availability of normative data for the studied population limited the selection of neuropsychological tests that can be used to reflect each cognitive domain. In addition, this cohort showed a lower lesion burden compared to previous reports. This might potentially limit the generalizability of the findings. Although FA and MD measures are indicative of altered tissue integrity (i.e. related to axonal degeneration and ischemic demyelination), they lack specificity regarding the concrete nature of the underlying histological changes. More sophisticated microscopic diffusion models have been developed to better inform the true nature of these changes and may be explored in future studies of SVD patients [9, 20]. In this regard, our results represent the first step towards a consensus analysis pipeline to generate comparable DTI histogram metrics capable of predicting cognitive impairment in SVD patients; this pipeline could be potentially applicable also to other clinical conditions such as multiple sclerosis [17].

5. Conclusion

In summary, our results support the hypothesis that FA and MD histogram-based metrics extracted from DTI are sensitive to differences between SVD patients and controls; and can also predict cognitive impairment in SVD patients. Importantly, our results emphasize the importance of properly selecting the mask used to extract the histograms. Overall, our findings support prior reports that the MD peak width has the best potential for correlation with both executive function and processing speed and should be used as a reference measure for the validation of novel biomarkers. In contrast, other DTI metrics are

more dependent on the underlying processing options. Finally, we recommend that the TBSS method be utilized to define the whole-brain WM mask, since it is a completely automated methodology that reduces CSF contamination when compared to non-skeletonized masks (NAWM or WM), especially when analyzing non-optimized clinical datasets. Overall, our results extend previous reports and further support the value of DTI histogram-based metrics as SVD biomarkers; but they also clearly highlight the importance of the processing methodology used, as well as the urgent need to mitigate the lack of standardized MRI data-processing pipelines.

Declarations

Competing interests: The authors declare no competing interests.

Author's Contributions

ARF – acquisition of data, analysis and interpretation of data, drafting of manuscript, and critical revision; **RGN** – study conception and design, analysis and interpretation of data, drafting of manuscript, and critical revision; **JP** – acquisition of data, analysis and interpretation of data, and critical revision; **LA** – patient recruitment, acquisition of data, and critical revision; **SC** – acquisition of data; **CG** – acquisition of data; **MR** – acquisition of data, and critical revision; **MVB** – patient recruitment, study conception and design, and critical revision; **PV** – study conception and design, analysis and interpretation of data, and critical revision; **PF** – study conception and design, analysis and interpretation of data, drafting of manuscript, and critical revision.

References

- [1] Charidimou A, Pantoni L, Love S. The concept of sporadic cerebral small vessel disease: A road map on key definitions and current concepts. *Int J Stroke* 2016;11:6–18.
- [2] Patel B, Markus HS. Magnetic resonance imaging in cerebral small vessel disease and its use as a surrogate disease marker. *Int J Stroke* 2011;6:47–59. <https://doi.org/10.1111/j.1747-4949.2010.00552.x>.
- [3] Prins ND, Scheltens P. White Matter hyperintensities, cognitive impairment and dementia: an update. *Nat Rev Neurosci* 2015;11:157–65. <https://doi.org/10.1038/nrneurol.2015.10>.
- [4] Benjamin P, Lawrence AJ, Lambert C, Patel B, Chung AW, MacKinnon AD, et al. Strategic lacunes and their relationship to cognitive impairment in cerebral small vessel disease. *NeuroImage Clin* 2014;4:828–37.
- [5] De Guio F, Duering M, Fazekas F, De Leeuw F-E, Greenberg SM, Pantoni L, et al. Brain atrophy in cerebral small vessel diseases: Extent, consequences, technical limitations and perspectives: The HARNESS initiative. *J Cereb Blood Flow Metab* 2020;40:231–45.

