Using DTI to assess white matter microstructure in Cerebral Small Vessel Disease (SVD) in multi-centre studies

Croall ID[1*], Lohner V[1*], Moynihan B[2], Khan U[2], Hassan A[3], O'Brien JT[4], Morris RG[5], Tozer DJ[1], Cambridge VC[1], Harkness K[6], Werring DJ[7], Blamire AM[8], Ford GA[9], Barrick TR[10], Markus HS[1]

[1]Stroke Research Group, Department of Clinical Neuroscience, University of Cambridge
[2]St. Georges NHS Healthcare Trust
[3]Leeds Teaching Hospitals NHS Trust
[4]Department of Psychiatry, University of Cambridge
[5]Kings College Institute of Psychiatry, Psychology and Neurosciences, London, UK
[6]Deptartment of Neurology, Royal Hallamshire Hospital, Sheffield, UK
[7]Stroke Research Centre, Department of Brain Repair and Rehabilitation, UCL Institute of Neurology, UK
[8]Magnetic Resonance Centre, Institute of Cellular Medicine, Newcastle University, UK
[9]Oxford University Hospitals NHS Foundation Trust & University of Oxford
[10]Molecular and Clinical Science Research Institute, St George's, University of London, UK

*These authors contributed equally to this work.
Corresponding author: Prof. Hugh Markus, email: hsm32@medschl.cam.ac.uk

Abstract

Diffusion tensor imaging (DTI) metrics such as Fractional Anisotropy (FA) and Mean Diffusivity (MD) have been proposed as clinical trial markers of cerebral small vessel disease due to their associations with outcomes such as cognition. However, studies investigating this have been predominantly single-centre. As clinical trials are likely to be multi-site, further studies are required to determine whether associations with cognition of similar strength can be detected in a multi-centre setting. 109 patients (mean age=68) with symptomatic lacunar infarction and confluent white matter hyperintensities (WMH) on MRI was recruited across 6 sites as part of the PRESERVE DTI sub-study. After handling of missing data, 3T-MRI scanning was available from 5 sites on 5 scanner models (Siemens and Philips), alongside neuropsychological and Quality of Life (QoL) assessments. FA median and MD peak height were extracted from DTI histogram analysis. Multiple linear regressions were performed, including normalised-brain volume, WMH lesion load, and n^{o} lacunes as covariates, to investigate the association of FA and MD with cognition and QoL. DTI metrics from all white matter were significantly associated with Global Cognition (standardised β=.268), Mental Flexibility (β=.306), Verbal Fluency (β=.376), and MoCA (β=.273). The magnitudes of these associations were comparable to those previously reported from single-centre studies found in a systematic literature review. In this multi-centre study, we confirmed associations between DTI parameters and cognition, which were similar in strength to those found in previous single-centre studies. This study supports the use of DTI metrics as biomarkers of disease progression in multi-centre studies.

Abbreviations: **CABG**: Coronary Artery Bypass Graft, **CSF**: Cerebrospinal Fluid, **DSC**: Digit Symbol Coding, **DTI**: Diffusion Tensor Imaging, **FA**: Fractional Anisotropy, **FDT**: FMRIB Diffusion Toolbox, **FLAIR**: Fluid Attenuated Inversion Recovery, **FLIRT**: FMRIB Linear Image Registration Tool, **FSL**: FMRIB Software Library, **GM**: Grey Matter, **MD**: Mean Diffusivity, **MoCA**: Montreal Cognitive Assessment, **NART**: National Adult Reading Test, **NAWM**: Normal-Appearing White Matter, **NBV**: Normalised Brain Volume, **RAVLT**: Rey Auditory Verbal Learning Test, **QoL**: Quality of Life, **SPM**: Statistical Parametric Mapping, **SSQoL**: Stroke Specific Quality of Life, **SVD**: Small Vessel Disease, **T1W**: T1-Weighted, **T2*W**: T2*-Weighted, **TMT**: Trail Marking Test, **TPM**: Tissue Probability Map, **WM**: (all) White Matter, **WMH**: White Matter Hyperintensity

Introduction

Cerebral Small Vessel Disease (SVD) causes a quarter of all ischaemic strokes, is the most common pathology underlying vascular cognitive impairment and dementia (1), and contributes to the severity of Alzheimer's Disease (2). SVD affects the small vessels of the brain and results in a number of characteristic radiological appearances best seen on MRI, including lacunar infarcts, T2-white matter hyperintensities (WMH), cerebral microbleeds, and brain atrophy (3,4). In terms of symptoms, cognitive impairment may be the most debilitating (5), with SVD characteristically associated with early deficits in executive function and processing speed, while episodic memory is relatively spared (1,2,6–9).

Despite the public health importance of SVD, there are few specific treatments (10). Furthermore, evaluating treatments represents a major challenge due to the variable rate of cognitive decline which can be slow in many patients, but occurs rapidly with progression to dementia in a subset. Whilst cognitive testing plays a central role in identifying the presence of cognitive impairment, it has proved to be relatively insensitive to longitudinal change (11). This has led to the suggestion that MRI might represent a useful surrogate marker to monitor disease progression and evaluate the efficacy of therapeutic interventions in smaller patient numbers prior to larger phase 3 trials (3,12).

Diffusion Tensor Imaging (DTI) has been shown to be particularly sensitive to white matter damage in SVD. Abnormalities have been shown not only within T2-WMH but also in apparently "normal appearing white matter" (13), and these changes correlate better with cognition than WMH lesion volume (8). In single-centre studies, change on DTI could be detected in SVD patients over follow-up periods of 1 to 3 years (14,15). This has led to the suggestion that DTI might provide a useful surrogate marker, and power calculations for phase 2 trials based on the rate of DTI change seen in these papers have shown that its use may allow evaluation of therapeutic interventions with much smaller samples sizes than if cognitive function was used as an outcome measure (11). However, studies conducted to date have been single-centre (12,15,16). Most therapeutic trials are likely to be multi-centre and involve acquisition of DTI across different sites. As image acquisition will be on different scanners this may present challenges (17). It is important to assess whether DTI is feasible in a clinical trial setting, and whether similar associations between MRI parameters and clinical and cognitive variables can be detected in the multi-centre setting. One way of assessing this

is to determine whether the strength of association between DTI and cognition in multi-centre studies is similar to that previously reported in single-centre studies.

To evaluate this we determined the association between DTI parameters and cognition in the baseline data of a multi-centre trial.

Methods

PRESERVE study

The PRESERVE study ("How intensively should we treat blood PRESsure in established cERebral small VEssel disease?") is a multi-centre randomised controlled trial comparing a strategy of intensive, versus standard, treatment of blood pressure on cognitive function over a two year follow-up period. Nested within the overall study is a DTI substudy in which patients additionally undergo multimodal MRI including DTI at baseline and at the end of the two year follow-up period. The baseline data from these individuals is presented in this paper.

Study Population

Inclusion criteria were a clinical lacunar stroke with an anatomically corresponding lacunar infarct on MRI, in addition to confluent WMH graded as ≥ 2 on the Fazekas scale (18). Patients were at least 40 years old with hypertension defined as either a systolic blood pressure >140mmHg, or a systolic blood pressure between 125-140mmHg while on antihypertensive treatment. Exclusion criteria were: a known single gene disorder causing SVD (e.g. CADASIL), symptomatic carotid stenosis or vertebral stenosis >50% , cortical infarction >2cm diameter, diagnosis of dementia, life expectancy of less than two years, symptomatic postural hypotension, women of childbearing potential and any inability to fulfil study data collection. All patients gave informed written consent. The study was approved by the Harrow NRES ethics committee (REC number: 11/LO/0458), and is registered with the UK Clinical Research Network (CRN number: 10962).

109 patients from 6 sites consented to participate in the PRESERVE DTI sub-study. The site sample sizes are as follows: Site 1 (N=48), Site 2 (N=29), Site 3 (N=14), Site 4 (N=11), Site 5 (N=6), Site 6 (N=1). Participants underwent baseline testing at least three months post-stroke.

Clinical Assessments

A stroke physician or vascular neurologist examined all participants. Cerebrovascular risk factors including a history of previous stroke, hypercholesterolaemia, diabetes, smoking (current and history), angina, myocardial infarction, coronary artery bypass grafts or coronary angioplasty were recorded.

Neuropsychological Assessment

Cognitive Testing: Assessment was performed by a neuropsychologist and occurred on the same day as MRI scanning, or as close to the scan as possible. A cognitive test battery was used which included tests sensitive to the characteristic impairments in processing speed and executive function associated with SVD (2), with additional testing of memory. This included for processing speed the Digit Symbol Coding test (DSC) (19), and for executive functioning the Trail Marking Test (TMT, (20)) to measure mental flexibility, and a phonemic verbal fluency task (FAS) (21) and a semantic verbal fluency task (animals) (22) to measure verbal generativity. Memory was measured using the Rey Auditory Verbal Learning Test (RAVLT, (23)). Premorbid IQ was estimated using the restandardised National Adult Reading Test (NART-R, (24)) and additional screening for cognitive impairment was conducted using the Montreal Cognitive Assessment (MoCA, (25)).

In addition the following assessments of disability and quality of life were performed; the Stroke Specific Quality of Life assessment (SSQoL) (26), and the EuroQoL (27).

Performance across neuropsychological tests was made comparable by transforming raw scores into z-scores using the best available age-scaled normative data (DSC; (19), TMT; (28), Letter Fluency; (21), Animal Fluency; (22), RAVLT; (28)). Tasks were grouped into four key domains (**Processing Speed**: WAIS coding total correct, TMT-A time to complete, **Mental Flexibility**: TMT-B time to complete, **Verbal Fluency**: total correct for "FAS" letter fluency and Animal fluency, and **Verbal Memory**: RAVLT "immediate" and "delayed" recall). Individual task z-scores were averaged across these groupings to create overall domain scores, while all domain scores were averaged to create a **Global Cognition** domain. SSQoL (total score), EuroQoL ("healthstate" rating) and the MoCA (total score) were analysed individually using raw scores.

