

**Evaluating Magnetic Resonance Imaging and Serum Markers
As Surrogate Endpoints for Clinical Trials in
Cerebral Small Vessel Disease**

Marco Tiber Egle

Hughes Hall

Department of Clinical Neuroscience

University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy
December 2020

Declaration

This thesis is the result of my own work and includes nothing, which is the outcome of work done in collaboration except as declared in the preface and specified in the text.

It is not substantially the same as any work that has already been submitted before for any degree or other qualification except as declared in the preface and specified in the text.

It does not exceed the prescribed word limit for the Clinical Medicine and Clinical Veterinary Medicine Degree Committee.

The clinical and MRI data used in this thesis was previously collected by others under the guidance of Prof. Hugh Markus (University of Cambridge), Prof. Frank-Erik de Leeuw (University of Nijmegen), Prof. Marco Duering (LMU Munich), Prof. Martin Dichgans (LMU Munich), Prof. Denis Chen (National University of Singapore), Prof. Reinhold Schmidt (University of Graz), Prof. Emmanuel Stamatakis (University of Cambridge), Dr. Jonathan Coles (University of Cambridge) and others. I helped processing the clinical data in the multicenter trial PRESERVE. Throughout all chapters described herein I designed the experiments and conducted the statistical analysis myself. Some MRI data was already analyzed for Hugh Markus and other principal investigators, such as white matter hyperintensity volume, brain volume, lacune count and cerebral microbleeds count, and some of the DTI data prior to the start of my PhD studentship. I analyzed the markers DSEG and PSMD of certain cohorts in Chapter 4 and the markers PSMD, DSEG, Geff in all 3 data sets in Chapter 5. The neurofilament assay analysis was done at the University of Basel under the guidance of Prof. Jens Kuhle and Prof. Nils Peters.

Abstract

Thesis title:

**Evaluating Magnetic Resonance Imaging and Serum Markers As
Surrogate Endpoints for Clinical Trials in Cerebral Small Vessel Disease**

Author:

Marco Tiber Egle

Cerebral small vessel disease (SVD) causes a quarter of all strokes and is the most common pathology underlying vascular cognitive impairment and dementia. White matter hyperintensities, lacunar infarcts, cerebral microbleeds and brain atrophy are characteristic features on conventional magnetic resonance imaging (MRI). More sensitive methods such as diffusion tensor imaging (DTI) show white matter ultrastructural damage and white matter tract disruption. As the vessel's pathology is largely unknown and is difficult to image and as there is a low incidence of clinical events such as lacunar stroke and dementia conversion, imaging markers have been proposed as surrogate marker for phase-2 clinical trials. Results have shown that imaging markers such as DTI and some conventional MRI measures are sensitive to change over a few years and are further associated with later clinical outcome events such as dementia conversion. Recently there has also been growing interest in circulating biomarkers such as serum neurofilament light chain as they are easier to obtain from patients than imaging markers and as they can be centrally computed also in multicenter trials.

The thesis aim is to evaluate imaging as well as circulating biomarkers as surrogate endpoints for clinical trials in SVD. Findings in Chapter 2 show that changes in conventional MRI markers may be more robust than in DTI markers for multicenter studies and therefore more suitable as a surrogate marker in this type of trial design. Evidence in Chapter 3 further indicates that DTI as a surrogate endpoint may depend on the disease severity and that conventional MRI such as brain atrophy and lacune incidences may increase the predictive accuracy for dementia conversion in severe SVD but not in mild

SVD or mild cognitive impairment (MCI). Findings in Chapter 4 demonstrate that recently developed automatic or semi-automatic DTI markers may be promising surrogate markers for a phase-2 clinical trial and may require the lowest sample size in sporadic SVD. Chapter 5 tests the reproducibility of the recently developed markers both between scanner types and over time. Results showed that DTI markers' reproducibility is only moderate between different scanner types. Findings of Chapter 6 show that the circulating biomarker serum neurofilament light chain may not be a suitable surrogate marker in SVD but may be used for selecting a SVD patient group with a higher risk for dementia conversion. All these findings indicate that automatic and semiautomatic DTI markers may be best suited as surrogate endpoint in a single-center clinical trial involving patients with severe sporadic SVD. The neurofilament light chain marker together with clinical markers as well as combined imaging marker scores may be employed to select severe SVD patients who have a higher risk for dementia conversion.

Acknowledgements

This work would have not been possible without the continuous support and encouragement I received.

To Hugh Markus, Dan Tozer and all members of the stroke research group, who I share many great memories with and who certainly made me a better scientist. Many thanks for this very nice and valuable experience.

To my parents Marie-Luise and Ulrich who always strongly believed in my academic abilities in good and difficult times and who I can rely on anytime without any doubts. They are the anchors in my life and will always be. To my brother Fabio and my sister Marlen who I share many wonderful memories with and who have challenged me throughout my life in order to make me a better person. Our lovely dog Millie is of course not to forget.

To Ho Yan whose strong emotional and academic support let me finish this work and let me climb this tall mountain. Without her support, this work would have not been possible.

To my grandparents Inge and Felix, who have been central in my life. They remind me that science has to give practical answers to complex problems and works best when it is inspired by the philosophy of humanism. To my grandparents Hilde and Paul, who would have been very proud of this academic work. Vascular dementia is such a terrible clinical outcome- I hope that there will be a treatment one day that will help future generations.

To my friends and mentors in the Netherlands, Germany and around the world, Juan Manuel Perafan, Katharina Hilger, Johanna Ehrenstein, Nausikaa Reimers, Lisa Marohn, Felicia Fehrentz, Hanna Botz, Timothy Sunday, Dorél Dobocan, Rita Romano Adao, Marloes Hellwich, Hasan Sari, Emily Chiang, Robert Wriedt, Defne Alfandari, Camilla Ascanelli, Lorenzo Hoppus, Michael Lerch, Iga Anna Bishop, Ahmed Abd El Baky, Ru Wang, Harry Ayton, Dewi Safitri, Pontus Leander, Brigit Toebes, Katrina Perehudoff, Nina Hansen, Rik Peters & Mark Nieuwenstein.

To the many inspiring people I met in Cambridge who gave me a very unique academic and emotional experience

Table of Contents

CHAPTER 1	13
1. PATHOGENESIS OF SMALL VESSEL DISEASE	13
1.1. ANATOMY.....	13
1.2 CEREBRAL HYPOPERFUSION.....	13
1.3 INCREASED BLOOD BRAIN BARRIER PERMEABILITY.....	14
2 SUBTYPES IN SVD	15
2.1 SPORADIC SVD.....	15
2.2 MONOGENETIC SVD (CADASIL).....	15
3. CLINICAL FEATURES OF SVD	16
3.1 COGNITIVE SYMPTOMS.....	16
3.2 PSYCHIATRIC & MOTOR SYMPTOMS.....	17
4. WHY ARE SURROGATE MARKERS NECESSARY FOR CLINICAL STUDIES IN SVD?	17
5. SURROGATE MARKERS	20
5.1 MRI MARKERS.....	20
5.2 BIOLOGICAL MARKERS IN SVD.....	29
6. CONCLUSION	32
7. AIMS OF THE THESIS	34
CHAPTER 2	35
1. INTRODUCTION	35
2. METHOD	36
2.1 PATIENTS.....	36
2.2 TRIAL DESIGN.....	36
2.3 OUTCOME MEASURES AND FOLLOW-UP ASSESSMENTS.....	37
2.4 ADVERSE EVENT RECORDING.....	38
2.5 MRI ACQUISITION.....	38
2.6 MRI ANALYSIS.....	38
2.7 COGNITIVE TESTING.....	44
2.8 STATISTICAL ANALYSIS.....	45
3. RESULTS	47
3.1 PATIENTS.....	47
3.2 BLINDED DATA.....	47
3.3 INTENTION TO TREAT ANALYSIS.....	49
3.4 PER PROTOCOL ANALYSIS.....	58
3.5 RELATIONSHIP BETWEEN CHANGE IN BP AND CHANGE IN IMAGING PARAMETERS.....	60
3.6 ADVERSE EVENTS.....	63
4. DISCUSSION	63
CHAPTER 3	65
1. INTRODUCTION	65
2. METHODS	65
2.1 PARTICIPANTS.....	65
2.2 MRI PROCESSING.....	74
2.3 STATISTICAL ANALYSIS.....	76
3. RESULTS	78
3.1 CROSS-SECTIONAL ASSOCIATION.....	78
3.2 LONGITUDINAL ASSOCIATION WITH BASELINE IMAGING MARKER.....	84
3.3 LONGITUDINAL ASSOCIATION WITH CHANGE IN IMAGING MARKER.....	84
4. DISCUSSION	91
CHAPTER 4	95

1. INTRODUCTION	95
2. METHODS	97
2.1 PARTICIPANTS	97
2.2 DTI IMAGING MEASURES	98
2.3 COGNITION AND DEMENTIA	103
2.4 STATISTICAL ANALYSIS.....	103
3. RESULTS	106
3.1 PCA BASELINE.....	106
3.2 CROSS-SECTIONAL RESULTS.....	106
3.3 LONGITUDINAL RESULTS.....	112
4. DISCUSSION.....	127
<u>CHAPTER 5</u>	<u>129</u>
1. INTRODUCTION	129
2. METHOD	130
2.1 PARTICIPANTS	130
2.2 MRI ACQUISITION.....	131
2.3 MRI ANALYSIS	132
2.4 STATISTICAL ANALYSIS.....	132
3. RESULTS	133
3.1 BETWEEN-SCANNER REPRODUCIBILITY IN BiOCOG.....	133
3.2 WITHIN- AND BETWEEN REPRODUCIBILITY IN WBIC DTI STUDY.....	133
3.3 BETWEEN-VISIT REPRODUCIBILITY IN THE DTI-NETWORK STUDY.....	134
3.4 BETWEEN-VISIT REPRODUCIBILITY AND AGE.....	134
4. DISCUSSION.....	146
<u>CHAPTER 6</u>	<u>149</u>
1. INTRODUCTION	149
2. METHODS	151
2.1 PATIENTS.....	151
2.2 SERUM NEUROFILAMENT LIGHT CHAIN.....	152
2.3 MRI ACQUISITION.....	152
2.4 CONVENTIONAL MRI MARKERS	153
2.5 DIFFUSION TENSOR IMAGING ANALYSIS.....	153
2.6 COGNITION AND DEMENTIA DIAGNOSIS.....	153
2.7 DISABILITY.....	153
2.8 STATISTICAL ANALYSIS.....	154
3. RESULTS	155
3.1 CROSS-SECTIONAL ANALYSIS.....	155
3.2 LONGITUDINAL ANALYSIS.....	156
4. DISCUSSION.....	169
<u>CHAPTER 7</u>	<u>172</u>
1. SUMMARY OF THE FINDINGS	172
2. FIVE RECOMMENDATIONS WHEN PLANNING FUTURE PHASE-2 CLINICAL TRIALS	194
3. HOW WOULD A FUTURE PHASE-2 CLINICAL TRIAL IN SVD LOOK LIKE?	197
4. CONCLUSION	200
<u>REFERENCE</u>	<u>201</u>

List of Abbreviations

AD	Axial diffusivity
AD	Alzheimer's disease
ADC	Apparent diffusion coefficient
Adj. R ²	Adjusted explained variance
AEs	Adverse events
AIC	Akaike information criterion
ANCOVA	Analysis of covariance analysis
ANOVA	Analysis of variance analysis
AUC	Area under the curve
BBB	Blood brain barrier
BOMBS	Brain Observer MicroBleed Scale
BP	Blood pressure
BV	Brain volume
CAA	Cerebral amyloid angiopathy
CADASIL	Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy
CBF	Cerebral blood flow
CMB	Cerebral microbleeds
CNS	Central nervous system
CRP	C reactive protein
CSF	Cerebral spinal fluid
CT	Computed tomography
DSEG	Diffusion tensor image segmentation
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders Version 4
DSM-V	Diagnostic and Statistical Manual of Mental Disorders Version 5
DTI	Diffusion tensor imaging
EF	Executive function
FA	Fractional anisotropy
Geff	Global efficiency network measure
Global	Global cognition
GM	Grey Matter
IADL	Instrumental activities of daily living
ITT	Intention-to-treat analysis
MARS	Microbleeds Anatomic Rating Scale
MCI	Mild cognitive impairment
MD	Mean diffusivity
MMSE	Mini-Mental State Examination
MRI	Magnetic resonance imaging
mRS	Modified Rankin score
NAWM	Normal-appearing white matter
NBV	Normalized brain volume
NfL	Serum neurofilament light chain
OPTIMAL	OPtimising mulTImodal MRI markers for use as surrogate markers in trails of Vascular Cognitive Impairment due to cerebrAl small vessel disease

PC1	First principal component score
PCA	Principal component analysis
PD	Parkinson's disease
PH	Normalized peak height of the histogram
pkval	Peak value of the histogram
PP	Per-protocol analysis
PRESERVE	How intensively should we treat blood PRESSure in established cEREbral small VEssel disease
PS	Processing speed
PSMD	Peak width of skeletonized mean diffusivity
PVS	Perivascular spaces
R ²	Explained variance
RD	Radial diffusivity
ROC	Receiver operating curve
SAEs	Serious adverse events
SVD	Cerebral small vessel disease
SVDp	Measure of white matter hyperintensity
TBSS	Tract-based spatial statistics
TE	Echo time pulse sequence parameter
TIA	Transient ischemic attack
TMT-A	Trail-making test A
TMT-B	Trail-making test B
TPM	Tissue Probability Map
TR	Repetition time pulse sequence parameter
VD	Vascular dementia
VIF	Variance inflation factor
WM	White matter
WMH	White matter hyperintensity
WML	White matter lesion

Chapter 1

Introduction to Cerebral Small Vessel Disease

1. Pathogenesis of Small Vessel Disease

1.1. Anatomy

Cerebral small vessel disease (SVD) refers to different pathological processes affecting the small vessels in the brain such as the small arteries, arterioles, capillaries and small veins which mainly supply the subcortical white matter (WM), deep grey matter nuclei, and brainstem with blood ^{1,2}. These vessels originate from medium sized arteries or directly from large arteries as arterial perforators ¹.