- [6] Charidimou A, Werring DJ. Cerebral microbleeds and cognition in cerebrovascular disease: An update. *J Neurol Sci* 2012;322:50–5.
- [7] Prins ND, Dijk EJ Van, Heijer T Den, Vermeer SE, Jolles J, Koudstaal PJ, et al. Cerebral small-vessel disease and decline in information processing speed, executive function and memory. *Brain* 2005;128:2034–41. <https://doi.org/10.1093/brain/awh553>.
- [8] Raja R, Rosenberg G, Caprihan A. Review of diffusion MRI studies in chronic white matter diseases. *Neurosci Lett* 2019;694:198–207.
- [9] Lyoubi-Idrissi AL, Jouvent E, Poupon C, Chabriat H. Diffusion magnetic resonance imaging in cerebral small vessel disease. *Rev Neurol (Paris)* 2017;173:201–10. <https://doi.org/10.1016/j.neurol.2017.03.005>.
- [10] Baykara E, Gesierich B, Adam R, Tuladhar AM, Biesbroek JM, Koek HL, et al. A Novel Imaging Marker for Small Vessel Disease Based on Skeletonization of White Matter Tracts and Diffusion Histograms. *Ann Neurol* 2016;80:581–92. <https://doi.org/10.1002/ana.24758>.
- [11] Zeestraten EA, Benjamin P, Lambert C, Lawrence AJ, Williams OA, Morris RG, et al. Application of Diffusion Tensor Imaging Parameters to Detect Change in Longitudinal Studies in Cerebral Small Vessel Disease. *PLoS One* 2016;11:1–16. <https://doi.org/10.1371/journal.pone.0147836>.
- [12] Zeestraten EA, Lawrence AJ, Lambert C, Benjamin P, Brookes RL, Mackinnon AD, et al. Change in multimodal MRI markers predicts dementia risk in cerebral small vessel disease. *Neurology* 2017;89:10–1212. <https://doi.org/10.1212/WNL.0000000000004594>.
- [13] Nitkunan A, Barrick TR, Charlton RA, Clark CA, Markus HS. Multimodal MRI in cerebral small vessel disease: Its relationship with cognition and sensitivity to change over time. *Stroke* 2008;39:1999–2005.
- [14] Croall ID, Lohner V, Moynihan B, Khan U, Hassan A, O'Brien JT, et al. Using DTI to assess white matter microstructure in cerebral small vessel disease (SVD) in multicentre studies. *Clin Sci* 2017;131:1361–73. <https://doi.org/10.1042/CS20170146>.
- [15] Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE, et al. Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. *Neuroimage* 2006;31:1487–505.
- [16] Lawrence AJ, Patel B, Morris RG, MacKinnon AD, Rich PM, Barrick TR, et al. Mechanisms of Cognitive Impairment in Cerebral Small Vessel Disease: Multimodal MRI Results from the St George's Cognition and Neuroimaging in Stroke (SCANS) Study. *PLoS One* 2013;8.
- [17] Vinciguerra C, Giorgio A, Zhang J, Donato I Di, Stromillo ML, Brocci RT, et al. Peak width of skeletonized mean diffusivity (PSMD) as marker of widespread white matter tissue damage in multiple sclerosis. *Mult Scler Relat Disord* 2019;27:294–7. <https://doi.org/10.1016/j.msard.2018.11.011>.

- [18] Mccreary CR, Beaudin AE, Subotic A, Zwiers AM, Alvarez A, Charlton A, et al. NeuroImage: Clinical Cross-sectional and longitudinal differences in peak skeletonized white matter mean diffusivity in cerebral amyloid angiopathy. *NeuroImage Clin* 2020;27:102280. <https://doi.org/10.1016/j.nicl.2020.102280>.
- [19] Duering M, Finsterwalder S, Baykara E, Tuladhar AM, Gesierich B, Konieczny MJ, et al. Free water determines diffusion alterations and clinical status in cerebral small vessel disease. *J Autism Dev Disord* 2017;47:549–62. <https://doi.org/10.1016/j.jalz.2017.12.007>.
- [20] Konieczny MJ, Dewenter A, Ter Telgte A, Gesierich B, Wiegertjes K, Finsterwalder S, et al. Multi-shell Diffusion MRI Models for White Matter Characterization in Cerebral Small Vessel Disease. vol. 96. 2021. <https://doi.org/10.1212/WNL.00000000000011213>.
- [21] Fazekas F, Chawluk JB, Alavi A, Hurtig HI, Zimmerman RA, Hurting H, et al. MR signal abnormalities at 1.5 T in Alzheimer's dementia and normal aging. *Am J Roentgenol* 1987;149:351–6. <https://doi.org/10.2214/ajr.149.2.351>.
- [22] Fernandes S. *Teste de Cores e Palavras de Stroop*. Lisboa: CEGOC-TEA; 2013.
- [23] Cavaco S, Gonçalves A, Pinto C, Almeida E, Gomes F, Moreira I, et al. Trail making test: Regression-based norms for the portuguese population. *Arch Clin Neuropsychol* 2013;28:189–98.
- [24] O'Sullivan M, Morris RG, Huckstep B, Jones DK, Williams SCR, Markus HS. Diffusion tensor MRI correlates with executive dysfunction in patients with ischaemic leukoaraiosis. *J Neurol Neurosurg Psychiatry* 2004;75:441–7.
- [25] O'Sullivan M, Barrick TR, Morris RG, Clark CA, Markus HS. Damage within a network of white matter regions underlies executive dysfunction in CADASIL. *Neurology* 2005;65:1584–90.
- [26] Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, et al. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 2004;23:S208–19.
- [27] Zhang Y, Brady M, Smith S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans Med Imaging* 2001;20:45–57.
- [28] Patenaude B, Smith SM, Kennedy DN, Jenkinson M. A Bayesian model of shape and appearance for subcortical brain segmentation. *Neuroimage* 2011;56:907–22.
- [29] Jenkinson M, Smith S. A global optimisation method for robust affine registration of brain images. *Med Image Anal* 2001;5:143–56.
- [30] Jenkinson M, Bannister P, Brady M, Smith S. Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *Neuroimage* 2002;17:825–41.

<https://doi.org/10.1006/nimg.2002.1132>.

[31] S.M. G, U.J. C, M.M. B, T.A. Y, C. K, H.R. J, et al. The Microbleed Anatomical Rating Scale (MARS): Reliability of a tool to map brain microbleeds. *Neurology* 2009;73:1759–66.