Where data were missing due to a subject being unable to complete a task the lowest available Z score was given; this applied to 15 individual tasks, across 13 participants (11.9%

of the sample size). If data were missing for any other reasons then the domain scores were calculated without that task; this applied to 3 participants (2.8% of the sample size).

Magnetic Resonance Imaging Acquisition

The aim was to test a study design for which MRI data was acquired using clinical scanners in different sites from different manufacturers. Within the 6 centres, 8 3-Tesla MR scanners were used (3 Philips Acheiva TX, 1 Philips Acheiva, 1 Philips Ingenia, 1 Siemens Verio, 1 Siemens Prisma, 1 Siemens Magnetom Prisma^{fit}). MRI acquisition included 3D T1-weighted (T1W), and DTI, T2*-weighted (T2*W), and Fluid Attenuated Inversion Recovery (FLAIR) scans for each participant. A rigorous quality control was implemented to ensure sequence acquisition parameters were as standardised as possible. T1W scans were acquired at 1mm^3 isotropic voxel resolution and TR and TE were optimised to ensure comparable T1 weighting and tissue contrast across sites. DTI scans (2mm^3 isotropic voxel resolution) had similar TEs and long TRs to avoid T1 relaxation effects. In addition to the b = 0 s mm^{-2} acquisitions, all DTI acquisition included 32 equally spaced, non-collinear diffusion gradient directions (b = 1000 s mm^{-2}) to ensure identical angular resolution and noise characteristics. T2*W sequences were TE matched and kept a similar TR to ensure comparable weighting. FLAIR sequences had identical inversion times and were also TE matched with long enough TR's to ensure no T1 weighting occurred. Resolution for T2*W and FLAIR sequences varied between sites; supplementary Table 1 gives an overview of the exact scanner and sequence details per site.

MRI data analysis

In addition to DTI, measures describing WMH, lacunes and brain volume are frequently investigated as potential markers of SVD (8,12,29–31). In the present study, these were analysed as a comparison to DTI.

WMH: WMH were defined as areas of increased signal on FLAIR images (excluding the rims of cavitated lacunes), and segmented by a single trained rater (I.D.C.) using a semi-automated, contouring technique in Jim image analysis software version 7.0_5 (Xinapse Systems Limited, http://www.xinapse.com/j-im-7-software/). Whole brain WMH lesions maps were generated and a WMH lesion load score was calculated as the percentage of WMH lesion volume against whole brain volume. To assess intra- and inter-rater reliability a test set of 10 FLAIR scans (from a previous study in SVD) with varying degrees of WMH

was used. In a randomised, blinded setting FLAIR images were each marked twice by I.D.C. and once by a second experienced rater (D.T.). The intraclass correlation coefficient (32) was calculated to assess inter-rater reliability (I.D.C. vs. D.T.) and intra-rater reliability providing coefficients of .988 and .998 respectively.

Lacunes: Lacunes were defined as cerebrospinal fluid (CSF) filled cavities at least 3mm in diameter. Additional features such as T2-hyperintense rims, shape and location were also considered to differentiate lacunes from similar imaging features such as perivascular spaces. The same single rater (I.D.C.) identified lacunes after training by a consultant neuroradiologist using a combination of T1W, T2*W and FLAIR scans.

Brain Volume: T1W scans were intensity non-uniformity corrected using "N4ITK" (33) and segmented into grey matter (GM), white matter (WM) and CSF tissue probability maps (TPM) using SPM12b (Statistical Parametric Mapping, http://www.fil.ion.ucl.ac.uk/spm/ (34)). Brain volume in native space was calculated from the soft segmentation of the GM and WM TPMs.

To obtain brain volume measures sensitive to atrophy, SIENAX ((35), a part of FSL; FMRIB Software Library, https://fsl.fmrib.ox.ac.uk/fsl (36)) was applied to T1W scans giving a scaling factor that describes the variation of brain size relative to the skull size. The native space brain volumes were multiplied by this scaling factor to provide normalised brain volumes (NBV). To minimise tissue misclassification of WMH as GM, the (normalised) volume of any GM which occurred within WMH was subtracted from the GM volume and added to the WM volume. Finally, whole NBV was calculated by adding GM and WM NBV's together.

DTI Histogram Analysis: FSL software (FDT; FMRIB's Diffusion Toolbox, http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FDT) was used for DTI pre-processing. Briefly; DTI scans were eddy current-corrected with eddy_correct using the 1st acquired b = 0 s mm^{-2} image as the reference. A binary brain mask in DTI space was calculated for each subject using BET on the same b = 0 acquisition. Fractional anisotropy (FA) and mean diffusivity (MD) maps were then calculated from these data using DTIFIT. Voxels with MD values above 0.0026mm^{2}s^{-1} were removed from analyses in case of them having been misclassified as CSF voxels by application of a diffusivity threshold. Likewise, spurious voxels with FA >1 were also removed. For each participant, FMRIB Linear Image Registration Tool (FLIRT, (37), using the normalised mutual information cost function in FSL) was used to register the

FLAIR to the T1W image, and the T1W to the b0 image (the average of all the b= 0 s mm^{-2} images in the DTI sequence). These affine transformation matrices were concatenated to create a third FLAIR-to-DTI transformation. TPM's and WMH lesion masks were registered into DTI space using the T1W-to-DTI (trilinear interpolation), and FLAIR-to-DTI (nearest neighbour interpolation) transforms for TPMs and binary WMH lesion masks, respectively.

A hard segmentation method was applied to generate maps of tissue classes. This was achieved by voxelwise comparison of the GM, WM and CSF TPMs, with each voxel being assigned to the highest probability tissue class. The WMH lesion masks were then added with these lesion voxels being automatically assigned to WMH. Finally, mask images of normal-appearing white matter (NAWM) and all white matter (WM) were generated from the hard segmentation map.

Histogram analysis was performed on FA and MD maps in both NAWM and WM. Normalised histograms with 1000 bins (FA range 0-1, bin width 0.001; MD range 0-4mm^2s^{-1} $x10^{-3}$, bin width 0.004mm^2s^{-1} $x10^{-3}$) were computed and median, peak height and peak value were extracted from these for both FA and MD. These metrics were chosen as summary measures as FA and MD are non-normally distributed in WM.

Statistical Analyses: All analyses were performed using IBM SPSS statistics version 23 (IBM Corp. Released 2013. IBM SPSS Statistics for Windows. Armonk, NY: IBM Corp, http://www.ibm.com/analytics/us/en/technology/spss/).

One measure for each MD and FA was chosen for the main study analyses. MD (normalised) peak height and FA median were picked due to previous studies which have shown these to be correlated with cognition (8,14), and sensitive to change in WM microstructure in SVD (11,15).

To compare MRI with cognitive parameters, "Simple" and "Complex" model linear regressions were conducted. This pipeline was structured as a method of selecting the most appropriate MRI measure per type (e.g. one brain tissue volume measurement, or MD / FA histogram parameter for DTI) so that contributions of MRI metrics could be assessed together while avoiding issues of multicollinearity. Thus, in Simple models the association of NBV, WMH lesion load, lacunes and histogram parameters (from NAWM and WM) were separately investigated against each outcome measure (cognitive domains, QoL and MoCA). As there were multiple NBV and DTI variables, the most significant of each type (or if *p* value was the same, the one with the largest β-value), per outcome measure, was selected and

used in the Complex model. Here, NBV, WMH lesion load, n^{o} lacunes, and DTI measures were included together to assess their contributions relative to each other. Separate Complex models were performed for each outcome measure, in WM and NAWM. These models controlled for confounding effects of age, gender, premorbid IQ, and were stratified by study site. Residuals were inspected for normality for all regression analyses while variance inflation factors were also calculated for the Complex Models to assess multicollinearity.

Further analyses compared DTI and outcome variables between sites, and repeated some Complex model analyses using site-specific data. These are detailed in the Supplementary Analysis.

Systematic Review: To allow comparison of the results with previous single-centre studies, a systematic review of previous literature was conducted on Pubmed (https://www.ncbi.nlm.nih.gov/pubmed/) using search terms of "cerebral small vessel disease diffusion tensor imaging", "white matter hyperintensities diffusion tensor imaging", and "leukoaraiosis diffusion tensor imaging" on the 16th of March, 2017. Criteria for inclusion were: **1.** Studies of sporadic SVD population (i.e. monogenic causes of SVD such as CADASIL were not included), **2.** Studies investigating the relationship between DTI metrics and cognitive performance, **3.** Studies investigating the cognitive domains analysed in the current study, **4.** Analysis controlling for at least 1 other confounding MRI measure, **5.** Results involved reporting of standardised β-values or partial correlation coefficients. Where a paper reported multiple associations against the same cognitive outcome, the strongest (i.e. largest β-value) was included. In cases where a study had published multiple papers based on the same participant data, the one which used the most similar metrics to those in the presented study was chosen.

Results

Profile of Participant Variables

Missing Data: Due to the low sample size (N=1), Site 6 was excluded from all statistical analyses. An additional 6 participants were excluded from analysis due to MRI data acquisition problems (2 cases from Site 1 due to excessive motion artefacts and corrupted data acquisition, and 4 cases from Site 4 where not all imaging sequences were acquired and some data were corrupted). Sample size was further reduced by incomplete cognitive data.

Verbal Fluency data was absent for one participant, Verbal Memory and NART in another, and (only) NART in a third. Sample size was therefore reduced by a further 3 for Verbal Fluency comparisons, and by 2 for all other comparisons. Consequently, complete DTI data were available in 102 participants, while sample size for main statistical analyses was N= 99 for testing Verbal Fluency, or N= 100 for all other outcome measures.

Demographics: Demographics, risk factors and clinical features are shown in Table 1.

All entry MRI scans were reviewed centrally by a consultant neurologist. All cases fitted the MRI inclusion criteria except for two which had WMH graded on the Fazekas scale of <2. Both were included in analysis as they had multiple lacunes consistent with severe SVD.