The pathology underlying SVD, particularly in the early stages, is not well understood since it is difficult to image and examine the small vessels in vivo ³. There are a number of reported arterial pathologies. These include: 1) Diffuse small vessel arteriopathy, also called arteriosclerosis, and 2) more focal areas of microatheroma in the large intracerebral arteries such as middle cerebral artery and at the origin of the perforating arteries ^{1,2}. Diffuse small vessel arteriopathy is characterized by a considerable loss in smooth muscle cells from the tunica media and deposits of fibro-hyaline material, thereby narrowing of the lumen, further thickening the small vessel wall ¹. It is thought to be primarily caused by hypertension ³. On the other hand, microatheroma in the middle cerebral artery or other larger intra-cranial arteries or perforating arterioles could occlude the perforating arteriole. Vessel leakage or rupture due to vessel wall damage, microaneurysms or amyloid deposition may further aggravate the condition, which can result in micro- and macrohaemorrhages ¹. As the death incidence from lacunar stroke, a widely accepted sign of SVD ¹, is low, most of the reported pathology likely reflect late rather than the early stage disease ⁴. An understanding of the early disease in humans is however essential in order to prevent brain damage and to minimize clinical consequences of the disease ⁵.

1.2 Cerebral Hypoperfusion

Vessel lumen restriction, wall thickening and a loss of smooth cells are thought to lead to impaired vasoreactivity and cerebral autoregulation as vessels become

elongated, tortuous and inflexible ^{1,3}. Impaired autoregulation results in reduced cerebral blood flow and chronic cerebral hypoperfusion. An ischemic-hypoxic mechanism may be triggered resulting in the degeneration of myelinated fibers and diffuse alterations in the WM ¹.

Reduced cerebral blood flow (CBF) has been associated with WM damage ⁶. Whether cerebral hypoperfusion is casual or merely secondary to tissue damage remains controversial. While some studies show that patients with high growth in WM damage over time demonstrate lower CBF at baseline ^{7,8,9}, there is also evidence in a large longitudinal study that reduced CBF is a mere consequence of significant WM damage ¹⁰.

1.3 Increased Blood Brain Barrier permeability

The blood brain barrier (BBB) is a semipermeable borderline of endothelial cells that hinders certain ions, molecules and cells in the blood from non-selectively passing into the extracellular fluid of the central nervous system (CNS) ¹¹. The BBB is essential for the normal neuronal function as it guards the CNS from potential toxins, inflammation, pathogens, injury and disease ¹¹.

The consequences of BBB leakage and barrier tightness in cerebral ischemic injury may involve ¹²: (1) the initiation of oxidative stress related signaling whereby free radicals irreversibly attack cellular lipids and proteins leading to cell death, (2) DNA damage mediating the disruption of BBB following ischemia, (3) alteration of the protein kinase C's activation targeted proteins responsible for mediating BBB permeability and (4) activation of transcription factors inducing vascular permeability.

Over the last 20 years it became clear that patients with SVD show an increased BBB permeability compared to healthy subjects. First evidence came from studies demonstrating that the CSF to serum albumin ratio, an indicator of BBB leakage, was significantly increased in SVD patients ¹³⁻¹⁵. Post-mortem studies strengthened this finding showing evidence of blood-borne proteins such as immunoglobulin and fibrinogen associated with BBB disruption in brains of patients with SVD ¹⁶. In-depth spatial MRI analysis further showed that BBB permeability was significantly increased in the whole WM and also in the normal-appearing WM (NAWM) ^{17,18}. On the other hand, it remains controversial whether BBB permeability is increased in

areas with WM damage showing inconsistent findings ^{17,19,20}. It is possible that methodological discrepancies may explain the evidence.

2 Subtypes in SVD

SVD includes a whole spectrum of clinical abnormalities in combination with various vascular and genetic risk factors, which are related to different disease subtypes ^{1,3}. The most common ones are described below.

2.1 Sporadic SVD

Sporadic SVD causes a quarter of all ischemic strokes, primarily lacunar infarction, and is the most common pathology underlying vascular cognitive impairment and contributes up to 45% of all dementia cases ³. It also results in disabilities such as motor and gait impairment, mood disorders and urinary problems ¹. The 2 major disease subtypes in sporadic SVD are arteriosclerosis and cerebral amyloid angiopathy (CAA) ¹. Arteriosclerosis is the most common type and primarily occurs in the context of ageing involving vascular risk factors such as hypertension and diabetes. CAA is a condition caused by the deposition of β -amyloid in the cerebral vessel walls of arteries, arterioles, and capillaries starting in the tunica media and adventitia gradually leading to a loss of smooth muscle cells ¹. It is prevalent among elderly patients, and associated with an increased risk of intracerebral bleeding ²¹. In both forms genetic and environmental risk factors increase the likelihood of getting the disease ^{1,22–25}.

2.2 Monogenetic SVD (CADASIL)

Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL) is an inherited small-vessel arteriopathy usually causing stroke in middle-aged patients ²⁶. It is caused by a mutation in the *NOTCH3* gene on the chromosome 19 ^{27,28}. Overall, CADASIL is classified as a rare disease with current prevalence estimate of mutation carriers of around 0.80-5 per 100.000 people ²⁹. Estimating the exact disease burden in the whole population is complicated by recent evidence suggesting that CADASIL may also be present in patients with milder clinical variants that presently are undiagnosed ³⁰.

Although CADASIL varies in clinical symptoms, four main symptoms associated with the condition are: 1) recurrent ischemic strokes, 2) cognitive decline, 3) migraine with aura and 4) psychiatric disturbances ^{29,31}. Cerebral transient ischemic attacks and

strokes are common clinical features in CADASIL affecting 60-85% of all symptomatic patients ^{26,32}. Reoccurring ischemic strokes may eventually lead to overall decline in cognition and motor functions, gait disturbances and pseudo bulbar palsy ³³. Cognitive impairment gradually worsens over time in CADASIL particularly in areas such as reasoning, language, visuospatial abilities and memory ³⁴. In around 50% of patients with CADASIL migraine with aura is observed ³⁵ and 25-40% of them experience psychiatric symptoms such as significant mood disturbances or apathy ^{26,36}.

To date, there is no disease-modifying treatment available and patients are symptomatically treated based on the current clinical practice ²⁹. As secondary prevention strategies, patients are advised to adopt healthy lifestyle factors such as restrict alcohol consumption, to do regular exercise, to have a balanced diet with the aim of avoiding overweight and to stop smoking ^{37,38}.

Given that CADASIL is marked by clinical, radiological symptoms and pathological features which are similar to sporadic SVD but without external confounding factors such as aging, the inherited condition has been proposed as an attractive model in better understanding sporadic SVD ¹.

3. Clinical Features of SVD

3.1 Cognitive symptoms

SVD is the important cause for vascular cognitive impairment and vascular dementia ^{1,5}. Patients with clinically diagnosed SVD and CADASIL often show cognitive deficits in executive function, i.e. goal-oriented behavior, and in processing speed. ^{39,40}. In contrast to Alzheimer disease (AD), long-term memory function is relatively spared in sporadic SVD patients ^{41,42}. However, a clear distinction between SVD and AD solely based on neuropsychological functioning alone may however be challenging and even misleading, as many elderly people have mixed disease ⁴³. Over time cognitive impairment may worsen eventually leading to dementia ^{44,45}. This, however, does not hold true for all SVD patients. Some patients show steeper declines, whereas others show little cognitive changes over the years ⁴². Even in severe symptomatic SVD only executive function but not processing speed or working memory significantly declined over a 3 years follow-up period ⁴².

3.2 Psychiatric & motor symptoms

Psychiatric symptoms such as depression and apathy are common in SVD with a significantly higher severity and prevalence of apathy in small-vessel compared to large-vessel vascular dementia ⁴⁶. Depression is one of the leading causes of disability worldwide with major depressive disorder defined by a list of symptoms in the DSM-V ⁴⁷. This includes depressed mood, loss of interest or pleasure, impaired concentration, insomnia, change in appetite, feeling of worthlessness, psychomotor agitation and thoughts of death. In contrast to depression, apathy is characterized by the following key diagnostic criteria ⁴⁸: (1) a diminishment of goal-directed behavior or cognitive activity, (2) a reduction of spontaneous emotions or emotional response to the environment and (3) the lack of interest/ engagement in social contexts. Both depression and apathy are syndromes that are often diagnosed in SVD with low prevalence of comorbidity ⁴⁹. Separating the symptoms is also relevant with regard to cognition and neuroimaging. While depression has been associated with WMH volume ⁵⁰, apathy but not depression has been associated with WM network disruption and dementia conversion after controlling for subjective memory complaints ⁵¹⁻⁵⁴. Similarly to sporadic SVD, CADASIL patients often experience psychiatric disturbances with mood disorders most often being reported ⁵⁵.

Studies show that patients with SVD also often demonstrate unsteadiness of gait and impaired balance on walking putting them at risk of falls ^{56,57}. In the LADIS study walking speed was related to WMH burden with patients being the slowest in the severe WMH category ⁵⁸. Recent evidence, however, shed some doubts regarding a causal role of gait impairment in sporadic SVD. Gait performance in CADASIL patients with severe WM lesions was relatively preserved suggesting that age-related comorbidities other than SVD may be central in explaining gait impairment in elderly sporadic SVD patients ⁵⁹.

4. Why are surrogate markers necessary for clinical studies in SVD?

Overall, there are still significant gaps in our understanding on how the disease progresses over time ³. This is primarily due to the fact that imaging the small vessels in vivo is still very challenging given the current technological advances ³. Not understanding the initial stages of the disease also prevents treatment options with the aim of preserving the initially damaged small vessels ³. As a result, clinical research on SVD patients usually starts when the first clinical symptoms already set

in such as when the patient suffers from a lacunar stroke. But even at that stage it is difficult to track the disease progression as patients with SVD only experience few lacunar strokes with a reported annual recurrent rate of 4-11%⁶⁰⁻⁶³ and show slow cognitive decline. It usually takes some years until a potential diagnosis of dementia can be made^{64,65}. Using lacunar strokes, dementia diagnosis or cognitive decline as clinical endpoints in a study is therefore very expensive as a large sample size in the 1000s and a long trial duration would be required for a clinical trial⁶⁶. Furthermore, dementia conversion is significantly increased in patients with comorbidities such as SVD and AD⁶⁷, which raises the question if the dementia diagnosis in some patients actually primarily reflects the progression of SVD.

To be able to run phase 2 clinical trials more efficiently, more cost-effective and quickly, surrogate markers for SVD are needed. These surrogate markers would need to fulfill a number of criteria⁶⁶ to be useful:

1. the marker needs to be able to predict clinical outcome measures where changes influenced by a therapy on a surrogate marker correspond to changes in a clinically meaningful end point such as stroke, dementia or cognitive decline
2. the marker's change has to be noticeable prospectively
3. the marker would require a reasonable sample size estimate for a clinical trial

Over the recent years surrogate markers in SVD have been developed and used. These include imaging measures as well as serum biomarkers. A summary of the pathogenesis and radiological manifestations of cerebral SVD can be found in Fig 1.

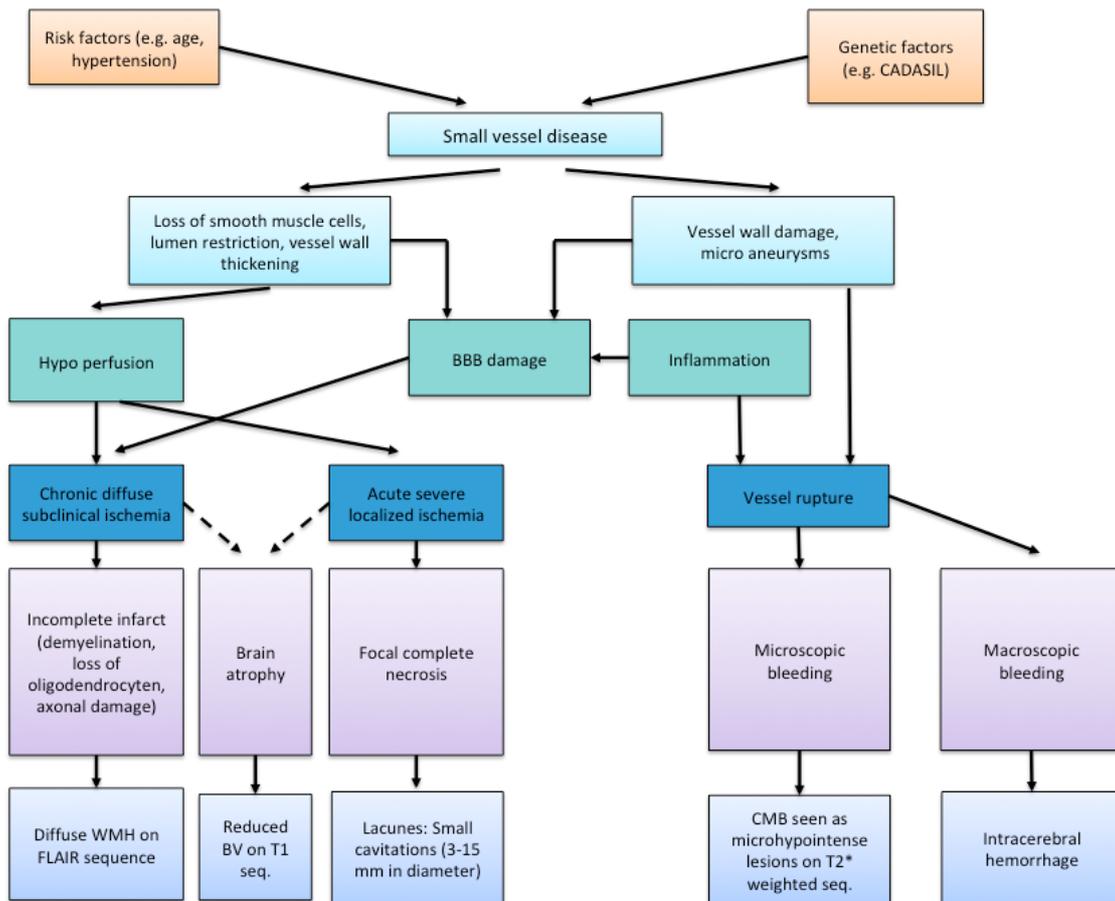


Fig 1. Pathogenesis and radiological manifestations of SVD (adapted from Pantoni, 2010). SVD is a heterogeneous disease that is understood. Pathological changes can result in ischemic and hemorrhagic incidences. Incomplete infarcts and focal complete necrosis are thought of being manifestations of the ischemic SVD process, which can be seen on various MRI sequences. Brain atrophy is another central feature of SVD acknowledged by the STRIVE neuroimaging initiative ⁶⁸. Studies further showed that blood-brain barrier (BBB) damage, subclinical inflammation may be involved.