<https://doi.org/10.1212/WNL.0b013e3181c34a7d>.

[32] Wardlaw JM. What Is a Lacune? *Stroke* 2008;39:2921–2.

[33] Andersson JLRR, Sotiropoulos SN. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *Neuroimage* 2016;125:1063–78.

<https://doi.org/10.1016/j.neuroimage.2015.10.019>.

[34] Andersson JLRR, Graham MS, Zsoldos E, Sotiropoulos SN. Incorporating outlier detection and replacement into a non-parametric framework for movement and distortion correction of diffusion MR images. *Neuroimage* 2016;141:556–72. <https://doi.org/10.1016/j.neuroimage.2016.06.058>.

[35] Behrens TEJ, Woolrich MW, Jenkinson M, Johansen-Berg H, Nunes RG, Clare S, et al. Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magn Reson Med* 2003;50:1077–88.

[36] Behrens TEJ, Berg HJ, Jbabdi S, Rushworth MFS, Woolrich MW. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *Neuroimage* 2007;34:144–55.

[37] Grabner G, Janke AL, Budge MM, Smith D, Pruessner J, Collins DL. Symmetric atlas and model based segmentation: an application to the hippocampus in older adults. *Med Image Comput Comput Assist Interv*, vol. 9, 2006, p. 58–66. https://doi.org/10.1007/11866763_8.

[38] Tuladhar AM, van Norden AGW, de Laat KF, Zwiers MP, van Dijk EJ, Norris DG, et al. White matter integrity in small vessel disease is related to cognition. *NeuroImage Clin* 2015;7:518–24.

[39] Freedman D, Diaconis P. On the histogram as a density estimator:L 2 theory. *Zeitschrift Für Wahrscheinlichkeitstheorie Und Verwandte Gebiete* 1981;57:453–76.

<https://doi.org/10.1007/BF01025868>.

[40] Metzler-Baddeley C, O’sullivan MJ, Bells S, Pasternak O, Jones DK. How and how not to correct for CSF-contamination in diffusion MRI. *Neuroimage* 2012;59:1394–403.

<https://doi.org/10.1016/j.neuroimage.2011.08.043>.

[41] Williams OA, Zeestraten EA, Benjamin P, Lambert C, Lawrence AJ, Mackinnon AD, et al. Diffusion tensor image segmentation of the cerebrum provides a single measure of cerebral small vessel disease severity related to cognitive change. *NeuroImage Clin* 2017;16:330–42.

<https://doi.org/10.1016/j.nicl.2017.08.016>.

[42] Andersson JLRR, Skare S, Ashburner J. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *Neuroimage* 2003;20:870–88.
[https://doi.org/10.1016/S1053-8119\(03\)00336-7](https://doi.org/10.1016/S1053-8119(03)00336-7).

[43] Cheng J, Shen D, Yap PT, Basser PJ. Single- and Multiple-Shell Uniform Sampling Schemes for Diffusion MRI Using Spherical Codes. *IEEE Trans Med Imaging* 2018;37:185–99.
<https://doi.org/10.1109/TMI.2017.2756072>.