Cognition: The cognitive profile of the participants, is shown in Figure 1. All 5 cognitive domains were significantly impaired compared to control performance levels (p= <.001 in all cases except for Verbal Fluency where p= <.05).

MoCA, QoL and MRI Results: Mean values for MoCA, SSQoL, EuroQoL, and MRI parameters are shown in Table 2. Qualitative comparison of histogram measures between the WM and NAWM tissue classes showed that the inclusion of WMH in the WM lowered the (normalised) peak height of FA and MD, increased the peak value and median of MD, and decreased the peak value and median of FA.

Relationship between MR variables and cognition

Simple Model Analyses: Full findings are shown in Table 3. FA median and MD peak height (in WM and NAWM) were significantly associated with all outcome measures, except for NAWM MD peak height with Processing Speed, both NAWM measures with SSQoL, and all DTI measures with Verbal Memory. Median FA held stronger associations than MD peak height in all cases except for EuroQoL in (all) WM. The directions of these relationships demonstrate that higher median FA and MD peak height were associated with better cognition or QoL in both tissue classes. There were no marked differences between the patterns or strengths of associations for DTI measures taken from within NAWM or the whole of the WM. Whole NBV held stronger associations than GM or WM NBV in all cases except EuroQoL, where WM was strongest.

Complex Model Analyses: "Complex Models" were performed to determine which MRI variables were independently associated with the outcome measures, and results are shown in Table 4. The Variance inflation factors of all models were smaller than 3 and deemed

acceptable. Median FA was significantly associated with Global Cognition, Mental Flexibility, Verbal Fluency and MoCA in both tissue classes. No other comparisons with DTI metrics reached significance. Considering the significant associations, the effect sizes of the WM comparisons (as indicated by the β-value) was always descriptively greater than the NAWM counterpart.

The number of lacunes was independently significantly associated with Global Cognition, Processing Speed, MoCA and SSQoL in both tissue class models. NBV only maintained a significant association with MoCA and EuroQoL (in both tissue class models). WMH lesion load was no longer significantly related to any outcome measures.

Systematic review: Comparison of strength of associations between DTI and cognition with that from previous studies

The search terms identified 230 papers, and after reading these abstracts 37 selected for review. An additional 5 papers were identified from reference lists. Eight of these 42 papers met inclusion criteria (8,12,29–31,38–40). Supplementary Table 2 details these papers and includes key findings from each study. Of note, one of these (39) is a multi-centre study across 3 sites using identical 1.5T scanners and acquisition sequences, with MoCA and MMSE used as cognitive measures.

Two of these papers reported 95% confidence intervals (CI) with their β values for associations between DTI metrics and cognition (12,38). Comparing the magnitude of the DTI-based β values (ignoring direction, as this will be influenced by the specific DTI parameter used, which differs between papers) from the presented study for the same cognitive domain shows that these fell within, or were higher than these previously reported CIs for Global Cognition (our β= **.268**, previous CIs= **-.22** to **.06** (12), and **-.38** to **.02** (38)) Executive Functioning (i.e. Mental Flexibility; our β= **.306**, previous CIs= **-.16** to **-.06** (12), and **.05** to **.39** (38)), Verbal Fluency (our β= **.376**, previous CIs= **-.21** to **-.02** (12)) and Verbal Memory (our β= **.099**, previous CIs= **-.28** to **-.06** (12)). Only the presented β for Processing Speed was lower than a previously reported CI (but only in one of these papers; our β= **.058**, previous CIs= **-.24** to **-.06** (12), and **-.33** to **.06** (38)). Conversely, previously reported β values from all 8 papers fell within the CIs found in the presented analyses in all instances except for one case of Verbal Memory being greater than our CI (previous β= **-.86** (31), our CI= **-.157** to **.355**) and one case of Verbal Fluency being lower than our CI (previous β= **-.11** (12), our CI= **.140** to **.612**).

Site-Specific Findings

These analyses are reported in full in the Supplementary Material

In order to assess any variation across individual sites, analyses were conducted on data from each site individually. FA median and MD peak height of each site were compared by one-way ANOVA, which returned a non-significant finding for each (FA $p= .424$, MD $p= .148$). Comparison of all outcome measures (i.e. cognitive domains, MoCA and QoL scales) between sites by one-way ANOVA and Kruskal-Wallis also showed no significant findings (p value range: $.192$ to $.827$).

"Complex Model" analyses were also repeated in Sites 1, 2 and 3 individually. These were repeated in cases where a DTI metric had been shown to have a significant relationship with a cognitive domain in the main analyses. These relationships were further visualised by scatterplot in *all* sites, with the 95% CI around the total regression line also included for comparison. Complex Model results showed Sites 1 and 2 to have β-values which were within, or higher than the 95% CI limits for the same comparison in the main analyses. While this was also true for Site 3 in the Global Cognition model, the Mental Flexibility and Verbal Fluency models gave a lower β-value than the CI limits. The scatterplot with the "weakest" (i.e. flattest) individual-site fit is included here as Figure 2. This shows the relationship between WM FA median and Mental Flexibility, with a weak fit for Site 4 (but not Site 3) in that its line falls outside of the total CI limits in a manner showing it to be flatter. Supplementary Figures 1 and 2 repeat this scatterplot for Global Cognition and Verbal Fluency comparisons, and likewise indicate Site 3 (but not Site 4) to have a weak fit in each. All other sites show either good fits (i.e. fall completely within the CI limits; see Site 1 in Figure 2), or "strong" ones (i.e. fall outside of the total CI limits in a manner showing them to have steeper slopes; see Sites 2, 3 and 5 in Figure 2). This suggests that the majority of sites do contribute to the main study findings. It is possible that individual cases of small Complex Model β-values, and unusually "weak" / "strong" scatterplot fits are due to lack of power from low sample sizes.

Discussion

In this analysis of baseline data from a multi-centre clinical trial of SVD, we found associations between DTI metrics and cognition of a similar magnitude to those reported in

previous single-centre studies. This provides support for the use of DTI measures as surrogate markers in clinical trials of SVD.

We found that both DTI markers and lacunar infarct count were independently associated with Global Cognition and MoCA results. Additionally, DTI markers were independently associated with Mental Flexibility and Verbal Fluency, and lacunes with Processing Speed and SSQoL. In contrast we found no independent associations between WMH lesion load and cognition, and only two for brain volume (with MoCA and EuroQoL). This is in line with most previous literature from single-centre studies, which has found weak or absent associations between WMH and cognition in patients with severe symptomatic SVD (8,29,39). However, it has been previously shown, as we also have, that the presence and number of lacunar infarcts (8,38), and the extent of diffuse WM damage assessed on DTI (8,12,29,31,38,39), are the strongest predictors of cognitive functioning. Furthermore, both have been shown to predict risk of dementia in longitudinal studies (41,42), while lacunes and the Apparent Diffusion Coefficient (a diffusion-weighted imaging measure highly similar to MD) have also been shown to predict future cognitive decline (43,44). Number of lacunes was chosen in the present study instead of lacune volume as it is a more practical measure to obtain in a clinical setting, and similar associations with cognitive performance have been found between these in a comparable severe SVD population (45).

Clinical trials of new agents in SVD will need to be multi-centre and if MRI is to be used as a surrogate marker it is important to evaluate how the different markers perform in a multi-centre setting. While research in other neurological disorders such as Parkinson's (46) and Huntington's (47) disease have shown that DTI markers of disease can be successfully applied in a mutli-centre study, there have been few studies addressing this issue in SVD. The use of multiple scanners, possibly from different manufacturers, is likely to add noise, and may diminish the statistical sensitivity of these metrics.

PRESERVE is one of the first studies to use advanced MRI imaging as a surrogate marker in SVD trials. In this setting we have shown that the magnitudes of associations between DTI and cognition are highly comparable to previous, single-centre studies, further validating the use of these metrics in this context. Additionally, while WM and NAWM DTI were always significantly associated with the same outcomes, the strengths of these associations was consistently descriptively greater in WM models. This indicates the simpler process of obtaining a WM mask is at least equally valid, and may be more practical in a clinical setting.

It should also be noted that previous research has indicated through power calculations that DTI parameters could detect change with much smaller sample sizes than lacunes, due to the frequency of new lacunes being relatively low (11). This suggests that DTI metrics may be the most powerful surrogate marker of the two.

Examination of individual site data did demonstrate some variation in the strength of associations between MR parameters and cognition from different centres, but the majority of these effect sizes were within (or greater than) the expected ranges as determined by 95% CIs for β-values and regression slopes from the main analyses. DTI metrics and cognition did not significantly differ between sites, meaning it is likely that a lack of power due to a low site sample size was a contributing factor to the instances where this was not the case. The similarity of DTI and cognitive metrics across sites also suggests good comparability between the centres involved in this study. With respect to the wider literature however, the authors do note that DTI metrics have sometimes been shown to differ in magnitude between manufacturers, such as one paper where MD values were found to be systematically higher on Siemens vs. Philips scanners (this would not impact peak height of MD as used in the presented study, but could affect measures of MD centrality (48)). Another paper (49) examined reproducibility of whole brain MD peak height between a 1.5T and a 3T Siemens scanner in a sample size of 7 CADASIL patients, which achieved an intraclass correlation coefficient of .752 (indicating "good" reliability (50)). A further paper has found scanner upgrades to affect DTI after scanning CADASIL patients (51). These findings show caution should be used when combining DTI data from different manufacturers or when taking measurements over time, and future research may wish to take this into account in analyses. These considerations also highlight the importance of conducting multi-centre scanner calibration and standardisation of acquisition protocols prior to study commencement, as well as on-going quality control checks during the study duration in multi-centre research of this nature.