5. Surrogate markers

5.1 MRI markers

Clinical research and care in SVD benefited enormously from magnetic resonance imaging (MRI) over the recent decades as a medical imaging technique to visualize structural brain changes in vivo (Fig 2) ³. Taking advantage of the differing water compositions of tissue, MRI generates contrast images of the brain by applying strong magnetic fields, magnetic field gradients and radio waves ^{69,70}. The MRI images not only differentiate between WM, grey matter (GM) and CSF but also capture various structural abnormalities in the brain that can be seen in SVD patients ⁶⁸. The radiological manifestations in SVD give indications about the disease's extent. These markers are: white matter lesions (WML), brain atrophy, lacunes, cerebral microbleeds (CMB) and enlarged perivascular spaces ^{1,68}.

5.1.1 WM Lesions

WML of presumed vascular origin, also termed as leukoaraiosis meaning 'diminution of the density of representation of the white matter' first seen on CT ⁷¹, is a key feature in SVD and is thought to be the result of the chronic diffuse injury reflected by an ischemic demyelination and an axonal degeneration ⁷². WML can be detected as hyper-intense regions (increase in water content in the injured regions) on a T2 or FLAIR (fluid-attenuated inversion recovery) MRI scan and are therefore termed as white matter hyperintensities (WMH) ⁶⁸. In a FLAIR MRI scan the signal from CSF is suppressed in order to increase the contrast between WMH and normal tissue. Being bilateral and symmetrical in shape, WMH are usually found in the periventricular and deep WM regions ⁶⁸. Other potential brain locations are the subcortical GM structures such as the basal ganglia and also the brain stem ⁶⁸. WMH are usually estimated either as WMH volume or WMH lesion load (i.e. the percentage of WMH volume in the brain). There are various methods to estimate the burden including both visual rating scales such as the Fazekas scale, ⁷³ and volumetric measurements. ⁷⁴. Volumetric measurements of WMH are usually labor-intensive, although provide a more accurate estimate.

WMH have been found to be associated with clinical outcome measures such as impaired cognition, gait, depression and urinary problems in elderly people ⁷⁵. WMH are frequently detected as incidental characteristics on a MRI scan in people in their 60s and older with varying extent and distribution ⁷⁶. In fact, more than 95 % of all

community-dwelling elderly people aged 65 or older without a reported stroke or TIA being scanned in a large study showed some indications of WM abnormalities ⁷⁷. On the clinical level, WMH have been shown to be associated with impaired cognitive function in healthy elderly people ⁷⁸. More detailed imaging analyses suggested that the exact location of the WMH may be critical for the manifestation of the symptoms. Whereas periventricular WMH were associated with cognitive dysfunction, higher mean arterial pressure and age, deep WMH were related to a history of mood disorders and higher body mass index (BMI) ^{50,79,80}. The location of the WMH may also indicate distinct underlying pathologies. While periventricular WMH, are marked primarily by significant gliosis, loosening of WM fibers and myelin loss, deep WMH are additionally related to more axonal loss, vacuolation and increased tissue loss ^{74,81,82}.

In severe SVD patients with lacunar stroke the association between WMH and cognition has to be shown weak or absent, ⁸³⁻⁸⁶ perhaps reflecting a threshold effect or a lack of power in a group who all had severe WMH.

WMH progress over time ⁸⁷⁻⁸⁹ and growth often involves expansion of existing lesions ⁹⁰. Predictors of a higher progression rate include extensive WMH at baseline, stroke, uncontrolled untreated hypertension and confluent instead of punctuate lesions ⁹¹⁻⁹⁴. Both WMH load at baseline and rate of progression of WMH have been associated with cognitive decline as well as global functional decline and disability ^{89,90,95,96}. WMH may also be comorbid with degenerative brain diseases together influencing the clinical outcome measures. For instance, higher periventricular WMH load at baseline in patients with amnesic mild cognitive impairment (MCI) significantly increased the likelihood to convert to AD within a three-years period ⁹⁷.

Although it is generally accepted that WMH grow over time, recent evidence showed that they are sometimes stable or shrink suggesting that WMH over time may be more dynamic than originally believed ⁹⁸. When it comes to clinical outcome measures, however, it has been shown that a regression in WMH was not related to improved cognitive function but likely reflected a more stable SVD condition in those patients ⁹⁸.

The ability to detect changes over a few years may suggest that WMH would be suitable as a surrogate marker to evaluate new therapies ⁹⁹. In the population-based

Austrian Stroke Prevention Study (ASPS) a sample size of 195 patients with confluent WMH would be required to detect a 20% treatment effect over a 3-years period. The sample size went up to 635 patients when enrolling only participants with early confluent WMH ¹⁰⁰. In lacunar stroke patients a sample size of 124, 178 or 279 per treatment arm would be required to detect a treatment effect of 30%, 25% or 20% in a clinical trial characterised by 3-years duration assuming a balanced design with measurements taken every year evenly ⁶⁶. In the imaging study arm of the positive SPRINT-MIND multicentre randomized clinical trial, where the change in WM lesion volume was the primary outcome measure, 449 patients had follow-up imaging data available around 4 years later ¹⁰¹.

5.1.2 *Lacunae*

Lacunae are defined as CSF filled cavities within the WM or subcortical matter, e.g. basal ganglia, pons or brainstem, of 3-15 mm in diameter and can be detected as regions of hypodensities on T1-weighted scan ⁶⁸. On FLAIR scans lacunae are detected as hypointense with a surrounding rim of hyperintensity ¹.

It is generally believed that lacunae are a result of lacunar infarcts ¹. In the acute phase lacunar infarcts appear as high signal on diffusion-weighted imaging (DWI) ¹⁰². However, not all lacunar infarcts result in lacunae. It has also been shown that one third of recent infarcts do not form CSF filled cavities but become WMH ^{68,103}. It is therefore assumed that the number of lacunae does not fully represent the disease burden caused by lacunar infarcts ¹⁰⁴, and lacune counts may underrepresent the true burden of disease. ¹⁰⁵.

Lacunae can be clinically silent or symptomatic. Clinical symptoms substantially vary based on the lesion location ¹⁰⁶. The most common symptoms are: pure motor hemiparesis, sensorimotor stroke, pure sensory stroke, dysarthria (clumsy-hand syndrome), and ataxic hemiparesis ¹⁰⁷.

Cognitive decline and the increased risk of converting to dementia are the long-term consequences of a lacunar infarct ^{108,109}. Silent infarcts have been demonstrated to be associated with specific deficits in cognitive and bodily function even in healthy elderly cohorts emphasizing their importance in the clinical outcome measures ¹¹⁰⁻¹¹². Coexisting morbidities such as AD, Parkinson's disease (PD) or recurrences of a

lacunar stroke may have synergic effects and may further accelerate the onset of dementia.

In SVD the changes in lacune counts significantly vary over time. Whereas some patients show stable lacunes counts, other patients demonstrate changes ⁴⁴. Employing lacune volume as a surrogate marker in a hypothetical 1.5 years clinical trial in CADASIL patients demonstrated that 11354 patients per treatment arm would be required to detect a 30% treatment effect. In comparison, 5387 patients would be required for the some clinical trial conditions if processing speed as a marker would be used ¹¹³.

5.1.3 Cerebral microbleeds

Cerebral microbleeds (CMB) are hemosiderin deposits from blood cells that presumably have leaked out of small vessels and often accompanied by spontaneous intracerebral hemorrhage ¹. Being between 2-5 *mm* in size, CMB can be detected on T2* weighted or susceptibility-weighted MRI sequences ⁶⁸. CMB burden is most often estimated employing visual rating scales, which has been shown to improve inter-rater reliability with regard their presence and their anatomical location ¹¹⁴. The most common ones are the Microbleeds Anatomic Rating Scale (MARS) ¹¹⁵ and Brain Observer MicroBleed Scale (BOMBS) ¹¹⁶.

The pathogenesis and etiology of CMB are not fully understood. It has been shown that patients with high CMB count have increased vessel wall thickness than patients with lower CMB burden ¹¹⁷. In terms of risk factors, CMB have also been found be related to age, hypertension, and cigarette smoking ^{118–121}. Furthermore, associations between CMB burden and other markers of SVD such as WMH and lacunes have been found found in different cohort studies ^{122–125}. Similar to lacunes and WMH, presence of any CMB at baseline was a strong predictor for increased CMB burden over a 3-years interval, with a risk factor of almost 5 times higher compared to patients without CMB ¹²⁶. Furthermore, initial CMB disease burden also predicted future cerebrovascular events in patients with stroke, and in community-dwelling elderly cohorts ^{127–131}.

In line with the cerebrovascular significance of the CMB, cross-sectional associations were found with cognitive impairment such as attention, processing speed, executive function and general cognition ^{114,132–134}. A more refined imaging analysis showed

that lobar versus deep CMB were connected to different cognitive impairments in SVD patients. Lobar CMB, which are primarily associated with CAA^{135,136}, were related to executive function and memory, whereas deeper CMB, most seen in patients with hypertensive SVD¹³⁷, were more associated with psychomotor speed and attention¹³². Looking at the CMB by brain regions, it has been also shown that CMB in the temporal lobe were more related to memory and attention¹³⁸. On the other hand, CMB located in the frontal areas were significantly associated with psychomotor speed, concept- shifting and attention¹³⁸.

On the longitudinal basis, changes in CMB have been associated with changes in cognition such as global cognition and executive function^{139–141}. However, longitudinal evidence is still sparse and more data is required to determine whether they do indeed predict cognitive decline and dementia, and whether this is similar to other MRI markers of SVD¹¹⁴.

5.1.4 Brain atrophy

On a T1 MRI scan, volumes of WM, GM and CSF can be obtained and brain atrophy over time can be estimated. In order to account for head size and to attenuate inter-individual variation, it is of standard practice to normalize brain volume to head size¹⁴².

Brain atrophy is a central and reliable imaging marker in SVD⁶⁸. In the context of cerebrovascular disease it is assumed that brain atrophy involves neuronal loss, cortical thinning, subcortical vascular pathology underlying WM shrinkage and secondary neurodegenerative changes^{143–145}. The marker has been shown to be associated with cognitive impairment, and cognitive decline over time, in patients with SVD^{146,147}.

Although brain atrophy as a distinct marker correlates with cognitive impairment, it may share significant underlying pathological aspects with WMH. It has been demonstrated that WMH progression is associated with increasing rates of regional GM atrophy, which is the most important factor explaining whole brain atrophy¹⁴⁸. The authors concluded that any therapies aimed at attenuating WMH might also significantly preserve secondary brain atrophy. Secondary degeneration of the cortex may be also influenced by acute ischemic infarcts. Damage to the WM tracts

connecting the acute infarcts with the distant cortex significantly correlated with thickness changes in that region ¹⁴⁹.

Using brain atrophy as a surrogate marker in a hypothetical clinical trial with lacunar infarct as an inclusion criterion resulted in a sample size estimation of 145, 208, 325 patients per treatment arm with a treatment effect of (30%, 25%, 20%) respectively ⁶⁶.

5.1.5 Enlarged perivascular spaces

Perivascular spaces (PVS) are fluid-filled spaces surrounding blood vessels in the WM and GM, and serve as a lymphatic drainage system for interstitial fluid from the brain ^{150,151}. Enlarged PVS are visible on MRI with a signal intensity similar to CSF on all sequences and can most often be detected in the basal ganglia, high convexities and midbrain ¹⁵². Dilated PVS (> 5mm) are found in around 1.6% of healthy people ¹⁵³. Some studies showed that the number of enlarged PVS were related to cognition in healthy elderly population cohorts ^{154,155}. Although enlarged PVS count has also been associated with symptomatic lacunar stroke syndrome and clinical markers of SVD such as CMB and WMH ^{156–159}, their overall clinical significance as a marker in SVD has not yet consistently been shown. In a 3-years longitudinal cohort study it has been shown that while PVS are highly prevalent in patients with lacunar stroke and associated with other baseline SVD imaging markers, they were not a predictor for cognitive decline over 5 years and showed no change in volume within the 3 follow-up years ¹⁶⁰.