Figures

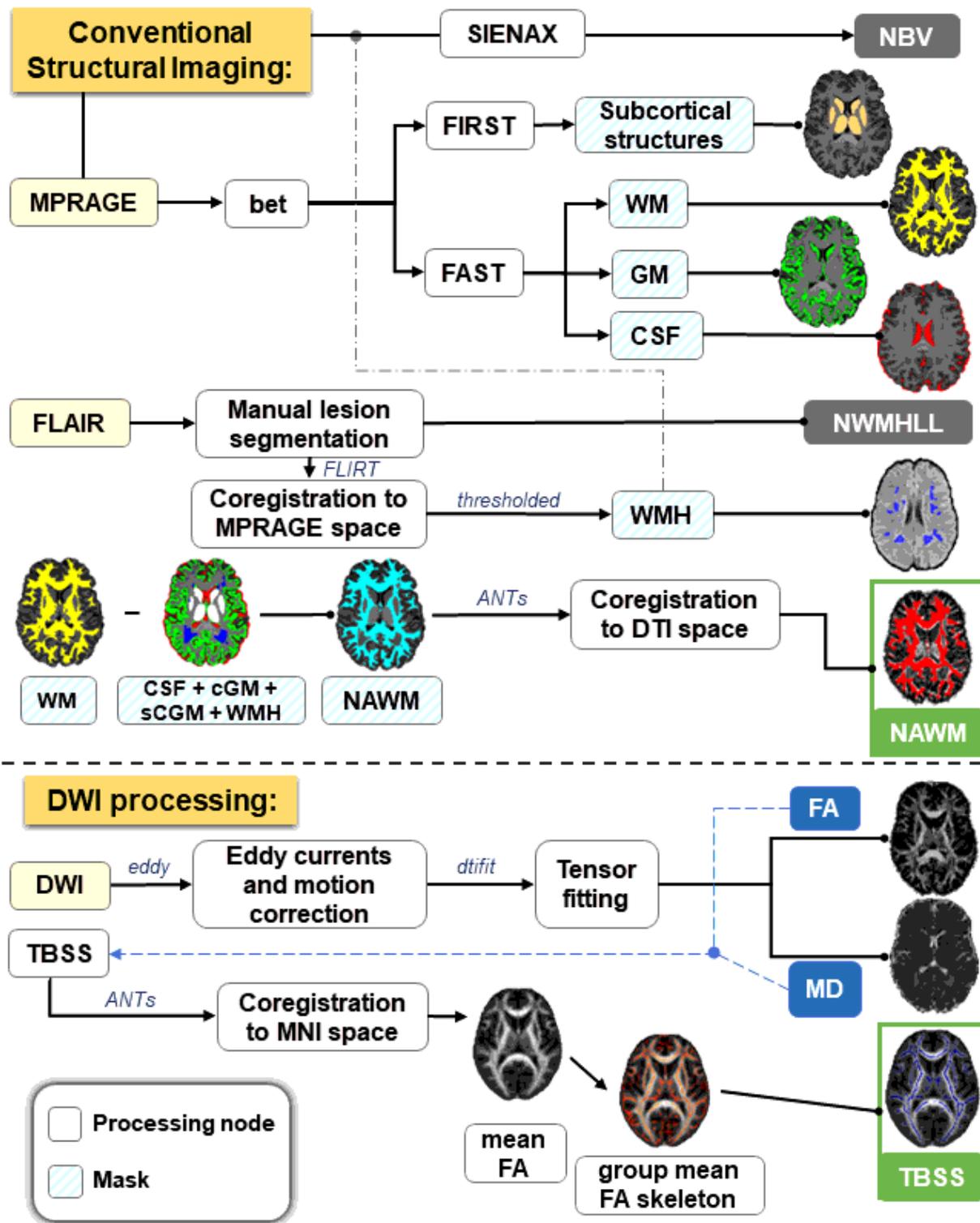
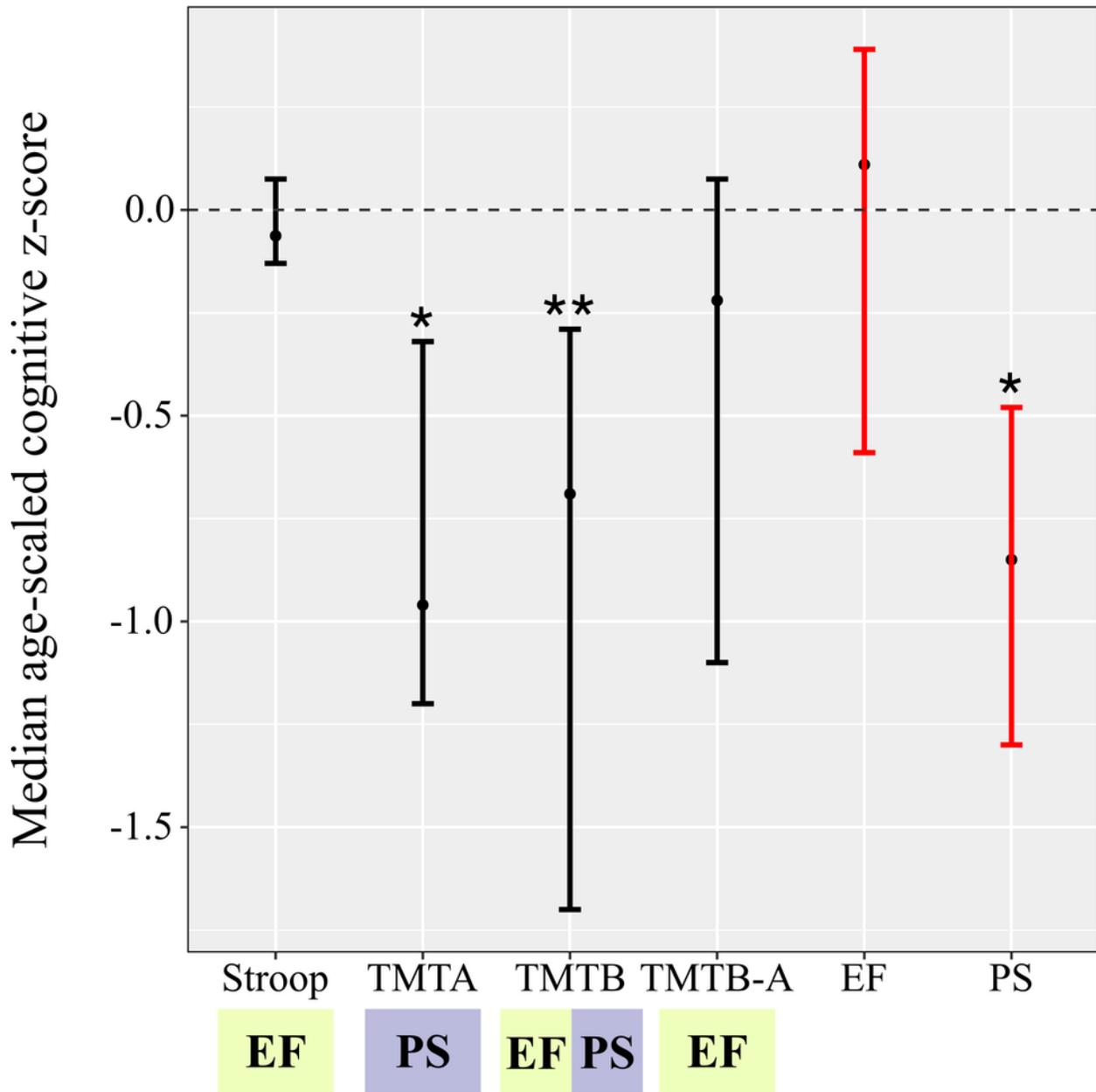