There were some limitations to this study. There were variable sample sizes across sites, meaning the influence of some centres is much stronger than others on our findings. In particular, having a greater number of participants scanned on non-Philips hardware would have provided more information about the comparability across scanners. The lack of data on inter-scanner reproducibility is also limiting and would have been valuable in more closely judging the sensitivity of these metrics across sites. It would also have been advantageous to acquire a field map with the DTI protocol so that corrections for susceptibility-induced

distortions could have been made. However registration to DTI space did appear good, so this is unlikely to have caused any major problems.

To conclude, in a multi-centre study we have shown that DTI metrics and lacune count correlate with cognition to a similar degree to that found in single centre studies. Our findings support the use of DTI as a surrogate marker of SVD in multi-centre studies.

Declarations of Interest

The authors report no declarations of interest.

Funding

The study was funded by a joint grant between the British Heart Foundation and the Stroke Association (TSA BHF 2010/01), and received additional support from the NIHR-funded Newcastle Biomedical Research Centre. H.S.M., G.A.F. & J.T.O. are supported by NIHR Senior Investigator awards. H.S.M. & J.T.O. are also supported by the Cambridge University Hospitals NIHR Comprehensive Biomedical Research Centre.

Author Contributions

Croall ID: Data analysis; Initial draft of manuscript

Lohner V: Data analysis; Initial draft of manuscript

Moynihan B: Data acquisition

Khan U: Data acquisition

Hassan A: Data acquisition

O'Brien JT: Obtained funding; Study design; Data acquisition

Morris RG: Obtained funding; Study design; Data acquisition; Data analysis

Tozer DJ: Data acquisition; Data analysis; Image analysis supervision;

Cambridge VC: Data acquisition

Harkness K: Data acquisition

Werring DJ: Data acquisition

Blamire AM: Obtained funding; Study design; Image analysis supervision; Data acquisition; Data analysis

Ford GA: Obtained funding; Study design; Data acquisition

Barrick TR: Obtained funding; Study design; Image analysis supervision; Data acquisition; Data analysis

Markus HS: Obtained funding; Study design; Data acquisition; Data analysis; Initial draft of manuscript; overall study supervision.

References

1. Pantoni L. Cerebral small vessel disease: from pathogenesis and clinical characteristics to therapeutic challenges. Lancet Neurol. 2010;9(7):689–701.
2. Román GC, Erkinjuntti T, Wallin A, Pantoni L, Chui HC. Subcortical ischaemic vascular dementia. Lancet Neurol. 2002;1(7):426–36.
3. 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(1):47–59.
4. Wardlaw JM, Smith EE, Biessels GJ, Cordonnier C, Fazekas F, Frayne R, et al. Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. Lancet Neurol. 2013;12(8):822–38.
5. Hollocks MJ, Brookes R, Morris RG, Markus HS, Banerjee S, Brookes RL, et al. Associations between the Brief Memory and Executive Test (BMET), Activities of Daily Living, and Quality of Life in Patients with Cerebral Small Vessel Disease. J Int Neuropsychol Soc. 2016;22(5):561–9.
6. Prins ND, van Dijk EJ, den Heijer T, 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(9):2034–41.
7. Lawrence AJ, Brookes RL, Zeestraten EA, Barrick TR, Morris RG, Markus HS. Pattern and Rate of Cognitive Decline in Cerebral Small Vessel Disease: A Prospective Study. Baron J-C, editor. PLoS One. 2015;10(8):e0135523.
8. 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(4):e61014.
9. Pantoni L. 2001-2011: A decade of the LADIS (leukoaraiosis and DISability) study: What have We learned about white matter changes and small-vessel disease? Cerebrovasc Dis. 2011;32(6):577–88.
10. Rosenberg G a, Wallin A, Wardlaw JM, Markus HS, Montaner J, Wolfson L, et al. Consensus statement for diagnosis of subcortical small vessel disease. J Cereb Blood Flow Metab. 2016;36(1):6–25.
11. Benjamin P, Zeestraten EA, Lambert C, Chis Ster IC, Williams OA, Lawrence AJ, et al. Progression of MRI markers in cerebral small vessel disease: sample size considerations for clinical trials. J Cereb Blood Flow Metab. 2016;36(1):228–40.
12. Van Norden AGW, De Laat KF, Van Dijk EJ, Van Uden IWM, Van Oudheusden LJB, Gons RAR, et al. Diffusion tensor imaging and cognition in cerebral small vessel disease. The RUN DMC study. Biochim Biophys Acta - Mol Basis Dis. 2012;1822(3):401–7.
13. Maillard P, Fletcher E, Harvey D, Carmichael O, Reed B, Mungas D, et al. White matter hyperintensity penumbra. Stroke. 2011;42(7):1917–22.

14. 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(7):1999–2005.
15. 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):e0147836.
16. 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.
17. De Guio F, Jouvent E, Biessels GJ, Black SE, Brayne C, Chen C, et al. Reproducibility and variability of quantitative magnetic resonance imaging markers in cerebral small vessel disease. J Cereb Blood Flow Metab. 2016;36(8):1319–37.
18. Fazekas F, Chawluk JB, Alavi A, Hurtig HI, Zimmerman RA. Mr Signal Abnormalities At 1.5-T in Alzheimer Dementia and Normal Aging. Am J Roentgenol. 1987;149(2):351–6.
19. Wechsler D. Manual for the Wechsler Adult Intelligence Scale – Fourth Edition. San Antonio: Pearson; 2012.
20. Reitan RM. The relation of the trail making test to organic brain damage. J Consult Psychol. 1955;19(5):393–4.
21. Delis D, Kaplan E, Kramer J. Delis-Kaplan Executive Function Scale (D-KEFS). San Antonio, TX: The Psychological Corporation; 2001.
22. Tombaugh TN, Kozak J, Rees L. Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. Arch Clin Neuropsychol. 1999;14(2):167–77.
23. Schmidt M. Rey Auditory and Verbal Learning Test: A handbook. Western Psychological Services, Los Angeles, CA; 1996.
24. Nelson HE, Willison JR. The Revised National Adult Reading Test - Test manual. Windsor, UK: NFER-Nelson; 1991.
25. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc. 2005;53(4):695–9.
26. Williams LS, Weinberger M, Harris LE, Clark DO, Biller J. Development of a stroke-specific quality of life scale. Stroke. 1999;30(7):1362–9.
27. EuroQoL Group. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy. 1990;16(3):199–208.
28. Mitrushina MN, Boone KB, D'Elia L. Handbook of normative data for neuropsychological assessment. 2nd ed. Oxford University Press; 2005.
29. 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 Mar;75(3):441–7.
30. Shimony JS, Sheline YI, Angelo GD, Epstein AA, Tammie LS, Mintun MA, et al. Diffuse Microstructural Abnormalities of Normal Appearing White Matter in Late Life Depression : a Diffusion Tensor Imaging Study. Biol Psychiatry. 2009;66(3):245–52.
31. Otsuka Y, Yamauchi H, Sawamoto N, Iseki K, Tomimoto H, Fukuyama H. Diffuse tract damage in the hemispheric deep white matter may correlate with global cognitive impairment and callosal atrophy in patients with extensive leukoaraiosis. Am J Neuroradiol. 2012;33(4):726–32.
32. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull. 1979;86(2):420–8.
33. Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA, et al. N4ITK:

improved N3 bias correction. IEEE Trans Med Imaging. 2010;29(6):1310–20.
34. Ashburner J, Friston KJ. Unified segmentation. Neuroimage. 2005;26(3):839–51.
35. Smith SM, Zhang Y, Jenkinson M, Chen J, Matthews PM, Federico A, et al. Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. Neuroimage. 2002;17(1):479–89.
36. 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(Suppl 1):S208-19.
37. Jenkinson M, Smith S. A global optimisation method for robust affine registration of brain images. Med Image Anal. 2001;5(2):143–56.
38. Moonen JE, Foster-Dingley JC, van den Berg-Huijsmans AA, de Ruijter W, de Craen AJ, van der Grond J van der MR. Influence of Small Vessel Disease and Microstructural Integrity on Neurocognitive Functioning in Older Individuals: The DANTE Study Leiden. Am J Neuroradiol. 2017;38(1):25–30.
39. Pasi M, Salvadori E, Poggesi A, Ciolli L, Del Bene A, Marini S, et al. White matter microstructural damage in small vessel disease is associated with montreal cognitive assessment but not with mini mental state examination performances: Vascular mild cognitive impairment tuscany study. Stroke. 2015;46(1):262–4.
40. Cao W, Wang Y, Dong Q, Chen X, Li Y, Zhou Y, et al. Deep microbleeds and periventricular white matter disintegrity are independent predictors of attention / executive dysfunction in non-dementia patients with small vessel. Int Psychogeriatr. 2016;1–11.
41. Vermeer SE, Prins ND, den Heijer T, Hofman A, Koudstaal PJ, Breteler MM. Silent brain infarcts and the risk of dementia and cognitive decline. N Engl J Med. 2003;348(13):1215–22.
42. van Uden IWM, Tuladhar AM, van der Holst HM, van Leijsen EMC, van Norden AGW, de Laat KF, et al. Diffusion tensor imaging of the hippocampus predicts the risk of dementia; the RUN DMC study. Hum Brain Mapp. 2016;37(1):327–37.
43. Jokinen H, Gouw AA, Madureira S, Ylikoski R, Van Straaten ECW, Van Der Flier WM, et al. Incident lacunes influence cognitive decline: The LADIS study. Neurology. 2011;76(22):1872–8.
44. Jokinen H, Schmidt R, Ropele S, Fazekas F, Gouw AA, Barkhof F, et al. Diffusion changes predict cognitive and functional outcome: The LADIS study. Ann Neurol. 2013;73(5):576–83.
45. 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.
46. Ofori E, Pasternak O, Planetta PJ, Burciu R, Snyder A, Febo M, et al. Increased free-water in the substantia nigra of Parkinson's disease: a single-site and multi-site study. Neurobiol Aging. 2015;36(2):1097–104.
47. Gregory S, Cole JH, Farmer RE, Rees EM, Roos RAC, Sprengelmeyer R, et al. Longitudinal Diffusion Tensor Imaging Shows Progressive Changes in White Matter in Huntington's Disease. J Huntingtons Dis. 2015;4(4):333–46.
48. Magnotta VA, Matsui JT, Liu D, Johnson HJ, Long JD, Bolster BD, et al. MultiCenter Reliability of Diffusion Tensor Imaging. Brain Connect. 2012;2(6):345–55.
49. 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(4):581–92.
50. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med. 2016;15(2):155–63.