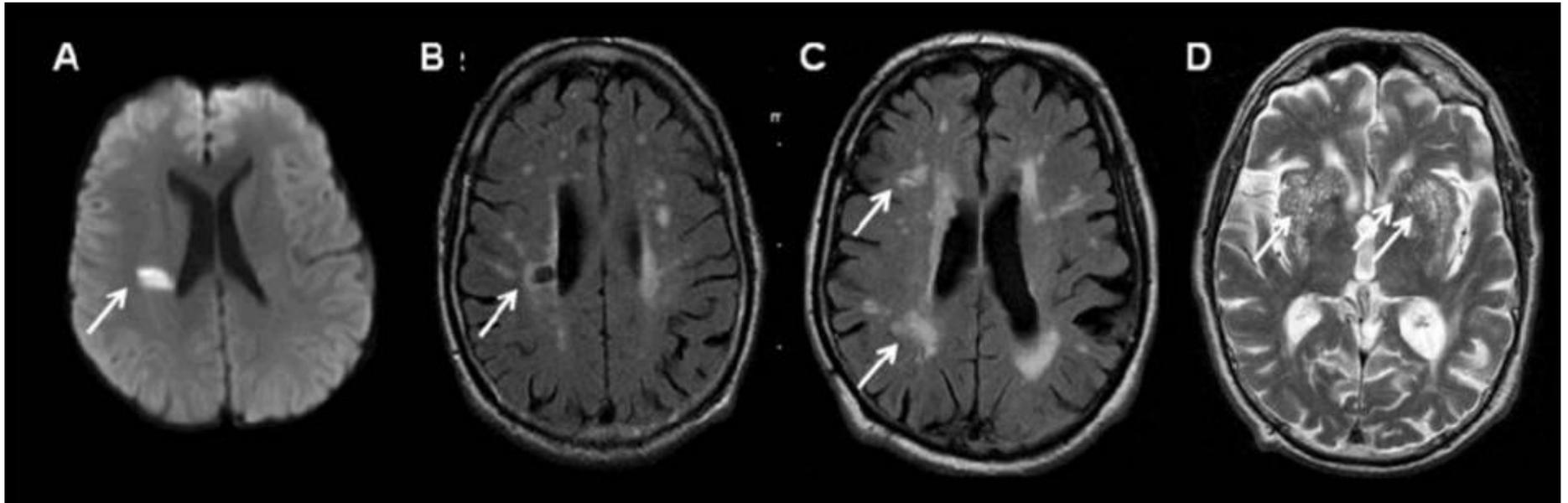


Fig 2. Key neuroradiological characteristics of SVD. A) diffusion-weighted image with an acute small deep lacunar infarct B) Lacune on FLAIR imaging C) WMH on FLAIR imaging D) Perivascular spaces on T2-weighted imaging, hyperintense due to containing CSF-like fluid. *Adapted from Wardlaw et al., 2013.*

5.1.6 Combination of imaging markers

The MRI markers usually do not exist in isolation in SVD, but are often found together in various combinations reflecting the heterogeneity of the disease. As a result, combined MRI scores have been created with the aim of capturing total brain damage in SVD patients. Total MRI scores have been shown to be associated with cognitive impairment, hypertension, age, gait, and other risk factors for SVD ^{57,161–164}. A combined MRI score has also been shown longitudinally to predict cognitive decline over 4 years and dementia conversion up to 5 years ^{165,166}. It could furthermore be demonstrated that, if it was used as an inclusion criteria, a higher total score would reduce the minimum sample sizes required for hypothetical clinical trials by around 40%–60% ¹⁶⁶.

5.1.7 Diffusion tensor imaging assessing microstructural WM damage

Damage to the WM does not only occur in areas of WMH seen on MRI images, as has been shown by diffusion tensor imaging (DTI) studies, demonstrating that the “normal-appearing” WMH is in fact not normal ¹⁶⁷. DTI is a further development of diffusion-weighted imaging and measures the random three-dimensional (3D) translational motion of water molecules in at least six non-collinear directions in the brain in order to acquire a tensor ^{168,169}.

The directionality of diffusion can be classed as either isotropic or anisotropic. Isotropic diffusion is defined as diffusion, which shows no preferential directionality and can be described by a single diffusion coefficient. In the brain isotropic diffusion is seen in the CSF for instance. On the other hand, anisotropic diffusion is characterized by increased diffusion in one principal direction. In oriented tissue such as in the WM tracts in the brain, anisotropic diffusion is observed, as water molecules are more likely to diffuse along the WM tracts than perpendicular to them.

Tensor analysis can be used to estimate the directionality of the diffusion in the WM. The eigenvectors and eigenvalues obtained can be fitted to a 3D ellipsoid, which reflects the average diffusion coefficient in each direction for each voxel. It is characterized by the length of the longest, middle and shortest axes quantified by eigenvalues $\lambda_1, \lambda_2, \lambda_3$ respectively together with their three orientations. The eigenvalues can then be used to create specific parameters.

The four principal scalar measures most often used are: 1) Mean diffusivity (MD) or Apparent Diffusion Coefficient (ADC); 2) Fractional anisotropy (FA); 3) Axial diffusivity (AD); 4) Radial diffusivity (RD).

MD measures the overall extent of water molecule diffusion in terms of how far the molecule moved in a prescribed amount of time; it is the average across all directions within the voxel. It is calculated as:

$$MD = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3} = \frac{D_{xx} + D_{yy} + D_{zz}}{3}$$

where D_{xx} , D_{yy} , D_{zz} are the diagonal terms of the diffusion tensor. The higher the MD the higher is the water diffusion.

FA represents a marker of the directionality of diffusion and quantifies the directional variation in diffusion around the ellipsoid:

$$FA = \sqrt{\frac{3}{2}} * \sqrt{\frac{(\lambda_1 - D)^2 + (\lambda_2 - D)^2 + (\lambda_3 - D)^2}{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}$$

where $D = (\lambda_1 + \lambda_2 + \lambda_3)/3$. It ranges from 0 (diffusivity equal in all directions) to 1 (completely unidirectional).

Axial diffusivity (AD) and radial diffusivity (RD) capture the diffusion along and perpendicular to its main diffusion direction, respectively:

$$AD = \lambda_1 > \lambda_2, \lambda_3$$

$$RD = \frac{\lambda_2 + \lambda_3}{2}$$

where $\lambda_1, \lambda_2, \lambda_3$ represent the eigenvalues of the length of the longest, middle and shortest axes.

In SVD, higher MD, AD and RD values and lower FA values in the NAWM as well as in the WM lesions were found in comparison to healthy controls suggesting a reduced WM integrity and more ultrastructural damage^{86,170-170}. This ultrastructural damage is believed to reflect axonal damage and demyelination¹⁷¹.

DTI parameters were significantly related to cognitive impairment such as executive function, processing speed and global cognition^{86,172,173} and were stronger predictors than conventional MRI measures such as WMH load^{85,170}. Adding MRI to the DTI

measures in a linear model significantly increased the predicted explained variance in cognitive impairment ranging between 57-74%^{85,86}.

Longitudinally, DTI measures have been demonstrated to be more sensitive to change in SVD than WMH load and brain volume over a 1-year period⁸⁵. Microstructural WM integrity on DTI at baseline also predicted cognitive decline over a period of 3 years¹⁷⁴. A later study also showed that in severe SVD changes in DTI measures over 3 years predicted cognitive decline and conversion to dementia over 5 years⁴⁴. DTI measures such as MD and FA have therefore been included in clinical trials as endpoints^{175,176}.

In severe SVD a minimum sample size of 128, 185 and 289 is required per treatment arm for a treatment effect of 30%, 25%, 20% in a clinical trial characterised by a 3-years duration with DTI assuming a balanced design with measurements taken every year evenly⁶⁶. For CADASIL patients the sample size estimation was significantly lower for a 1.5 years clinical trial with a sample size of 96 and a treatment effect size of 30%¹¹³.

5.2 Biological markers in SVD

5.2.1 Why we have to look beyond imaging markers

As pointed out in the previous section, imaging measures are central in SVD allowing us to understand better disease progression through different stages of the disease. Imaging markers have furthermore served as surrogate markers of SVD in clinical studies. Despite their vital importance, imaging also has significant limitations and this may have implications in terms of the interpreting the findings in a clinical trial. First, using MRI as surrogate endpoint may result in patient selection bias, where only certain participants may be enrolled in the clinical trial. Criteria for excluding a patient may solely be technical as certain body implants make it too dangerous to be exposed to strong magnetic field. Another reason may be that the patient does not feel comfortable enough doing an MRI or even suffer from anxieties. Second, despite a significant reduction in sample size estimation when using MRI measures compared to clinical measures in SVD, employing imaging measures as surrogate markers in a clinical trial still involves relatively high cost and efforts. Third, as clinical trials often involve multicenter designs, harmonization of MRI measures coming from different scanners can be challenging. A lot of efforts and time investment may be

needed to account for confounding factors originating from different types of scanners. This may even compromise the trial's statistical power to be able to detect a possible treatment effect.

On the other hand circulating biomarkers, often purely based on blood sampling, may address some of these limitations. There may be less selection bias, it may be significantly cheaper and easier to process and samples can be analyzed in one centralized reference laboratory.

5.2.2 Promising circulating biomarkers

Elevated levels of circulating markers have been found SVD patients. These include C reactive protein (CRP) marker, serum albumin and albuminuria and serum neurofilament light chain (NfL) ^{177,178}.

5.2.2.1 C reactive protein marker

Although inflammation is a response to infections and injuries, it can have deleterious effects in healthy tissue. CRP is a marker of systematic inflammation ¹⁷⁷. It has been demonstrated that elevated peripheral levels of CRP together with cytokine interleukin 6 are related to a higher risk of vascular dementia but not to AD ¹⁷⁹. Further evidence showed that higher values in CRP are also associated with lower WM integrity and the presence as well as progression of WMH independent of cardiovascular risk factors and carotid atherosclerosis ^{180,181}. Despite these findings, the importance of the CRP especially in individuals with WMH is still inconclusive ¹⁸²⁻¹⁸⁵.

5.2.2.2 Serum albumin and albuminuria

Albumin is a common protein produced in the liver and transports various substances such as hormones and fatty acids among others. Albuminuria is a state of excreting too much albumin in the urine. Epidemiological studies demonstrated that albuminuria was related to cognitive decline in patients with vascular disease and with diabetes Miletus ¹⁸⁶ and to both Alzheimer's and vascular dementia conversion in community-dwelling Japanese elderly ¹⁸⁷. Recent evidence also showed that the combination of albuminuria and WMH lesions additively increased the likelihood of cognitive decline, dementia and all-cause mortality ¹⁸⁸. The authors suggest that urinary albumin may be a potential useful biomarker sensitive to future development

of stroke and dementia in SVD although more evidence is needed to characterize the association between albumin and SVD specific imaging markers.

5.2.2.3 Neurofilament light chain

NfL is a marker for neuroaxonal damage, which is released into the extracellular space, CSF and blood ¹⁸⁹ (Fig 3). The marker has been established in a number of neurodegenerative diseases such as AD ¹⁹⁰ and frontotemporal dementia ¹⁹¹. In a population-based cohort study NfL levels significantly increased in individuals beyond age 60, and were associated with greater brain atrophy both in the cross-sectional as well as longitudinal setting ¹⁹². The authors suggest that this association may be explained by subclinical comorbid pathologies. Higher baseline, but not changes in NfL, also predicted cognitive decline in the entire community-based cohort ¹⁹². Higher serum NfL levels were associated with greater disease burden and were related to cognitive impairment in SVD ¹⁹³. Furthermore higher NfL levels were related to greater disability and neurological disability in monogenic SVD patients ¹⁹³. It is yet to be tested whether baseline and change in NfL levels predict future dementia risk and cognitive decline in severe SVD. If it is to be used as a surrogate marker in clinical trials, a change in NfL level within a 2-3 years period typical of a clinical trial must be shown to be associated with clinical progression.

6. Conclusion

SVD is a prevalent condition causing lacunar stroke, vascular cognitive impairment and dementia. Overall there are still significant gaps in our understanding about the disease's progression and there are only a few treatment options available. One challenge in testing new therapies is to define the clinical endpoints in the study. Using lacunar strokes, dementia diagnosis or cognitive decline as clinical endpoints is very challenging as a large sample size would be required for a clinical trial. Over the recent years surrogate imaging markers in SVD have been developed and used.

DTI measuring ultrastructural damage in the WM and WM tract disruption including in areas outside the WMH have been shown to be strongly associated with cognitive impairment, cognitive decline and dementia conversion in SVD. Recent evidence also has shown that blood-based markers correlate with SVD disease burden and may therefore show promise as a potential biomarker in a clinical trial with SVD.

Testing the most often studied DTI markers in a clinical trial with SVD patients and comparing them with more advanced DTI markers across different degrees of SVD are crucial to determine the optimal surrogate marker as a clinical endpoint. A robust sensitive SVD marker may then allow us to test new therapies in order to slow down the disease's progression and ultimately reduce the risk of stroke, cognitive decline, disability and dementia conversion.