Figure 1

Schematic diagram of the image processing pipeline. The main steps to generate both TBSS and NAWM masks (highlighted in green boxes) used for histogram computation are illustrated.



EF = Executive Function
 PS = Processing Speed

Figure 2

Patients' cognitive profile. The group median age-scaled z-score is represented for each considered cognitive test (i.e., Stroop, TMTA, TMTB and TMTB-A, in black) and for the composite scores resulting from the combination of Stroop+TMTB-A (executive function) and TMTA+TMTB (processing speed), in red. Individual scores reflecting processing speed (PS) are labelled in blue and those reflecting executive functions (EF) are labelled in green. Bars represent the BCa (bias-corrected accelerated) Bootstrap

confidence interval at 95%. The dashed line (z -score=0) represents the expected performance for a healthy population. Significant impairment is represented by $** (p < 0.01)$ and $* (0.01 < p < 0.05)$. Performance is significantly impaired in TMTA and TMTB. In general, processing speed is significantly impaired, but executive function is not.

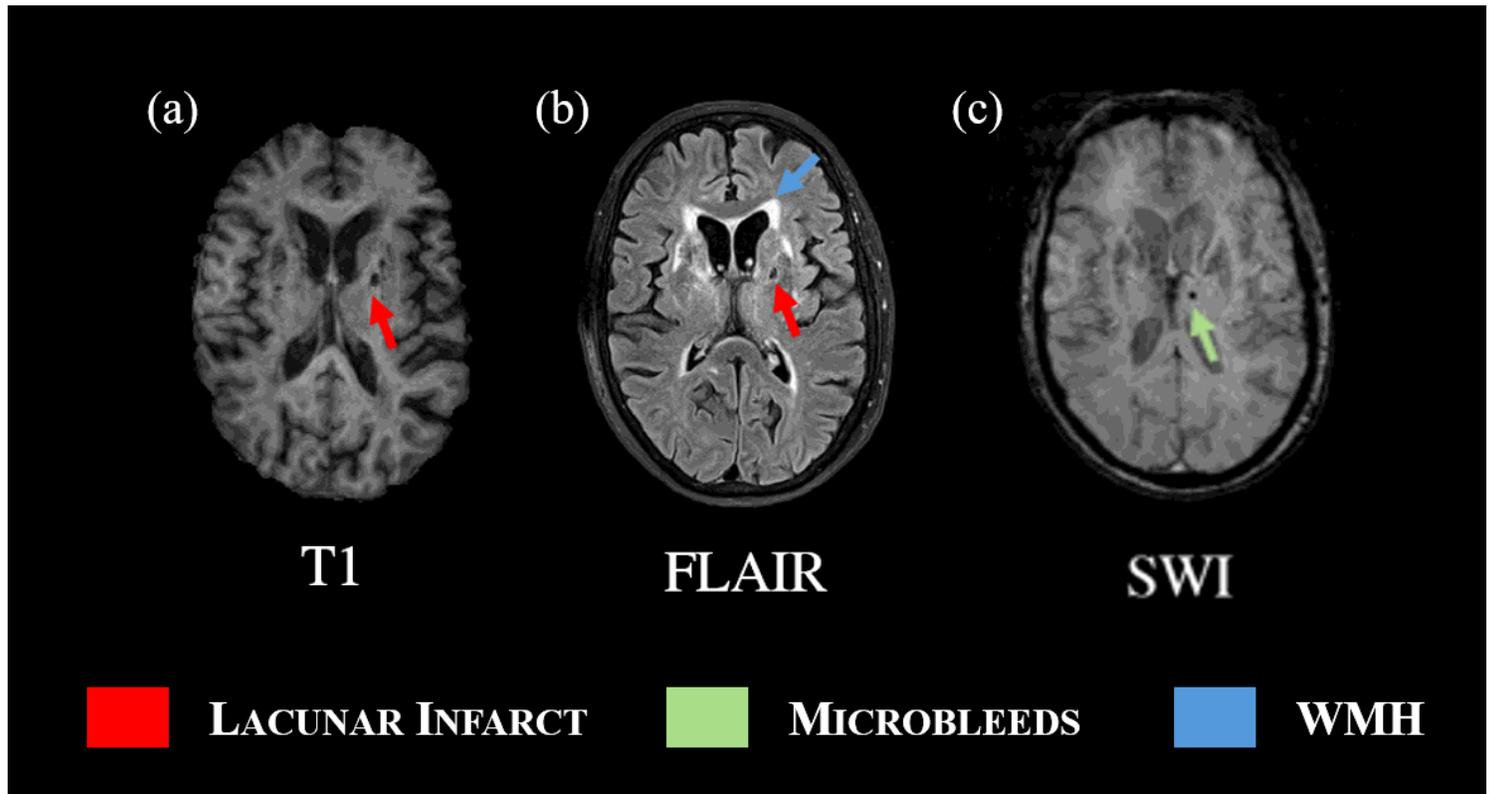


Figure 3

Illustrative examples of cerebral SVD lesions on conventional structural MRI from a representative patient. The image represents illustrative axial slices showing: (a) a lacune (red arrow) on T1WI; (b) a lacune (red arrow) and WMH (blue arrow) on FLAIR imaging; (c) a microbleed (green arrow) on SWI. T1WI = T1-weighted imaging; WMH = white matter hyperintensities; FLAIR = fluid-attenuated inversion recovery; SWI = susceptibility-weighted imaging

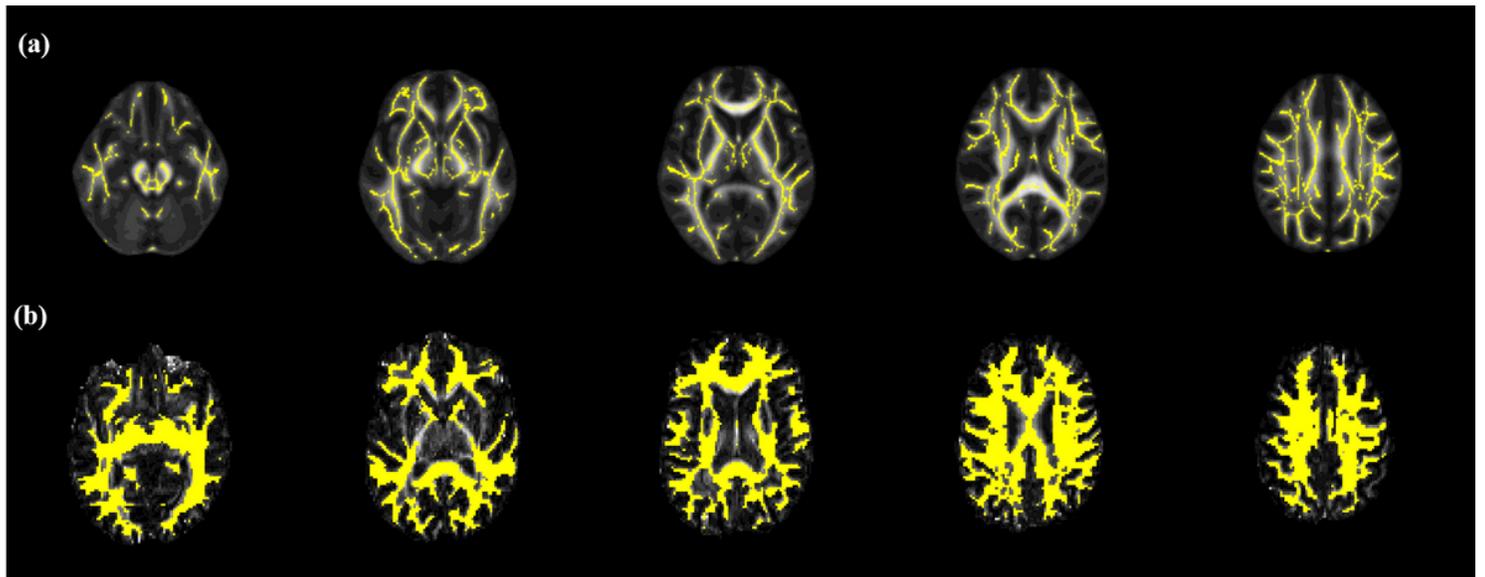


Figure 4

Illustrative examples of the different masks selected for histogram computation: TBSS and NAWM. (a) For TBSS, we show axial slices of the group mean FA map (MNI space) overlaid with the mean FA skeleton mask (yellow). (b) For NAWM, we show axial slices of the FA map of an individual (native space) overlaid with the NAWM mask (yellow). The images for TBSS and NAWM masks are not to scale.