51. Gunda B, Porcher R, Duering M, Guichard JP, Mawet J, Jouvent E, et al. ADC histograms from routine DWI for longitudinal studies in cerebral small vessel disease: A field study in CADASIL. PLoS One. 2014;9(5).

Table 1: **Baseline characteristics of the study population.** CABG= Coronary Artery Bypass Graft.

Demographic variable	**Mean (SD) / Number (%)**
Age, mean (SD) years	68.2 (9.07)
Male, n (%)	64 (58.7%)
Premorbid IQ	115.8 (8.12)
MoCA <26	54 (49.5%)
Systolic Blood pressure (mmHg)	150 (13)
Diastolic Blood pressure (mmHg)	85 (12)
Previous Stroke, n (%)	21 (19.3%)
Hypercholesterolaemia, n (%)	84 (77.1%)
Diabetes, n (%)	24 (22.0%)
Current Smokers, n (%)	16 (14.7%)
Former smokers, n (%)	40 (37.7%)
Angina, n (%)	7 (6.4%)
Myocardial infarction, CABG, or Coronary Angioplasty, n (%)	6 (5.5%)
Peripheral vascular disease, n (%)	2 (1.9%)
History of depression, n (%)	20 (18.3%)

(Other missing data not previously reported; Former smoker = 3; Peripheral vascular disease = 1)

Table 2. **Mean scores for key individual variables using all available data.** SSQoL=Stroke-Specific Quality of Life; NBV=Normalised Brain Volume; WMH=White Matter Hyperintensity; NAWM=Normal Appearing White Matter; WM=(all) White Matter; FA=Fractional Anisotropy; MD=Mean Diffusivity.

Variable	**Mean (SD), Range**
Cognitive / QoL Variables	
MoCA	24.9 (3.5), 11-30
SSQoL	190.6 (32.8), 93-244
EuroQoL	69.3 (19.1), 0-100
MRI Variables	
NBV (whole brain, ml)	1355.84 (107.70)
Grey matter normalised volume (ml)	714.49 (73.48)
White matter normalised volume (ml)	641.35 (70.39)
WMH volume (ml)	34.74 (22.27)
WMH lesion load (% brain)	3.41 (2.22)
Lacunes (number)	4.41 (4.73)
FA Height $x10^{-3}$ (NAWM / WM)	3.27 (.26) / 3.24 (.25)
MD Height $x10^{-2}$ (NAWM / WM)	1.42 (.21) / 1.33 (.23)
FA Value (NAWM / WM)	.320 (.042) / .311 (.047)
MD Value mm^2s^{-1} $x10^{-3}$ (NAWM / WM)	.761 (.040) / .762 (.040)
FA Median (NAWM / WM)	.342 (.026) / .335 (.028)
MD Median mm^2s^{-1} $x10^{-3}$ (NAWM / WM)	.774 (.039) / .787 (.044)

Table 3. **Results from linear regression "Simple Model" analyses. All numbers are standardised β-values (*p*-values). Significant relationships are shown in bold while the most strongly associated MR variable per outcome, per-category is underlined.** NBV=Normalised Brain Volume; WMH=White Matter Hyperintensity; NAWM=Normal Appearing White Matter; WM=(all) White Matter; FA=Fractional Anisotropy; MD=Mean Diffusivity.

MR Variable		Global Cog.	Proc. Speed	Mental Flex.	Verbal Fluency	Verbal Memory	MoCA	SSQoL	EuroQoL
Volume Measures	Whole NBV	**.330 (.003)**	**.361 (.002)**	**.286 (.016)**	.171 (.153)	**.245 (.041)**	**.421 (<.001)**	**.273 (.036)**	.200 (.115)
	Grey NBV	.167 (.109)	.109 (.321)	.111 (.312)	.147 (.177)	.177 (.109)	**.339 (.001)**	.084 (.485)	-.016 (.888)
	White NBV	**.199 (.036)**	**.283 (.004)**	**.202 (.043)**	.049 (.627)	.101 (.320)	.142 (.150)	.213 (.051)	**.225 (.034)**
WMH Measure	Lesion Load	**-.288 (.001)**	**-.312 (.001)**	**-.245 (.009)**	**-.240 (.011)**	-.132 (.170)	**-.196 (.035)**	**-.249 (.015)**	**-.248 (.013)**
Lacune Measure	N° Lacunes	**-.357 (<.001)**	**-.389 (<.001)**	**-.268 (.006)**	**-.286 (.003)**	**-.233 (.018)**	**-.333 (<.001)**	**-.323 (.002)**	-.195 (.062)
NAWM DTI Measures	FA Median	**.352 (<.001)**	**.247 (.009)**	**.338 (<.001)**	**.374 (<.001)**	.167 (.081)	**.332 (<.001)**	.196 (.058)	**.253 (.011)**
	MD Peak Height	**.267 (.005)**	.186 (.063)	**.241 (.016)**	**.275 (.006)**	.160 (.115)	**.262 (.007)**	.170 (.121)	**.244 (.021)**
WM DTI Measures	FA Median	**.371 (<.001)**	**.282 (.002)**	**.354 (<.001)**	**.375 (<.001)**	.174 (.067)	**.329 (<.001)**	**.213 (.037)**	**.267 (.007)**
	MD Peak Height	**.303 (.001)**	**.247 (.011)**	**.273 (.005)**	**.293 (.003)**	.162 (.101)	**.262 (.006)**	**.218 (.042)**	**.291 (.005)**

Table 4. **Results from linear regression "Complex Model" analyses. All numbers are standardised β-values (*p*-values) [95% standardised β confidence interval], with overall model significance being given on the bottom row. Models are separated into those which test NAWM and WM metrics horizontally. Significant associations are shown in bold.** WMH=White Matter Hyperintensity; NBV=Normalised Brain Volume; NAWM=Normal Appearing White Matter; WM=(all) White Matter; FA=Fractional Anisotropy; MD=Mean Diffusivity.

Tissue class model	MR Variable	Global Cog.	Proc. Speed	Mental Flex.	Verbal Fluency	Verbal Mem.	MOCA	SSQoL	EuroQoL
NAWM	Whole NBV	.134 (*.227*) [-.085 : .353]	.197 (*.098*) [-.037 : .432]	.112 (*.363*) [-.131 : .356]	-.041 (*.735*) [-.280 : .198]	.163 (*.213*) [-.095 : .422]	**.284 (*.014*) [.058 : .509]**	.137 (*.323*) [-.137 : .410]	-
	WM NBV	-	-	-	-	-	-	-	**.230 (*.030*) [.023 : .437]**
	WMH Lesion Load	-.029 (*.775*) [-.230 : .172]	-.126 (*.246*) [-.341 : .089]	-.009 (*.938*) [-.232 : .215]	.006 (*.954*) [-.212 : .225]	.033 (*.784*) [-.204 : .270]	.116 (*.270*) [-.091 : .323]	-.100 (*.428*) [-.351 : .150]	-.144 (*.315*) [-.354 : .125]
	N° Lacunes	**-.251 (*.006*) [-.429 : -.072]**	**-.287 (*.004*) [-.477 : .096]**	-.166 (*.099*) [-.365 : .032]	-.192 (*.057*) [-.390 : .006]	-.186 (*.082*) [-.397 : .024]	**-.247 (*.009*) [-.431 : -.063]**	**-.245 (*.031*) [-.467 : .-022]**	-.047 (*.672*) [-.266 : .172]
	FA Median	**.227 (*.023*) [.032 : .421]**	.038 (*.717*) [-.170 : .247]	**.253 (*.022*) [.037 : .470]**	**.333 (*.002*) [.121 : .546]**	.085 (*.463*) [-.145 : .315]	**.244 (*.018*) [.043 : .445]**	.032 (*.796*) [-.211 : .275]	.196 (*.099*) [-.037 : .428]
	MD Peak Height	**-**	-	**-**	**-**	-	-	-	**-**
	Model sig. (p value, Adj. R^2)	***<.001*, .429**	***<.001*, .334**	***<.001*, .292**	***<.001*, .317**	***.001*, .202**	***<.001*, .392**	***.029*, .108**	***.004*, .164**
WM	Whole NBV	.131 (*.236*) [-.087 : .349]	.194 (*.103*) [-.040 : .428]	.107 (*.380*) [-.135 : .350]	-.041 (*.730*) [.280 : .197]	.162 (*.215*) [-.096 : .421]	**.284 (*.014*) [.058 : .509]**	.139 (*.316*) [-.135 : .412]	-
	WM NBV	-	-	-	-	-	-	-	**.225 (*.033*) [.018 : .432]**
	WMH Lesion Load	.025 (*.882*) [-.195 : .245]	-.110 (*.359*) [-.346 : .127]	.055 (*.655*) [-.189 : .299]	.075 (*.536*) [-.165 : .315]	.052 (*.693*) [-.209 : .313]	.164 (*.155*) [-.063 : .392]	-.100 (*.474*) [-.376 : .176]	-.081 (*.545*) [-.347 : .184]
	N° Lacunes	**-.248 (*.007*) [-.425 : -.70]**	**-.285 (*.004*) [-.476 : -.*095*]**	-.163 (*.104*) [-.360 : .034]	-.190 (*.059*) [-.388 : .007]	-.186 (*.083*) [-.396 : .025]	**-.245 (*.009*) [-.429 : -.062]**	**-.245 (*.031*) [-.468 : -.023]**	-.047 (*.669*) [-.267 : .172]
	FA Median	**.268 (*.016*) [.052 : .484]**	.058 (*.621*) [-.174 : .290]	**.306 (*.013*) [.066 : .546]**	**.376 (*.002*) [.140 : .612]**	.099 (*.445*) [-.157 : .355]	**.273 (*.017*) [.049 : .497]**	.026 (*.849*) [-.245 : .297]	-
	MD Peak Height	**-**	**-**	**-**	**-**	-	-	-	.209 (*.112*) [-.050 : .468]
	Model sig. (p value, Adj. R^2)	***<.001*, .433**	***<.001* .345**	***<.001*, .299**	***<.001*, .319**	***.001*, .202**	***<.001*, .392**	***.029*, .108**	***.004*, .162**