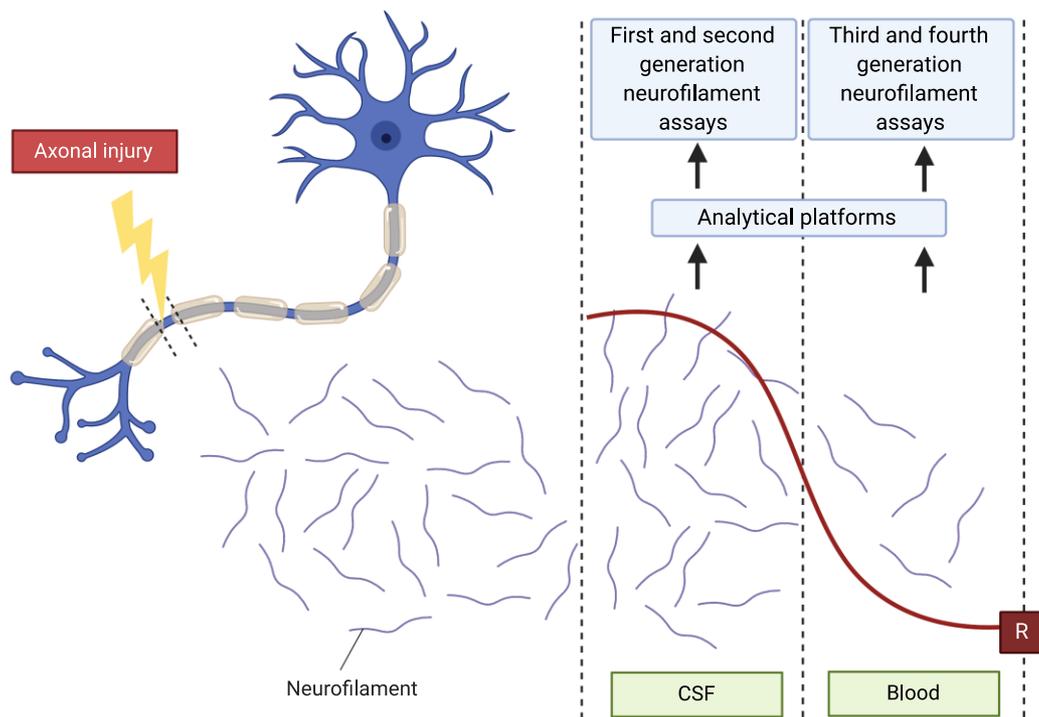


Fig 3. Neurofilament release after axonal damage. Upon axonal damage cytoskeletal proteins such as neurofilaments are released into the extracellular space and afterwards into the CSF and into the blood with a low concentration. While the first 2 generations of enzyme-linked immunosorbent assay (ELISA) immunoassays only had a limited sensitivity of neurofilament detection in the CSF, the third- and fourth generational device allowed reliable detection of neurofilament light chain in the blood which correlated with the ones found in the CSF. *Adapted from Khalil and colleagues (2018). Created by biorender.com*

7. Aims of the thesis

This thesis aims to address the following research questions:

1. Are diffusion tensor imaging measures robust surrogate markers in a multicenter study with SVD patients? (Chapter 2)
2. How important is white matter microstructural damage measured by diffusion tensor imaging to cognitive impairment and dementia conversion in patients with varying degrees of SVD severity? (Chapter 3)
3. Which of the ways to analyze diffusion data best predict future dementia risk, and how do traditional methods and new semi and fully automated approaches compare? (Chapter 4)
4. How reproducible are the recently developed automated DTI measures? (Chapter 5)
5. Is the serum neurofilament marker a suitable surrogate marker for a clinical trial in SVD patients? (Chapter 6)

Chapter 2

Evaluating DTI as a Surrogate Marker in a Randomized Multicenter Trial in SVD

1. Introduction

A major risk factor in symptomatic SVD is hypertension ¹⁹⁴. It is, however, unclear how intensively hypertension should be treated. Previous evidence showed that treating blood pressure (BP) intensively (systolic BP < 120 *mmHg*) vs. standardly (systolic BP < 140 *mmHg*) was associated with a lower rate of cardiovascular events in the SPRINT trial ¹⁹⁵ and a reduced risk of stroke in secondary prevention in the PROGRESS trial, where the systolic and diastolic BP was reduced by 9 and 4 *mmHg* respectively compared to the placebo control group ^{196,197}. It is, however, unclear whether this treatment would do more good than harm in patients with severe SVD who show impaired autoregulation of the small vessels ¹⁹⁸. Intensive BP lowering may significantly lower their CBF and as a result lead to even more brain damage in the WM and negatively impact cognitive function. However, this link has not been shown and evidence on the contrary coming from the SPS3 trial has demonstrated that there was no change in cognition between intensive systolic BP (<130 *mmHg*) lowering and standard systolic BP (130-149 *mmHg*) target administration in mild SVD patients ¹⁹⁹. Further evidence showed that systolic BP lowering was associated with a reduced risk of MCI and with less WMH progression in the intensive group (<120 *mmHg*) compared to the standard treatment group (<140 *mmHg*) in a stroke-free hypertensive population ^{101,200}.

The PRESERVE study ('How intensively should we treat blood PRESsure in established cERebral small VEssel disease?') is a multicenter randomized controlled trial that investigated whether intensive BP treatment (systolic BP < 125 *mmHg*) is associated with more WM microstructural damage than standard BP treatment (systolic BP= 130-140 *mmHg*). WM microstructural damage was measured by DTI at baseline and at 24 months. Previous evidence has shown that DTI measures were sensitive to change in a longitudinal study called SCANS involving severe SVD patients, ²⁰¹ and that DTI changes were associated with later clinical outcomes in

terms of dementia conversion ⁴⁴. Other surrogate markers in the trial were the conventional imaging markers NBV, WMH lesion load, lacune and CMB count. Patients were furthermore assessed in terms of their cognitive function over the 2 years.

PRESERVE is the first multicentre clinical trial in SVD employing DTI as a marker. An analysis in PRESERVE including only data at baseline demonstrated relationships with similar standardized effect sizes between DTI measures and cognitive function to those shown in previous single-center studies ²⁰². This chapter's goal is to evaluate DTI as a surrogate marker in a multicenter trial setting

1. by testing whether intensive vs. standard BP lowering is associated with more WM microstructural damage
2. by comparing the performance of the DTI measure to cognitive and conventional MRI measures

2. Method

2.1 Patients

167 patients were randomised in the clinical trial. Of these, 111 individuals also consented to enter the MRI substudy, and therefore had imaging analysis available (Figure 1) ¹⁷⁶. In addition, there was also another imaging substudy in PRESERVE employing the arterial spin labeling technique to measure CBF ²⁰³. The trial's inclusion criteria were clinical lacunar stroke with an anatomically corresponding lacunar infarct on MRI, in addition to confluent WMH graded ≥ 2 on the Fazekas scale at a minimum of 3 months post stroke ⁷³. Patients were at least 40 years of age with either systolic BP > 140 *mmHg* if they were on no anti-hypertensive medication or 125-140 *mmHg* if treated with antihypertensive medication. Patients were excluded from the study if they had a known single gene disorder causing SVD, a cause of stroke other than SVD, a clinically diagnosis of dementia, a life expectancy of less than 2 years, symptomatic postural hypertension, were unable to take part in the regular assessments, or were a woman of childbearing potential.

2.2 Trial Design

PRESERVE had a parallel trial design. Patients were randomly allocated to the intensive (systolic BP < 125 *mmHg*) or standard (systolic BP 130-140 *mmHg*) treatment regime by the local clinician in a 1:1 ratio via a centralized system. Treatment allocation was unblinded to the patients and clinical staff. In contrast, the

imaging and statistical analysis were blinded to the treatment allocation performed at the beginning of the trial.

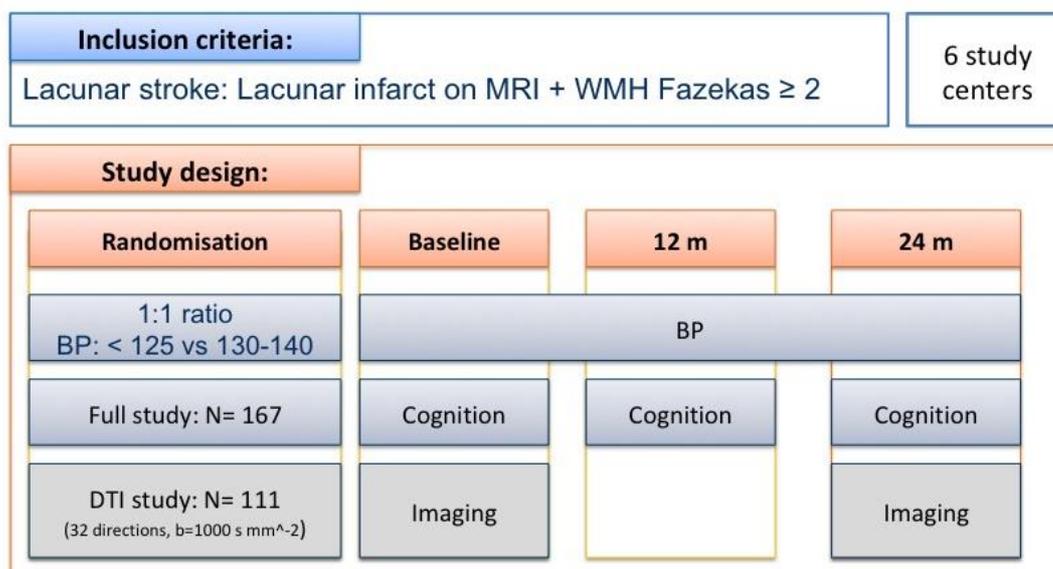


Figure 1. An overview of the study design with the MRI substudy included. Patients with lacunar stroke were enrolled in the study from 6 study sites. 167 Patients were randomized to the intensive (blood pressure (BP) < 125 mmHg) or to the standard treatment arm (BP = 130-140 mmHg). BP was repeatedly assessed over the 2 years. A subgroup of patients (N= 111) also underwent MRI imaging at both time points.

2.3 Outcome measures and follow-up assessments

The initial endpoint was a global cognitive score with imaging as a secondary endpoint. There was a pre-planned review after the expected publication of the SPS3 cognition study¹⁹⁹, which was published in 2014. SPS3 showed that there was no change in cognition over two years in 2916 patients with lacunar stroke. Following this, the steering committee met, and with funders agreement, halted recruitment to the cognitive only arm which had a planned sample size of 422, and only recruited to the combined imaging substudy (which had a sample size of 180), with the primary endpoint of the overall study becoming DTI. No interim analyses were conducted. Recruitment, however, stopped when 111 patients were reached and when the trial funding ended.

Outcome measures with imaging measures were available across 6 study sites. Clinical assessments and BP monitoring happened at 1, 3, 6, 12, 18 and 24 months. Patients were prescribed increase in antihypertensive medication when the study's BP target level was not met. Patients had MRI sessions at baseline and at 2 years. The primary clinical endpoint of the intention-to-treat (ITT) and per-protocol analysis (PP) was change in the DTI histogram measure mean diffusivity normalized peak height (MDPH). Patients' cognitive functions were assessed at baseline, 12 and 24 months.

2.4 Adverse event recording

At each visit, reported adverse events (AEs) and serious adverse events (SAEs) were recorded. Events of falls or postural dizziness were separately recorded. Instances of stroke and death outcomes were collected on a proforma and reviewed by two adjudicators blinded to treatment.

2.5 MRI acquisition

Different 3-Tesla MRI scanners were employed across the 6 centres (3 Philips Acheiva TX, 1 Philips Acheiva, 1 Philips Ingenia, 1 Siemens Verio, 1 Siemens Prisma, 1 Siemens Magnetom Prisma). 3D T1-weighted, DTI, T2*-weighted and FLAIR scans were obtained. T1W scans were obtained at 1 mm^3 isotropic voxel resolution and TR and TE optimized to make sure to have comparable T1 weighting and tissue contrast. DTI scans with a 2 mm^3 isotropic voxel resolution had similar TEs and long TRs to prevent T1 relaxation effects. Every DTI acquisition had 32 equally spaced, non-collinear diffusion gradient directions ($b=0$ s/mm^2 , $b=1000$ s/mm^2) to guarantee the same angular resolution and noise characteristics. Exact scanner and sequence details are shown in Table 1a-c.

2.6 MRI analysis

2.6.1 Brain volume and WMH lesion load

T1W scans were segmented into grey matter, WM and CSF tissue probability maps employing SPM12b (Statistical Parametric Mapping (SPM), <http://www.fil.ion.ucl.ac.uk/spm/>). Soft segmentations of the grey matter and WM tissue probability maps were used to compute brain volume. Normalized brain volume was computed from native T1 images as an estimate of size of the brain relative to the skull size with scaling factors derived from SIENAX, part of the FMRIB software library¹⁷⁶. To obtain a measure of WMH a semiautomated program JIM

(Xinapse Systems Limited; www.xinapse.com) was used to segment WMH regions. The whole brain lesion maps were used to compute a measure of lesion load as the percentage of WMH lesion volume against whole brain volume ¹⁷⁶.

2.6.2 DTI histogram measure

The image analysis is summarized in Figure 2. Eddy correct software from “FDT”, FMRIB’s Diffusion Toolbox, (<http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FDT>) was employed for DTI preprocessing. Mean diffusion (MD) and fractional anisotropy (FA) maps were created with ‘DTIFIT’. FLAIR to T1W and T1W to b0 registrations were performed using the FMRIB Linear Image Registration Tool (FLIRT) and the affine transformation matrices were concatenated to produce a FLAIR-to-DTI transformation ²⁰⁴. The tissue probability maps were registered into DTI space using these transformations. A hard segmentation method was applied to generate maps of tissue classes. Histogram analysis was conducted on the MD maps in all WM regions. The summary histogram measure mean diffusivity normalized peak height (MDPH) was derived from normalized histograms with 1000 bins and an upper threshold of 2.6×10^{-3} .