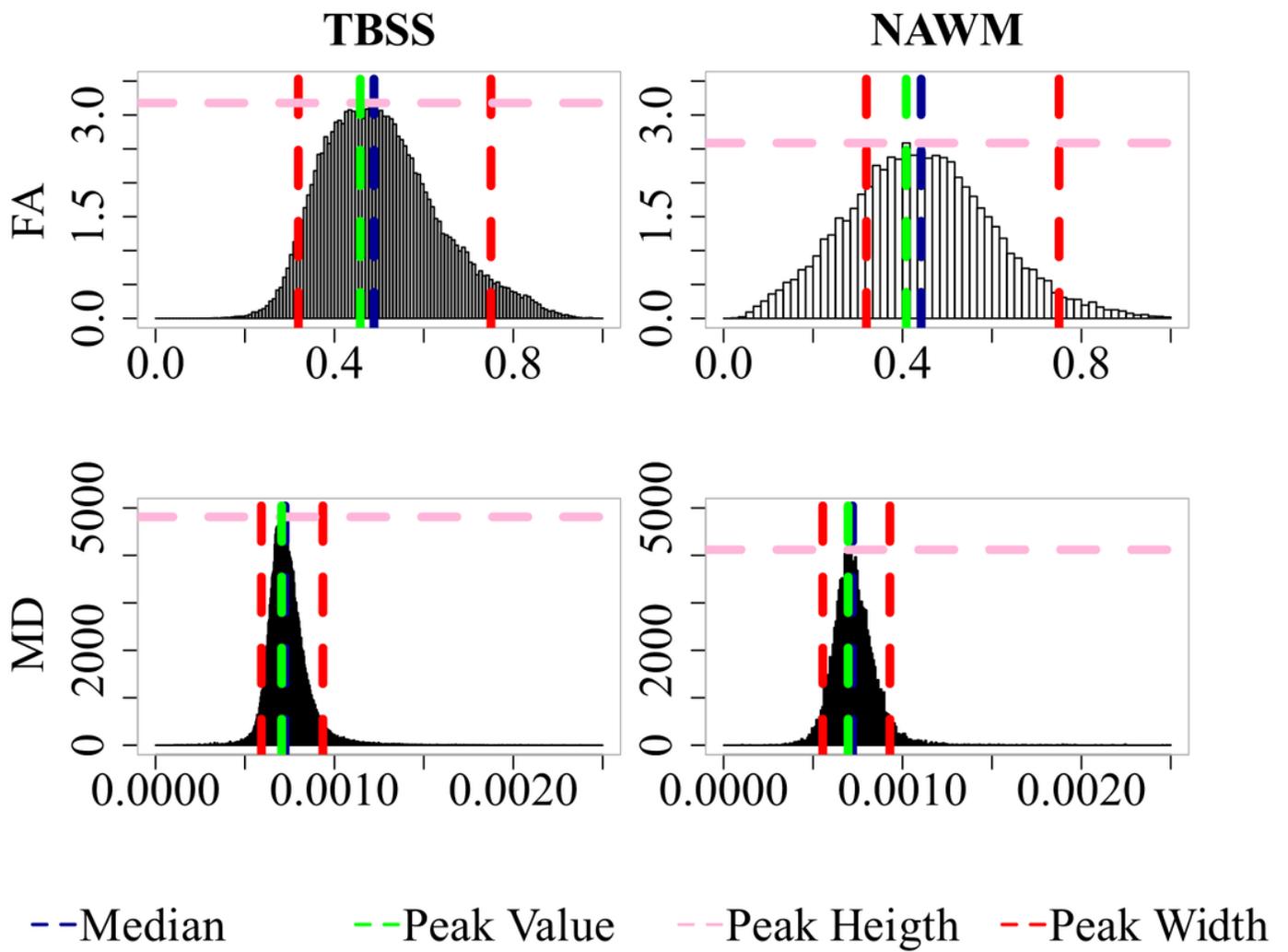


Figure 5

Illustrative examples of FA and MD histograms obtained using the different masks: TBSS (left) and NAWM (right), demonstrating the extraction of the histogram-based metrics (median, peak height, peak value and peak width).