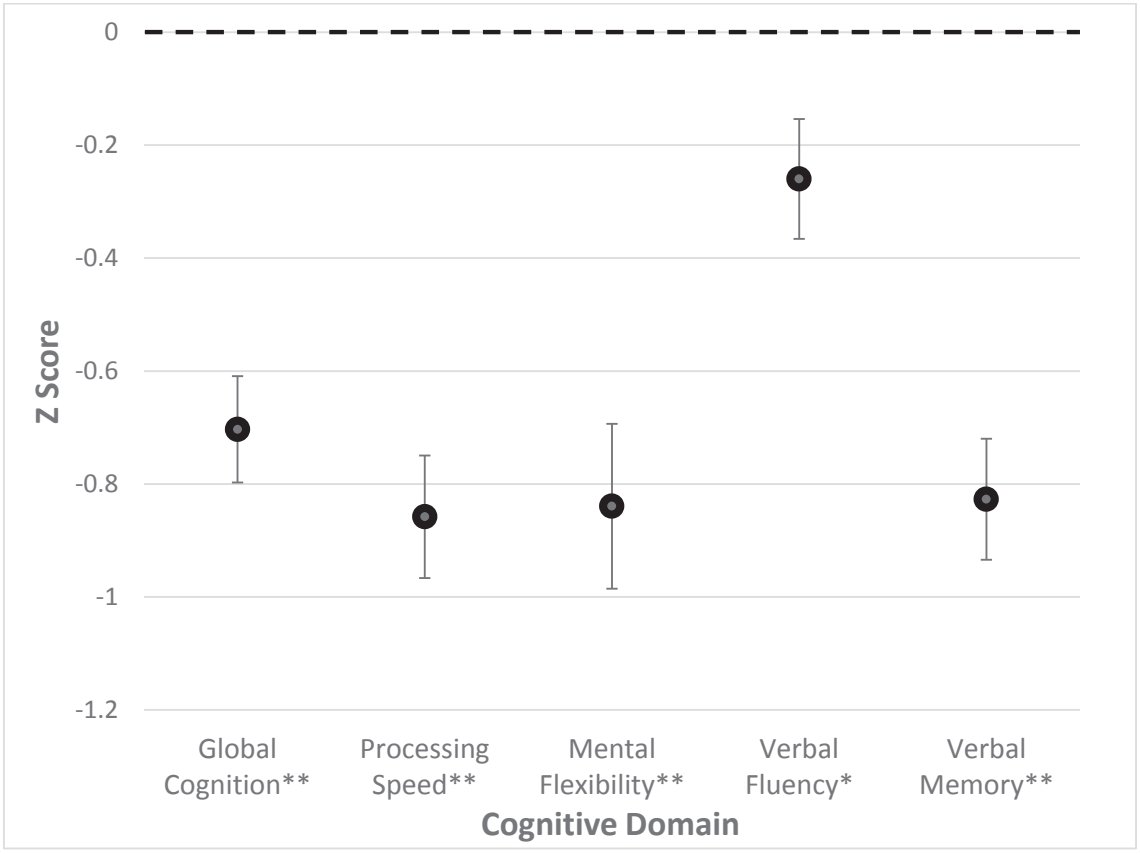


Figure 1. **Cognitive profile of the SVD patient group. This figure shows average, age-matched z scores for cognitive indices. Error bars represent +/- 1 standard error of the mean. Index score significantly different from zero: **=p<.001, *=p<.005.**

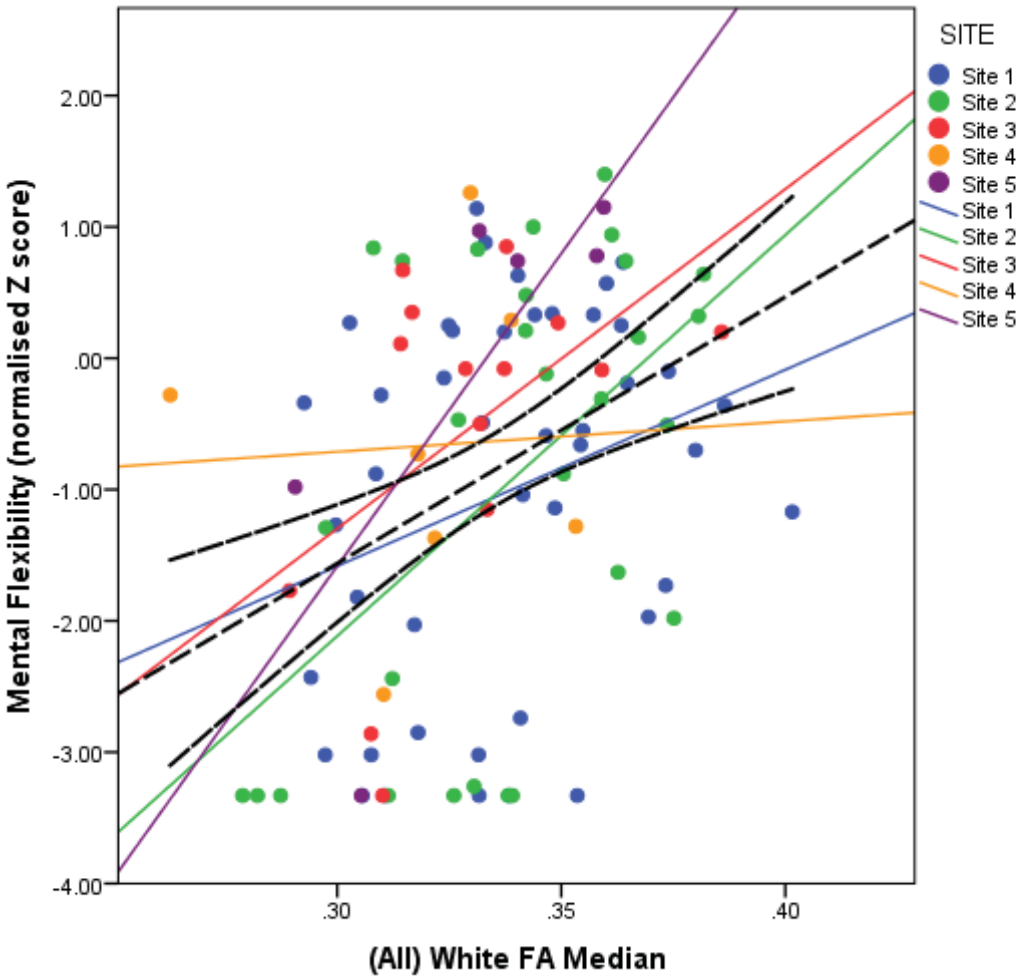


Figure 2. **A scatterplot showing the relationship between WM FA median and Mental Flexibility, stratified by study site. In addition to individual site regression lines, the regression line for the total is also included with accompanying 95% CI limits (black, dashed line).** FA = Fractional Anisotropy; WM = (all) White Matter.

Supplementary Material

Supplementary Analysis

To assess variation across individual sites, analyses were conducted on data from each site individually. Comparison of DTI data by one-way ANOVA between sites gave non-significant results for both FA median (p= .424) and MD peak height (p= .148).

Performance in all outcome measures was also compared between sites by Kruskal-Wallis (for Mental Flexibility, which was non-normally distributed) and one-way ANOVA (for all other metrics). No findings were significant: Global Cognition (p= .661), Processing Speed (p= .437), Mental Flexibility (p= .229), Verbal Fluency (p= .827), Verbal Memory (p= .641), MoCA (p= .678), SSQoL (p= .192), EuroQoL (p= .272).

To assess the strength of associations with cognition on DTI measures obtained from individual sites, "Complex Model" analyses were repeated in these sites individually. Data from Sites 4, 5 and 6 were not included due to small sample sizes. Complex models selected for this were ones where the cognitive domain had held a significant association with a DTI metric in the main study analysis. The WM DTI model was chosen for replication over the accompanying NAWM model due to the former always holding the larger association. Complex models were therefore repeated for Global Cognition, Mental Flexibility, and Verbal Fluency. The full findings are displayed in Supplementary Table 3. Briefly, this shows some variability in the strengths of β-values between sites; notably Site 2 appears to hold relatively strong associations (smallest / largest β-value, ignoring direction: .402 / .510) and Site 3 relatively weak ones (smallest / largest β-value: .047 / -.195). All β-values from Sites 1 and 2, and the Global Cognition model for Site 3, fall within the 95% CI reported from the main study analysis for the same comparison (see Table 4). However the β-values for Mental Flexibility and Verbal Fluency models from Site 3 are lower.

These relationships are further visualised in Figure 1, and Supplementary Figures 1 and 2, which respectively show the relationships between WM FA median and Mental Flexibility, Global Cognition and Verbal Fluency, stratified by site while also including Sites 4 and 5. The 95% CI around the regression line for the total fit (i.e. ignoring site) is also shown. These again show some variability. Notably, with respect to the total regression line, Figure 2 shows a "weak" fit for Site 4, and Supplementary Figures 1 and 2 show a "weak" fit for Site 3 (i.e. the individual site regression lines lie at least partially outside the total CI limits in a manner indicating them to have a flatter slope). All other site lines in all Figures show "good" fit by

falling either completely within the total CI limits, or outside of them in a manner indicating them to have a steeper slope.