Table 1a. An overview of the scanners and sequence parameters employed at each site. Axial DTI sequences with 32 diffusion gradient directions at b-value = 1000 s/mm^2 , isotropic voxel resolution 2 mm^3 . FOV=Field of View; FLAIR=Fluid Attenuated Inversion Recovery

Axial DTI	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6
3 T Scanner(s)	Phillips Achieva TX	Phillips Achieva, Phillips Archieva TX	Phillips Achieva TX	Siemens Prisma	Phillips Ingenia	Siemens Verio Siemens Magnetom Prisma ^{fit}
TR	6850ms	6850ms	6850ms	9500ms	9100ms	11500ms
TE	75ms	75ms	75ms	93ms	82ms	93ms
In-plane FOV	224x224mm ²	224x224mm ²	224x224mm ²	192x192mm ²	224x224mm ²	192x192mm ²
N° slices	60	60	60	81	60	75
N° b0	8	8	8	2	8	2
Max Gradient Strength	80mT/m	80mT/m	80mT/m	40mT/m	45mT/m	45/80mT/m
Parallel imaging factor	3	3	3	2	3	2
N° headcoil channels	8	8	8	12	15	32

Table 1b. An overview of the exact scanners and sequence parameters employed at each site. Sagittal 3D T1-weighted with isotropic voxel resolution 1 mm^3

Sagittal 3D T1-weighted	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6
3 T Scanner(s)	Phillips Achieva TX	Phillips Achieva, Phillips Archieva TX	Phillips Achieva TX	Siemens Prisma	Phillips Ingenia	Siemens Verio Siemens Magnetom Prisma ^{fit}
TR	8.27ms	9.81ms	11ms	2200ms	8.53ms	2200ms
TE	4.61ms	4.60ms	4.61ms	2.94ms	4.61ms	2.97ms
In-plane FOV	$240^2 \times 170 \text{ mm}^3$	$240^2 \times 170 \text{ mm}^3$	$240^2 \times 170 \text{ mm}^3$	$256^2 \times 208 \text{ mm}^3$	$240^2 \times 170 \text{ mm}^3$	$256^2 \times 208 \text{ mm}^3$
Inversion Time				900ms		900ms

Table 1c. An overview of the exact scanners and sequence parameters employed at each site. Axial FLAIR with inversion time= 2800 ms

Axial FLAIR	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6
3 T Scanner(s)	Phillips Achieva TX	Phillips Achieva, Phillips Archieva TX	Phillips Achieva TX	Siemens Prisma	Phillips Ingenia	Siemens Verio Siemens Magnetom Prisma ^{fit}
TR	11000ms	11000ms	11000ms	8000ms	11000ms	8000ms
TE	120ms	120ms	120ms	121ms	120ms	124ms
In-plane FOV	230x230mm ²	230x230mm ²	230x230mm ²	208x230mm	230x230mm ²	208x230mm
Voxel size	0.48 ² x3mm ³	0.48 ² x3mm ³	0.48 ² x3mm ³	0.45 ² x3mm ³	0.48 ² x3mm ³	0.45 ² x3mm ³
N° slices	57	57	57	60	57	60

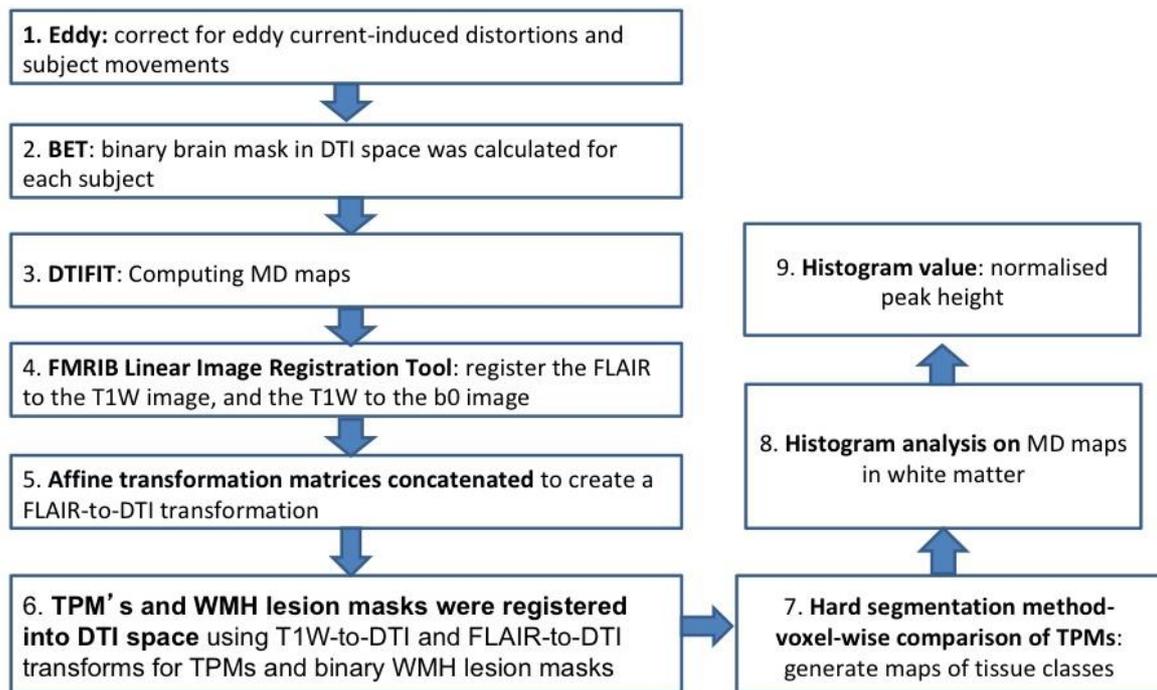


Figure 2. Flow Chart about the DTI histogram processing pipeline. DTI scans were pre-processed using eddy-correction and a binary brain mask was computed. DTIFIT was run on DTI fitting a diffusion tensor model at each voxel. Images with different MRI sequences were registered and affine transformation matrices were concatenated to create a FLAIR-to-DTI transformation. Tissue probability maps (TPM) and WMH lesion masks were registered into DTI space and a hard segmentation method was applied. A histogram analysis on MD maps were run and the histogram metric normalized peak height was computed.

2.6.3 Lacunes and cerebral microbleeds

In all cohorts lacune and CMB were identified by trained raters blinded to all clinical data. Lacunes were defined as CSF filled cavity within the WM or subcortical areas between 3-15 *mm* in diameter on T1-weighted, T2*-weighted and FLAIR images ⁶⁸. CMB were identified on T2* weighted GRE (Gradient echo) images as focal areas up to 10 *mm* in diameter ⁸⁶.

2.7 Cognitive Testing

Patients were cognitively tested in three neuropsychological assessments over the 2 years, which were time-wise matched as closely as possible to the MRI session. The cognitive tests used have been shown to be sensitive to impairments in SVD. The cognitive test battery included the following tests: WAIS-III Digit Symbol Coding test (DSC), Trail Making Test (TMT), phonemic and semantic verbal fluency task (FAS) and Rey Auditory Verbal Learning Test (RAVLT). Furthermore, the re-standardized National Adult Reading Test (NART-R), a test of premorbid IQ, was administered.

DSC is a coding task where participants pair a number from 1 to 9 into predefined geometric figures ²⁰⁵. A key at the beginning of each page indicates which character corresponds to each numeric value. The outcome measure was obtained by counting the number of pairs completed within a fixed period of 120 seconds. The DSC has been used many times and was originally developed as a mean to understand human associative learning ²⁰⁶. Its methodological strengths are the lack of confounding factors such as language, culture, and education in the cognitive assessment ²⁰⁷.

TMT is composed of 2 parts, TMT-A and TMT-B ²⁰⁸. In both types the patients were asked to connect a set of dots as quickly as possible. While in TMT-A these dots needed to be connected according to the sequence of rising numbers, in TMT-B patients needed to connect the dots while alternating between rising numbers and letters according to their position in the alphabet. In both tasks participants were instructed not to lift the pen and work as quickly and as accurately as possible.

FAS is a task which participants were asked to generate as many words as possible for a predefined semantic or phonemic category within a set period of time ²⁰⁹. In PRESERVE patients were asked to produce as many words starting with a specific

letter and to generate as many expressions as possible referring to the animal category.

RAVLT is a test of short-term memory ²¹⁰. In this test participants were asked to recall as many words of the 15 noun-word list as possible, which were read to them. Participants were instructed that the order of the recall was not important. This procedure was repeated 5 times and the recall was assessed at each trial round. (list A) Afterwards an interference list of 15 new nouns was read to the subject and participants were again asked to recall as many words as possible. Participants were then asked to recall list A, measuring short recall, and were again asked to recall list A 30 min later. In the final round a recognition test was conducted which consists of presenting participants with 30 words. 15 words of the recognition test were previously presented in list A and the number of words being correctly recognized by the participant gets recorded.

All test scores were transformed into standardized scores employing normative age-scaled data. The lowest available Z-score was given, when the patient was unable to complete a task and a test score was therefore missing. This applied to 33 individual tasks across 6 participants

The scores were grouped into 4 cognitive index scores:

1. Processing speed (WAIS III coding total correct, TMT-A time to complete)
2. Executive function (TMT-B time to complete, total correct number of words for FAS)
3. Verbal memory (RAVLT 'immediate' and 'delayed' recall)
4. Global cognition (Average of the 3 domain-specific cognitive index scores)

If scores were missing due to other reasons than the patient's inability to complete the task, the domain scores were computed without that task. This applied to 14 participants (12.6% of the cohort).

2.8 Statistical analysis

Demographic variables and risk factors were compared between groups employing t-tests, Mann-Whitney test or chi-square tests as appropriate. Differences in MRI and DTI marker between baseline and 24 months were tested using paired t-test and Wilcoxon Signed Rank test. Employing an analysis of covariance analysis (ANCOVA) model with permutation, it was furthermore determined whether the change in imaging marker depended on the study site while controlling for baseline

imaging. Permutation test was employed as the permutational ANCOVA models is robust against outliers. To see individual differences between sites, a post-hoc analysis with Tukey's test was computed.

Changes in cognition across time were assessed through a linear mixed model²¹¹. Fixed effect variation was explained by visit, and random effect variation allowed for remaining inter-individual differences. The intercept of each patient's was allowed to vary with both fixed and random effects. Changes in systolic and diastolic BP were compared across visits by a linear mixed model with visit as a fixed factor and patient as a random intercept. To test whether changes in BP were dependent on the treatment group, a fixed interaction term between visit and treatment group was added to the linear model.

The primary analysis was intention-to-treat (ITT). The DTI marker for the primary endpoint was chosen as MDPH because this was shown to be the most sensitive to change in longitudinal study in severe SVD allowing the detection of treatment effects with a low sample size²⁰¹. To assess treatment effect at 24 months while accounting for baseline values, analysis of covariance (ANCOVA) was employed, with study site added as a factor. WMH lesion load was log transformed to normalize the distribution. Differences between treatment groups in lacune and CMB count were tested using an ANCOVA with permutation. For cognition linear mixed models were employed to examine whether change in cognition over time was dependent on treatment group²¹¹. The interaction between the fixed factors study visit and treatment group was included to estimate treatment effects.

A per-protocol (PP) analysis was performed, limited to subjects who reached their BP target at 3 months, defined as systolic BP <125 *mmHg* in the intensive group, and systolic BP ≥130 *mmHg* in the standard group. To determine any relationship between change in BP and MRI markers, multivariable regression was conducted between change in systolic BP and change in the imaging markers while accounting for study site. Change in BP was estimated per patient by a linear mixed model with study duration as a fixed factor and patient nested in study duration as a random factor²¹¹. The proportions of patients with side effects were compared by Fishers exact test.

3. Results

3.1 Patients

Recruitment took place from 29.02.2012-30.10.2015; follow-up was completed on 1.11.2017. Patient flow is shown in Figure 3. 111 participants (56 standard, 55 intensive) were recruited for the imaging study. One subject did not meet MRI criteria on baseline central MRI review and was withdrawn. Three died during follow up, 1 developed other serious illness and could not continue, 6 withdrew consent, and 2 were lost to follow-up. Baseline MRI was not performed in one, and follow-up MRI not performed in two. Therefore, 90 subjects remained with baseline and follow-up MRI scans. Of these, DTI sequences were available for 86 (42 standard, 44 intensive). After excluding 5 scans (3 standard, 2 intensive) of inadequate quality for DTI analysis, 81 pairs remained for the analysis (intensive 42, standard 39). There were no differences in baseline demographics, or cardiovascular risk factors between the treatment arms (Table 2).

3.2 Blinded data

3.2.1 Change in imaging markers

A significant decrease in MDPH was detectable over the two-year period. Brain volume, WMH lesion load, lacune count and CMB count also all significantly progressed (Table 3). ANCOVA with permutation showed that there was a difference between study sites when it comes to changes in WMH lesion load ($p= 0.01$) and MDPH ($p= 0.01$) while controlling for their baseline imaging measures (Figure 4). In contrast, there were no differences in study sites for change in NBV ($p= 0.12$). Post-hoc analysis with Tukey's test showed that change in MDPH was significantly different between study site 6 where patients were scanned with a Siemens MRI machine and sites where a Phillips machine was used (site 01, site 02, site 03, site 05). Change in WMH lesion load was significantly different between the two biggest study sites (site 01 and site 02).

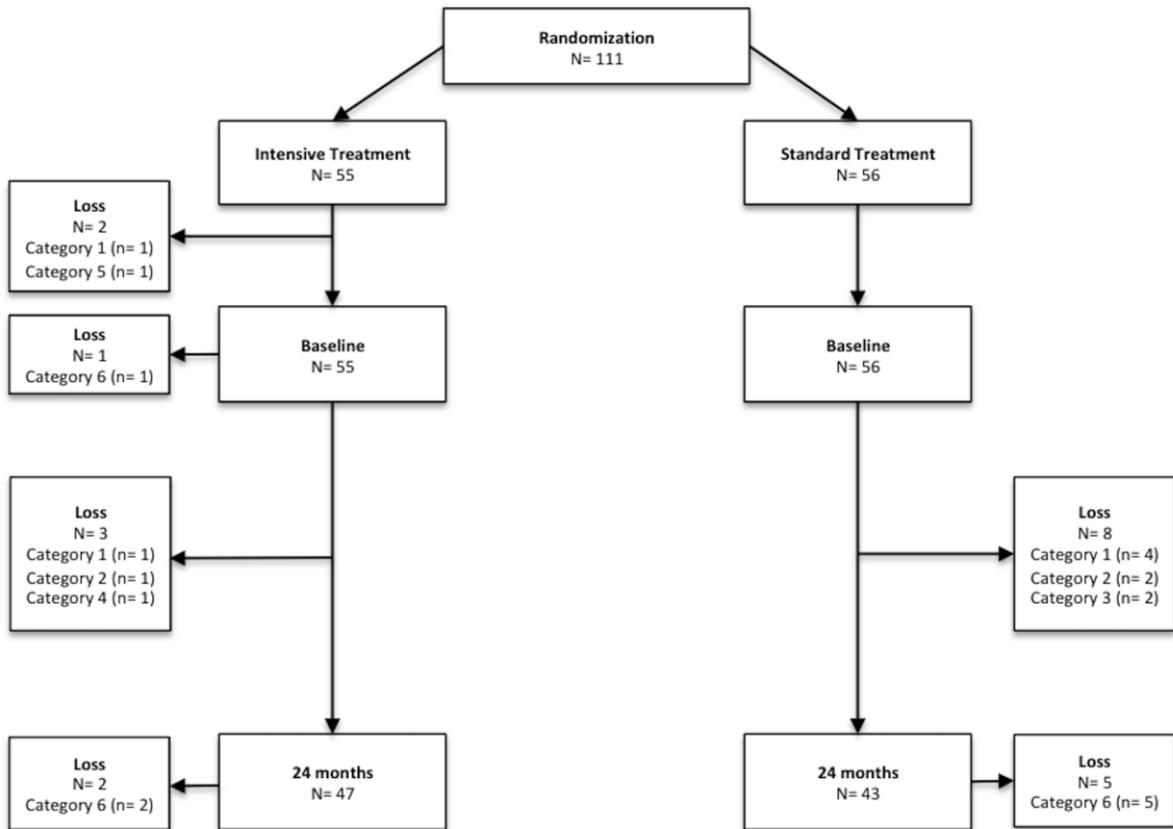


Figure 3: An overview of the patient flow. The inner boxes show the number of patients enrolled in the clinical trial at each TP after being randomized. The Primary analysis included all patients (intention-to-treat). Of the 111 patients randomized, 108 patients had a baseline scan and were cognitively tested. Between baseline and 24 months 18 patients were lost due to withdrawal of consent ($N= 5$), death ($N= 3$), no follow-up ($N= 2$), developing other serious illness ($N= 1$), and no scans being performed ($N= 7$).