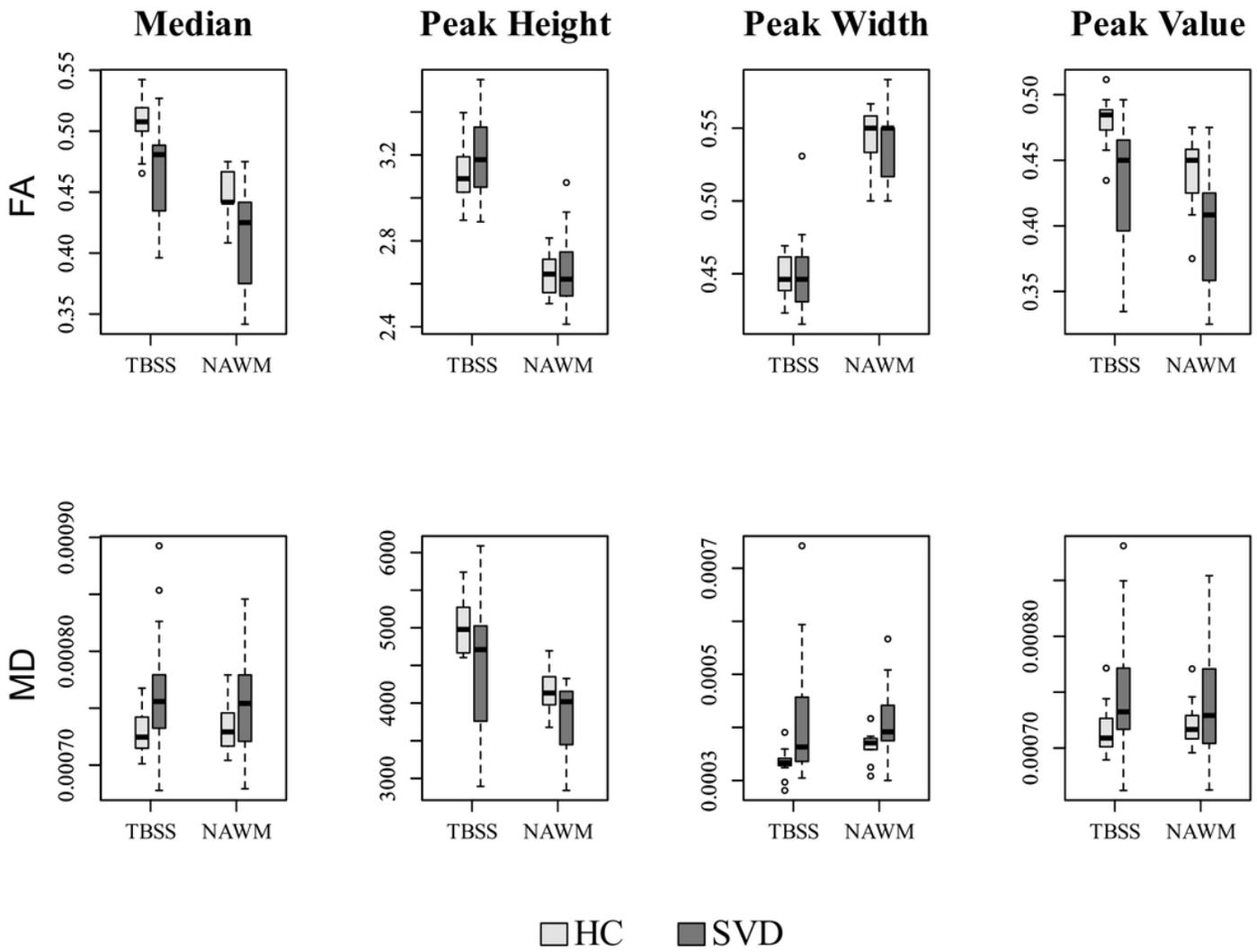
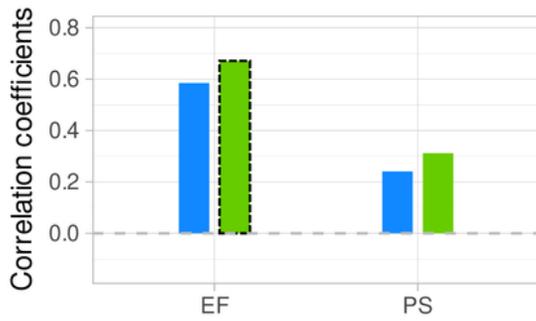
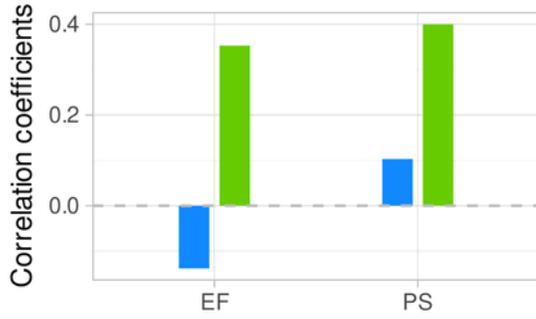


Figure 6

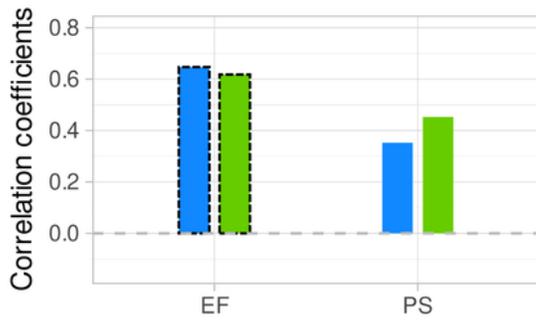
Boxplots showing the distributions across healthy subjects (HC) and SVD patients (SVD) of FA and MD histogram metrics (median, peak height, peak width and peak value), obtained using each mask (TBSS and NAWM).



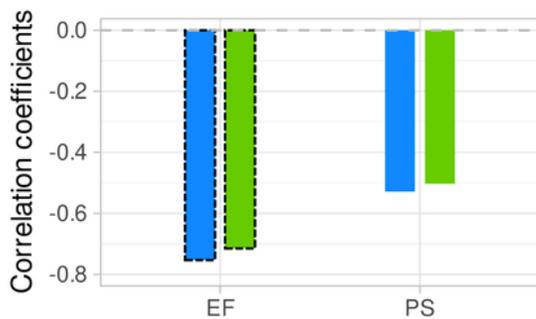
(a) FA Median



(b) FA peak height



(c) MD peak height



(d) MD peak width

■ TBSS ■ NAWM

Figure 7

Partial correlation analysis between DTI histogram-based metrics (FA median and peak height, MD peak height and peak width) and cognitive measures representing executive function and processing speed, while correcting for differences in age and sex. Significant correlations after correction for multiple comparisons are highlighted by boxes with dashed black borders.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryMaterialMAGMA.docx](#)