Considering these Figures and the repeated Complex Model analyses, this is suggestive that while a minority of comparisons at the individual site level do not appear representative of the significant cognitive domain associations reported in Table 4, the majority of site data *does* contribute to the main study finding. Given that neither the cognitive scores or DTI metrics differ between sites, that Sites 3 and 4 only have non-representative associations in some (but not all) comparisons, and that these sites have low sample sizes, it is likely this is due to a lack of power.

Supplementary Table 1. **An overview of the exact scanners and sequence parameters used at each site.** FOV=Field of View; FLAIR=Fluid Attenuated Inversion Recovery

Site (N)	**Site 1 (48)**	**Site 2 (29)**	**Site 3 (14)**	**Site 4 (11)**	**Site 5 (6)**	**Site 6 (1)**
3T Scanner(s)	**Philips Achieva TX**	**Philips Achieva** (N=24), **Philips Achieva TX** (N=5)	**Siemens Verio** (N=8), **Siemens Magnetom Prisma^{fit}** (N=6)	**Philips Achieva TX**	**Philips Ingenia**	**Siemens Prisma**
Axial DTI (32 diffusion weighted gradient directions at b-value = 1000 s mm^{-2}, Isotropic voxel resolution 2mm^3)	**DwiSE** **TR =** 6850ms **TE** = 75ms **In-plane FOV:** 224×224mm^2 **N^o slices:** 60 **N^o b0s:** 8 **Max. Gradient Strength**: 80mT/m **Parallel Imaging Factor**: 3 **N^o headcoil channels:** 8	**DwiSE** **TR =** 6850ms **TE** = 75ms **In-plane FOV:** 224×224mm^2 **N^o slices:** 60 **N^o b0s:** 8 **Max. Gradient Strength**: 80mT/m **Parallel Imaging Factor**: 3 **N^o headcoil channels:** 8	**Twice-refocussed** TR = 11500ms TE = 93ms **In-plane FOV:** 192×192mm^2 **N^o slices:** 75 **N^o b0s:** 2 **Max. Gradient Strength (Verio/Prisma)**: 45/80mT/m **Parallel Imaging Factor**: 2 **N^o headcoil channels:** 32	**DwiSE** **TR =** 6850ms **TE** = 75ms **In-plane FOV:** 224×224mm^2 **N^o slices:** 60 **N^o b0s:** 8 **Max. Gradient Strength**: 80mT/m **Parallel Imaging Factor**: 3 **N^o headcoil channels:** 8	**DwiSE** **TR =** 9100ms **TE** = 82ms **In-plane FOV:** 224×224mm^2 **N^o slices:** 60 **N^o b0s:** 8 **Max. Gradient Strength**: 45mT/m **Parallel Imaging Factor**: 3 **N^o headcoil channels:** 15	**Twice-refocussed** TR = 9500ms TE = 93ms **In-plane FOV:** 192×192mm^2 **N^o slices:** 81 **N^o b0s:** 2 **Max. Gradient Strength**: 40mT/m **Parallel Imaging Factor**: 2 **N^o headcoil channels:** 12
Sagittal 3D T1-weighted (Isotropic voxel resolution 1mm^3)	**Turbo Field Echo** **TR =** 8.27ms **TE** = 4.61ms **Field of View** 240^2×170mm^3	**Turbo Field Echo** **TR =** 9.81ms **TE** = 4.60ms **Field of View** 240^2×170mm^3	**MP RAGE** **TR =** 2200ms **TE** = 2.97ms **FOV:** 256^2×208mm^3 **Inversion Time (TI) =** 900ms	**Turbo Field Echo** **TR =** 11ms **TE** = 4.61ms **Field of View** 240^2×170mm^3	**Turbo Field Echo** **TR =** 8.53ms **TE** = 4.61ms **Field of View** 240^2×170mm^3	**MP RAGE** **TR =** 2200ms **TE** = 2.94ms **FOV:** 256^2×208mm^3 **Inversion Time (TI) =** 900ms

Axial T2*-weighted	**Fast Field Echo** **TR = 1800ms** **TE = 20ms** **Voxel size:** $0.5^2 \times 3mm^3$ **In-plane FOV:** $240 \times 240mm^2$ **N° slices:** 51	**Fast Field Echo** **TR = 1800ms** **TE = 20ms** **Voxel size:** $0.5^2 \times 3mm^3$ **In-plane FOV:** $240 \times 240mm^2$ **N° slices:** 51	**Spoiled Gradient Echo** **TR = 1570ms** **TE = 20ms** **Voxel size:** $0.94^2 x 3mm^3$ **In-plane FOV:** 195x240mm **N° slices:** 50	**Fast Field Echo** **TR = 1800ms** **TE = 20ms** **Voxel size:** $0.5^2 \times 3mm^3$ **In-plane FOV:** $240 \times 240mm^2$ **N° slices:** 51	**Fast Field Echo** **TR = 1800ms** **TE = 20ms** **Voxel size:** $0.54^2 \times 3mm^3$ **In-plane FOV:** $240 \times 240mm^2$ **N° slices:** 51	**Spoiled Gradient Echo** **TR = 1570ms** **TE = 20.7ms** **Voxel size:** $0.94^2 x 3mm^3$ **In-plane FOV:** 195x240mm **N° slices:** 50
Axial FLAIR **(**Inversion time = 2800ms**)**	**Fluid Attenuated Inversion Recovery (FLAIR)** **TR =** 11000ms **TE** = 120ms **In-plane FOV:** $230 \times 230mm^2$ **Voxel size:** $0.48^2 \times 3mm^3$ **N° slices:** 57	**FLAIR** **TR =** 11000ms **TE** = 120ms **In-plane FOV:** $230 \times 230mm^2$ **Voxel size:** $0.48^2 \times 3mm^3$ **N° slices:** 57	**Turbo Inversion Recovery** **TR =** 8000ms **TE** = 124ms **In-plane FOV:** 208x230mm **Voxel size:** $0.45^2 x 3mm^3$ **N° slices:** 60	**FLAIR** **TR =** 11000ms **TE** = 120ms **In-plane FOV:** $230 \times 230mm^2$ **Voxel size:** $0.48^2 \times 3mm^3$ **N° slices:** 57	**FLAIR** **TR =** 11000ms **TE** = 120ms **In-plane FOV:** $230 \times 230mm^2$ **Voxel size:** $0.48^2 \times 3mm^3$ **N° slices:** 57	**Turbo Inversion Recovery** **TR =** 8000ms **TE** = 121ms **In-plane FOV:** 208x230mm **Voxel size:** $0.45^2 x 3mm^3$ **N° slices:** 60

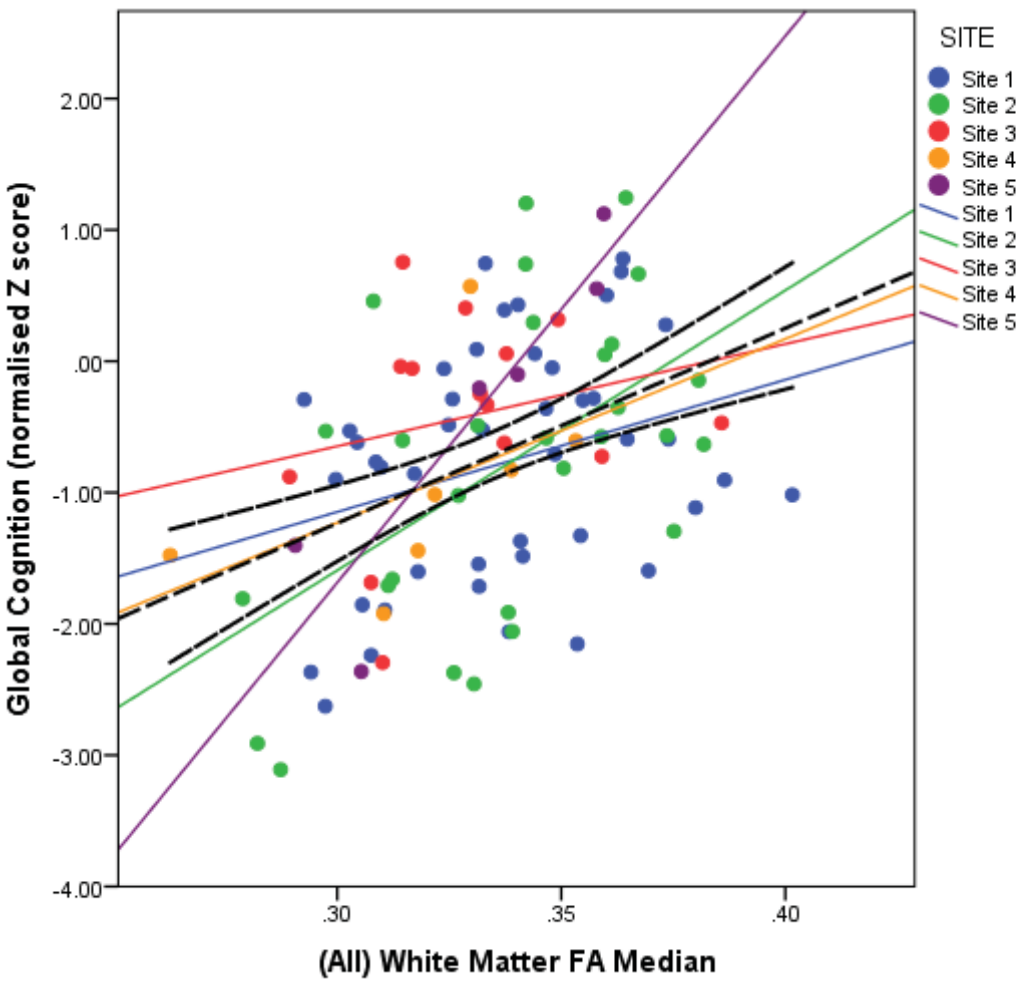
Supplementary Table 2. **A summary of results from previous single-centre studies looking at the relationship between DTI metrics and cognition in SVD for comparison with the presented study (included at the top). 95% CIs for the β's have been reported where available. From the present study, "Mental Flexibility" has been renamed Executive Functioning, while Verbal Fluency has been kept as a separate domain. This is in order to allow better comparison with previous literature.** CI = Confidence Interval; FA = Fractional Anisotropy; HDWM = Hemispheric Deep White Matter; MD = Mean Diffusivity; NART = National Adult Reading Test; NAWM = Normal Appearing White Matter; NBV = Normalised Brain Volume; PV = Periventricular; RAVLT = Rey Auditory Verbal Learning Test; TMT = Trail Marking Test; WM = (all) White Matter; WMH = White Matter Hyperintensity