TABLE 2. Baseline demographics and risk factors in the two treatment arms.

Values show the mean (standard deviation) and proportion (%) of the demographic measures between treatment groups. Differences between treatment groups were tested using the t-test for continuous variables and the chi-square test for categorical measures. There were no differences between the treatment groups on any demographics.

Demographics	Standard (N= 56)	Intensive (N= 55)	p-value
Age, (years)	67.1 (8.4)	69.1 (9.8)	0.27
Male Sex	36 (64%)	31 (57%)	0.59
Systolic BP (mmHg)	150 (13)	150 (13)	0.83
Hypercholesterolaemia	46 (82%)	39 (72%)	0.31
Current smoker	9 (16%)	8 (15%)	1
Former smoker	22 (41%)	19 (36%)	0.75
Diabetes mellitus	1 (2%)	1 (2%)	1
Myocardial infarction, CABG, or coronary angioplasty	3 (5%)	4 (7%)	0.96
Peripheral vascular disease	1 (2%)	1 (2%)	1
History of treated depression	13 (23%)	8 (15%)	0.38

BP- blood pressure, CABG- Coronary artery bypass surgery.
Hypercholesterolaemia was defined as on drug treatment

3.2.2 Change in cognition

No decrease in any cognition domain was detectable (Table 3). On the other hand, global cognition and verbal memory significantly increased over the 2 years. Changes in cognition were far from being homogeneous across the patient cohorts showing increases and decreases over time (Figure 5).

3.3 Intention to treat analysis

3.3.1 Blood Pressure changes in the Treatment Arms

Target BP difference was achieved by 3 months (intensive 127 *mmHg*, standard 140 *mmHg*), and maintained for two years (Table 4). Mean (SD) systolic BP was reduced by -15.3 (15.4) and by -23.1 (22.0) *mmHg* in the standard and intensive groups, respectively ($p < 0.001$). Systolic BP over time between treatment groups is shown in Figure 6.

3.3.2 *Clinical Endpoints on imaging*

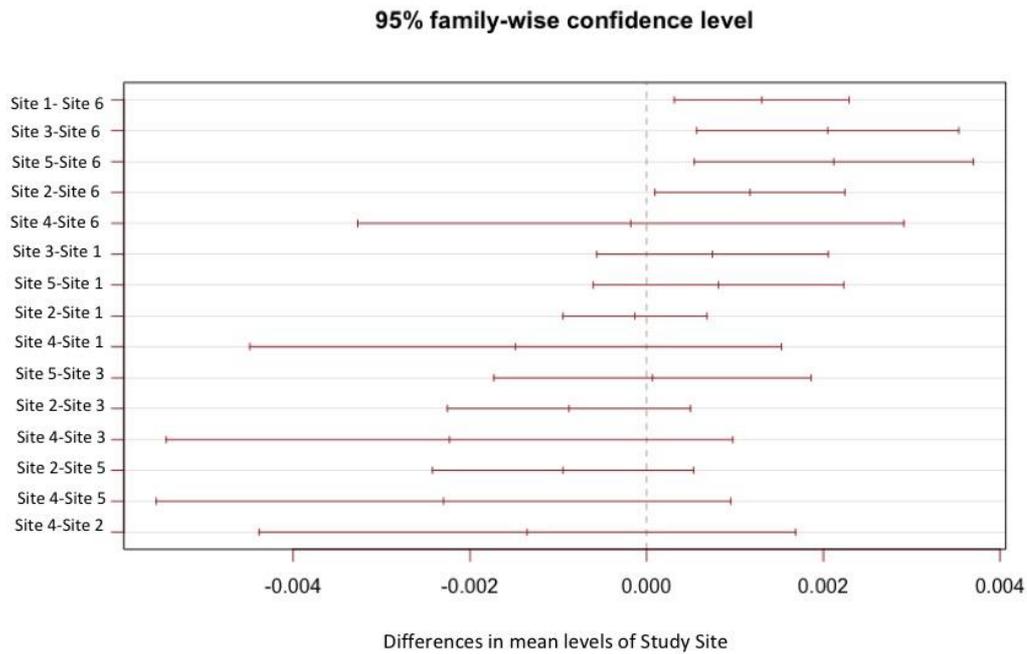
On ITT analysis there was no difference between treatment groups for the primary endpoint MDPH: standard, adjusted mean (SE)= 12.5 (0.2); intensive, 12.5 (0.02), $p=0.92$ (Table 5).

TABLE 3. Change in cognitive and MRI measures over the 2 years follow-up period in the whole imaging population. All MRI parameters changed over the 2 years. There was no decline in cognition over time. In contrast, global cognition and verbal memory significantly increased over time.

MRI parameter	Raw Mean (SD/ IQR*)			Change Raw Mean, (SD, IQR*) (%)	p-value
	Baseline	12 months	24 months		
				2 years	
MDPH ($\times 10^{-3}$ mm ² /s)	13.3 (2.4)	--	12.5 (2.3)	-0.8 (1.2) (-6.0 %)	3.39×10^{-8}
NBV (whole brain, ml)	1353.8 (112.6)	--	1322.6 (117.7)	-31.3 (30.3) (-2.3 %)	9.09×10^{-16}
WMH lesion load (% brain)	3.4 (2.2)	--	3.8 (2.5)	0.4 (0.8) (11.8 %)	9.30×10^{-6}
Lacunae (count)	4.3 (5.0)	--	6.2 (7.5)	1.9 (3.0) (44.2 %)	9.28×10^{-8}
CMB (count)	4.0 (6.25)	--	4.5 (5.0)	0.5 (1.0) (12.5 %)	0.01
Cognition	Estimated Mean (SE)			Change Estimated Mean (SE) (%)	p-value
	Baseline	12 months	24 months		
				2 years	
Global Cognition	-0.75 (0.10)	-0.61 (0.10)	-0.56 (0.10)	0.20 (0.07) (26.7 %)	0.01
Executive Function	-0.52 (0.12)	-0.43 (0.12)	-0.40 (0.12)	0.13 (0.09) (25.0 %)	0.31
Processing Speed	-0.87 (0.12)	-0.84 (0.12)	-0.76 (0.12)	0.11 (0.08) (12.6 %)	0.37
Verbal Memory	-0.85 (0.12)	-0.59 (0.13)	-0.50 (0.13)	0.35 (0.10) (41.2 %)	0.002

*Interquartile range instead of standard deviation was computed for count data.
MDPH- Normalised peak height mean diffusivity, NBV- normalized brain volume,
WMH – white matter hyperintensity, CMB- cerebral microbleeds

(A)



(B)

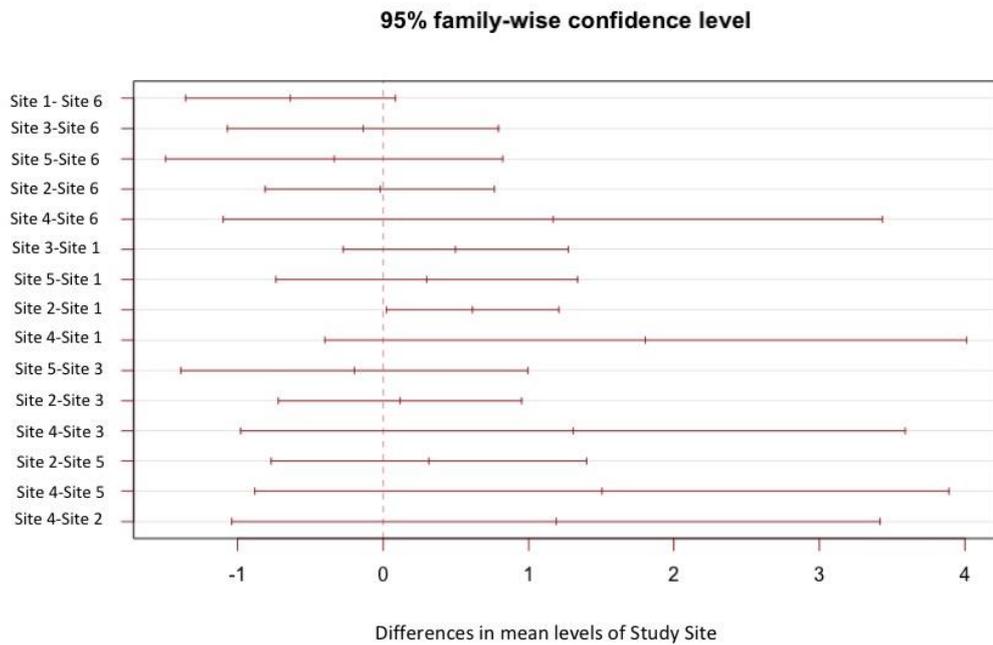


Figure 4. Testing for differences in change in markers between study sites with post-hoc Tukey's Test analysis. There were significant differences in MDPH change between the Siemens scanner site 6 and the Phillips scanner sites 1, 2, 3, 5 (panel A). Change in WMH lesion load was significantly different between the 2 biggest study sites 1 and 2 (panel B).

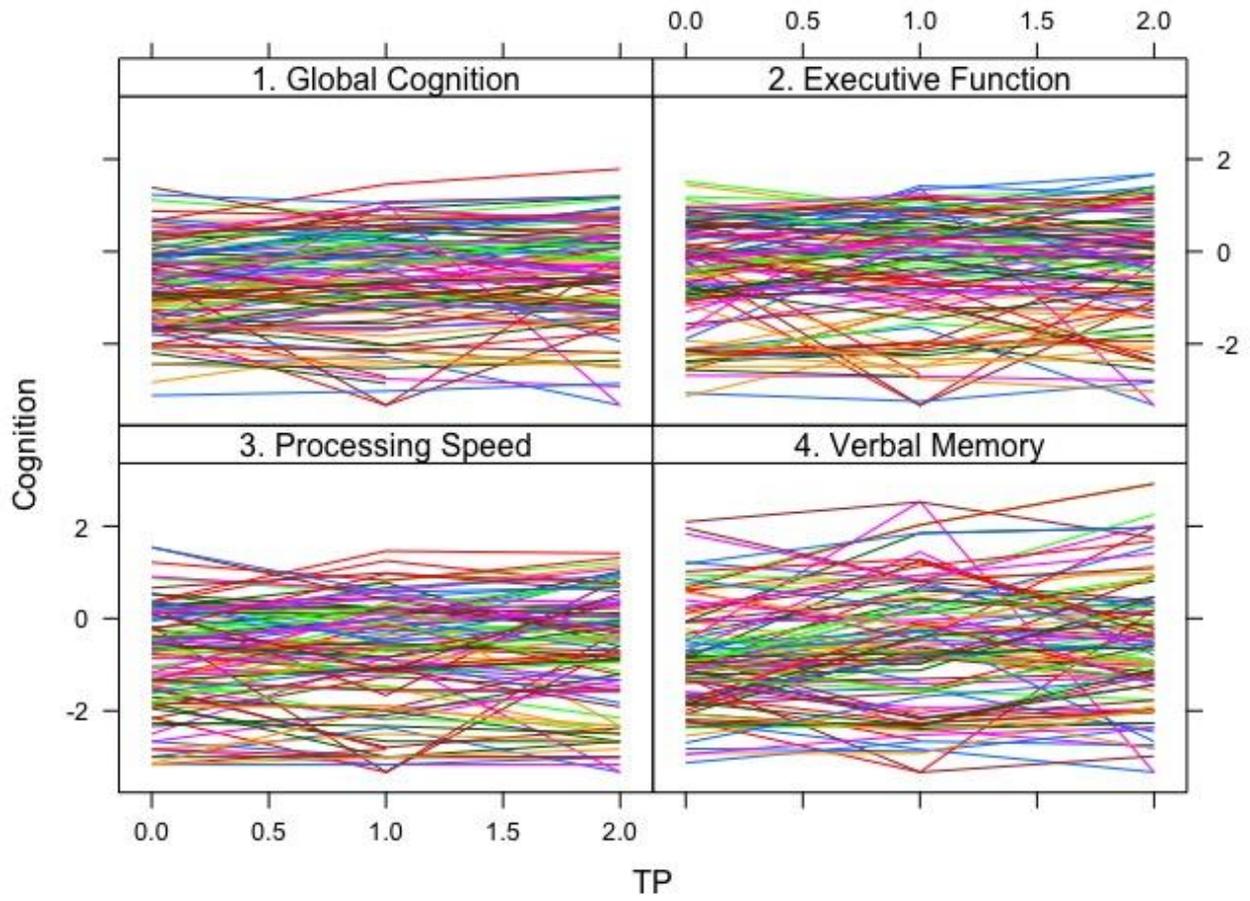


Figure 5. Cognitive index scores as a function of time (TP) grouped for all patients. Figures represent 'spagetti-plots' where the patient's scores over time is marked by one line. Cognitive scores significantly fluctuated over the 2 years showing no evidence of cognitive negative change in any index score in the imaging subgroup

TABLE 4. Systolic Blood pressure over time between treatment groups. Blood pressure in both treatment groups decreased over time. Change in BP was significantly dependent on the treatment group.