Study	Study Cohort	N	DTI Metric	Cognitive Measure	Model and Additional Variables	Finding
The Presented One	Lacunar infarcts and confluent WMH	100	WM MD Median	**Global Cognition** (composite score)	**Multiple Linear Regression**: Age, Gender, Site, NART IQ, NBV, WMH Lesion Load, Lacune N^{o}	**β= .268**, *p= .016* **CI = .052 : .484**
			NAWM FA Peak Height	**Processing Speed** (TMT-A, WAIS digit symbol)		**β= .058**, *p= .621* **CI = -.174 : .290**
			WM MD Median	**Executive Functioning** (i.e. "*Mental Flexibility"*; TMT-B)		**β= .306**, *p= .013* **CI = .066 : .546**
			WM MD Median	**Verbal Memory** (RAVLT)		**β= .099**, *p= .445* **CI = -.157 : .355**
			WM FA Peak Height	**MoCA**		**β= .273**, *p= .017* **CI = .049 : .497**
		99	WM MD Median	**Verbal Fluency** (Verbal Fluency Task)		**β= .376**, *p= .002* **CI = .140 : .612**
(29)	Lacunar infarcts and confluent WMH	36	NAWM MD Mean	**Executive Function** (Wisconsin Card Sorting Task errors)	**Multiple Linear Regression**: Age, Gender, Brain Volume, T1 & T2 Lesion Load	**β= -.41**, *p= .046*
(30)	Vascular risk factors with depression and WMH	67	NAWM Prefrontal MD Mean	**Processing Speed** (Digit Symbol Substitution, Stroop colour naming subset task, TMT-A)	**Partial Correlation**: Age, Gender, WMH Lesion Volume	**r= −.27**, *p= .034*
(31)	Lacunar infarcts and confluent WMH	24	HDWM Mean FA	**Verbal Fluency** (Verbal Fluency Task)	**Multiple Linear Regression**: Age, Gender, WM NBV, WMH volume	**β= .56**, *p= .006*
			HDWM Mean MD	**Verbal Memory** (Wechsler Memory Scale Revised)		**β= -.86**, *p= <.002*

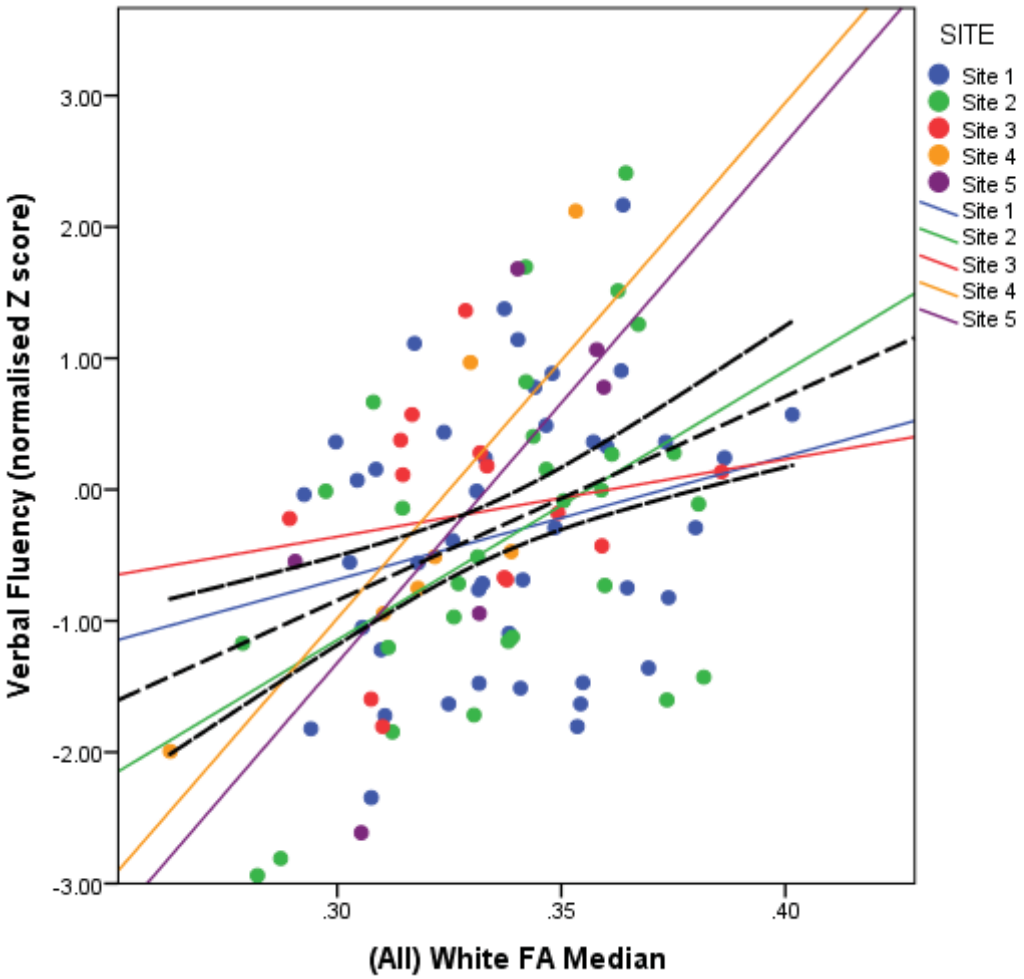
(12)	RUN DMC	503	NAWM MD Mean	**Global Cognition** (composite score)	**Multiple Linear Regression**: Age, Gender, Education, Depressive symptoms, NBV, Lacune N^{o}, WMH volume	**β= -.18,** *p= <.01* **CI = -.22 to .06**
				Verbal Memory (RAVLT)		**β= -.18**, *p= <.01* **CI = -.28 to -.06**
				Verbal Fluency (Verbal Fluency Task)		**β= -.11**, *p= <.05* **CI = -.21 to -.02**
				Executive Function (Stroop task 3)		**β= -.10**, *p= <.05* **CI = -.16 to -.06**
				Processing Speed (Digit Symbol Substitution, Paper-Pencil Memory Scanning, Stroop reading subset task)		**β= -.18**, *p= <.01* **CI = -.24 to -.06**
(8)	SCANS	115	NAWM RD Peak Height	**Executive Function** (TMT-B, Verbal Fluency, Modified Wisconsin Card Sorting)	**Multiple Linear Regression**: Age, Gender, NART IQ, NBV, WMH Lesion Load, Microbleed N^{o}, Lacune N^{o}	**β= -.21**, *p= .046*
			NAWM MD Peak Height	**Processing Speed** (Speed of Information Processing, Digit Symbol Substitution, Grooved Pegboard Task)		**β= -.085**, *p= .41*
(39)	VMCI-Tuscany	76	WM Median MD	**MoCA**	**Partial Correlation:** Age, Gender, Education, WMH rating, Global & Temporal lobe atrophy ratings	**r= −.28**, *p= .023*
(38)	DANTE Study Leiden	195	WM MD Mean	**Global Cognition** (composite score)	**Multiple Linear Regression**: Age, Gender, Education, NBV, WMH Volume, Microbleed N^{o}, Lacune N^{o}	**β= -.18**, *p= .08* **CI = -.38 to .02**
			WM RD Mean	**Processing Speed** (Letter-digit substitution)		**β= -.14**, *p= .17* **CI = -.33 to .06**
			WM FA Mean	**Executive Function** (interference score from abbreviated Stroop, TMT; B minus A)		**β= .22**, *p= .01* **CI = .05 to .39**
(40)	Lacunar infarcts and confluent WMH	55	PV MD Mean	**Executive Function** (TMT, Stroop, Category Fluency)	**Multiple Linear Regression (stepwise):** Age, Gender, Education, Depressive State (binary), Hypertension (binary), NBV, Microbleed N^{o}	**β= -.457**, *p= <.01*
			PV MD Mean	**Verbal Memory** (RAVLT)		**β= -.314**, *p= .02*

Supplementary Table 3. **Results of "Complex Model" analyses conducted in each site. For clarity, only the association (i.e. standardised β-values, with accompanying 95% CI) of the included DTI metric is reported. *p* values are not reported as varying sample sizes would have a large effect on these. All comparisons were made using WM FA median.** WM=(all) White Matter; FA=Fractional Anisotropy.

Site	**Global Cognition**	**Mental Flexibility**	**Verbal Fluency**
Site 1 (N= 48) β-value with 95% CI	.126 [-.280 : .532]	.218 [-.225 : .662]	.255 [-.182 : .692]
Site 2 (N= 29) β-value with 95% CI	.402 [.077 : .727]	.510 [.088 : .931]	.417 [.046 : .788]
Site 3 (N= 14) β-value with 95% CI	-.195 [-.848 : .457]	.047 [-.442 : 535]	.063 [-.914 : 1.039]



Supplementary Figure 1. **A scatterplot showing the relationship between WM FA median and Global Cognition, stratified by study site. In addition to individual site regression lines, the regression line for the total is also included with accompanying 95% CI limits (black, dashed line).** FA = Fractional Anisotropy; WM = (all) White Matter.



Supplementary Figure 2. **A scatterplot showing the relationship between WM FA median and Verbal Fluency, stratified by study site. In addition to individual site regression lines, the regression line for the total is also included with accompanying 95% CI limits (black, dashed line).** FA = Fractional Anisotropy; WM = (all) White Matter.