Treatment	TP	BP Mean (SD)	F-test DF= (6, 607.94)	p-value
Intensive	Baseline	149.45 (13.41)	4.41	< 0.01
	1 month	136.81 (14.47)		
	3 months	126.62 (10.74)		
	6 months	127.06 (15.18)		
	12 months	126.73 (13.32)		
	18 months	127.94 (18.36)		
	24 months	126.34 (15.24)		
Standard	Baseline	149.98 (12.88)	4.41	< 0.01
	1 month	140.24 (13.43)		
	3 months	139.74 (12.70)		
	6 months	140.16 (11.41)		
	12 months	137.98 (12.58)		
	18 months	137.90 (12.20)		
	24 months	135.10 (11.48)		

TP- time point, DF- degrees of freedom, BP- blood pressure, SD- standard deviation

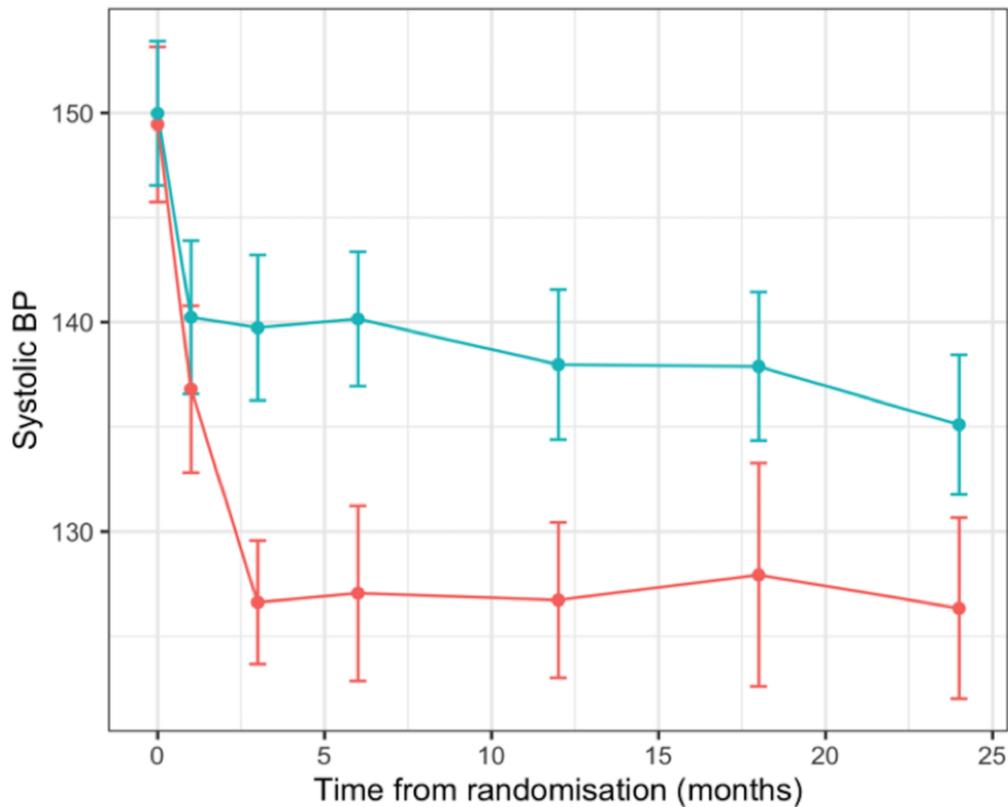


Figure 6. Systolic Blood Pressure over Time between Treatment Groups. Mean and 95% confidence intervals are shown. Change in systolic blood pressure significantly depended on the treatment group allocation.

There was no difference between treatment arms in WMH lesion load, brain volume, lacune count, or CMB count (Table 5).

3.3.4 Clinical Endpoints on cognition

There was no difference in change in global cognition, or any cognitive subdomains, between groups (Figure 7). Change in cognition did not depend on the treatment group for global cognition ($F(2, 196.64) = 0.42, p = 0.66$), executive function ($F(2, 197.04) = 1.01, p = 0.37$), processing speed ($F(2, 196.69) = 0.55, p = 0.58$) or verbal memory ($F(2, 195.37) = 1.12, p = 0.33$).

TABLE 5. MRI measures at baseline and 2 years in the two treatment groups-intention to treat analysis. There was no treatment effect between any imaging measures while accounting for study site.

MRI parameter	Raw, Mean (SD, IQR ^a)				Change Mean (SD, IQR) (%)	
	Standard		Intensive		Standard	Intensive
	Baseline	24 months	Baseline	24 months	Baseline – 24 months	Baseline – 24 months
MDPH (x10 ⁻³ mm ² /s)	13.5 (2.6)	12.7 (2.6)	13.2 (2.3)	12.3 (2.1)	-0.8 (1.4) (-5.9 %)	-0.9 (1.0) (-6.8 %)
NBV (whole brain, ml)	1368.0 (131.0)	1342.5 (135.1)	1340.9 (92.2)	1304.3 (97.0)	-25.5 (27.8) (-1.9 %)	-36.6 (31.9) (-2.7 %)
WMH lesion load (% brain)	3.2 (2.1)	3.7 (2.4)	3.5 (2.4)	3.9 (2.6)	0.5 (0.8) (15.6 %)	0.4 (0.8) (11.4 %)
Lacunae (count)	4.3 (4.1, 5.0 ^a)	5.8 (5.2, 7.0 ^a)	4.4 (5.3, 4.0 ^a)	6.5 (6.6, 8.0 ^a)	1.5 (2.4, 3 ^a) (34.9 %)	2.1 (3.6, 3 ^a) (47.7 %)
CMB (count)	4.0 (8.0, 5.0 ^a)	4.6 (9.1, 5.8 ^a)	4.0 (6.9, 6.8 ^a)	4.4 (7.7, 4.75 ^a)	0.6 (1.6, 1 ^a) (15.0 %)	0.4 (2.2, 1 ^a) (10.0 %)
	Adjusted, Mean (SE) (95% CI) at 24 months ^e				p-value ^b	
MRI parameter	Standard		Intensive			
MDPH (x10 ⁻³ mm ² /s)	12.5, (0.2) (12.1, 12.8)		12.5, (0.2) (12.2, 12.8)		0.92	
NBV (whole brain, ml)	1327.2 (45.5) (1318.2, 1336.3)		1318.3 (43.5) (1309.6, 1326.9)		0.16	
WMH lesion load (% brain) ^c	0.59 (0.02) (0.56, 0.62)		0.56 (0.02) (0.53, 0.59)		0.17	
Lacunae ^d (count)	6.1 (0.4) (5.3, 6.9)		6.2 (0.4) (5.5, 7.0)		> 0.05	
CMB ^d (count)	4.7 (0.27) (4.2, 5.2)		4.3 (0.3) (3.8, 4.8)		> 0.05	

MDPH=Mean diffusivity normalized peak height; NBV= Normalised whole brain volume; WMH= White matter hyperintensity,

^aInterquartile range additionally computed for count data.

^b Analysis of covariance testing the difference between treatment groups at 24 months while adjusting for the baseline value and study site

^c MRI variable log10 transformed

^d Permutational analysis of covariance testing the difference between treatment groups at 24 months while adjusting by the baseline value and study site

^e Means, standard errors and 95% confidence intervals adjusted by the baseline value and study site

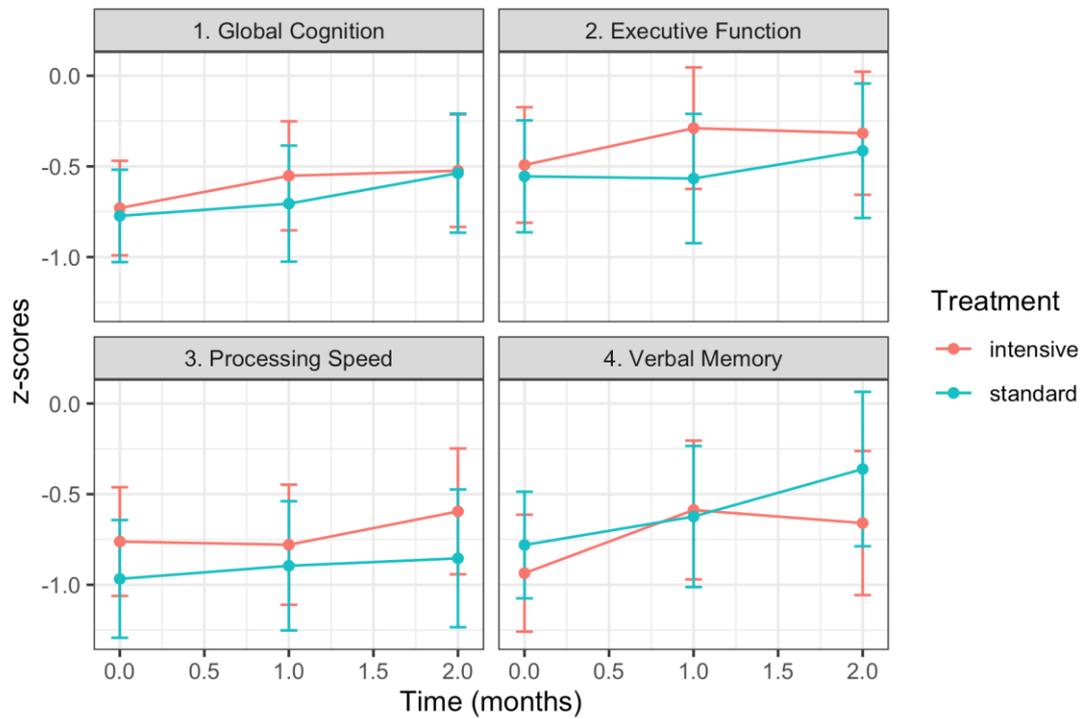


Figure 7: Cognition in the two treatment arms - intention to treat analysis. Cognition in the two treatment arms- per-protocol analysis. The means and 95% confidence intervals in each treatment group are shown per time point for each normalised cognitive index score. Testing whether the change in cognition depends on treatment groups, a linear mixed model with a fixed effect interaction visit x treatment and patients as a random intercept was employed. Change in cognition did not depend on the treatment group for Global Cognition ($F_{(2, 196.64)} = 0.42, p = 0.66$), Executive Function ($F_{(2, 197.04)} = 1.01, p = 0.37$), processing speed ($F_{(2, 196.69)} = 0.55, p = 0.58$) or verbal memory ($F_{(2, 195.37)} = 1.12, p = 0.33$).

3.4 Per protocol analysis

3.4.1 Clinical endpoints on imaging

On the PP analysis restricted to those achieving BP targets, there was no difference between treatment groups in the primary DTI endpoint or any secondary imaging endpoints (Table 6).

3.4.2 Clinical endpoints on cognition

Change in cognition did also not depend on the treatment group for global cognition ($F_{(2, 142.85)} = 0.12, p = 0.89$), executive function ($F_{(2, 143.29)} = 0.33, p = 0.72$), processing speed ($F_{(2, 142.42)} = 0.21, p = 0.81$) or verbal memory ($F_{(2, 141.33)} = 0.62, p = 0.54$) (Figure 8).

TABLE 6. Imaging marker at baseline and 2 years in the two treatment groups-per-protocol analysis. There was no treatment effect between any imaging measures while accounting for study site.

MRI parameter	Raw, Mean (SD, IQR ^a)				Change Mean (SD, IQR) (%)	
	Standard		Intensive		Standard	Intensive
	Baseline	24 months	Baseline	24 months	Baseline – 24 months	Baseline – 24 months
MDPH (x10 ⁻³ mm ² /s)	13.1 (2.2)	12.5 (2.6)	13.1 (2.5)	12.2 (2.0)	-0.06 (0.1) (-0.5 %)	-0.09 (0.1) (-0.7 %)
NBV (whole brain, ml)	1358.6 (124.9)	1330.9 (127.7)	1347.9 (94.9)	1313.4 (102.0)	-27.7 (28.4) (-2.0 %)	-34.5 (32.5) (-2.6 %)
WMH lesion load (%brain)	3.3 (2.1)	3.8 (2.5)	3.7 (2.4)	4.1 (2.6)	0.6 (0.7) (18.2 %)	0.4 (0.9) (10.8 %)
Lacunae (count)	4.6 (4.1, 5.5 ^a)	6.1 (5.0, 6.5 ^a)	4.9 (6.4, 5.5 ^a)	6.3 (7.1, 7.5 ^a)	1.5 (2.3, 3.0 ^a) (32.6 %)	1.4 (2.2, 2.0 ^a) (28.6 %)
CMB (count)	4.3 (8.6, 4.5 ^a)	4.9 (9.6, 6.0 ^a)	4.6 (7.5, 8.0 ^a)	4.9 (8.0, 5.0 ^a)	0.6 (1.7, 1.0 ^a) (14.0 %)	0.3 (2.2, 1.0 ^a) (6.5 %)
	Adjusted, Mean (SE) (95% CI) at 24 months				p-value	
MRI parameter	Standard		Intensive			
MDPH (x10 ⁻³ mm ² /s)	12.4 (0.2) (12.0, 12.8)		12.3 (0.2) (11.9, 12.7)		0.70 ^b	
NBV (whole brain, ml)	1324.8 (50.4) (1314.7, 1334.9)		1320.1 (56.3) (1309.7, 1332.2)		0.61 ^b	
WMH lesion load (%brain)	0.61 (0.02) (0.57, 0.64)		0.56 (0.02) (0.53, 0.60)		0.07 ^{b, c}	
Lacunae (count)	6.4 (0.35) (5.6, 7.1)		6.0 (0.4) (5.2, 6.8)		>0.05 ^d	
CMB (count)	5.0 (0.3) (4.4, 5.6)		4.7 (0.4) (4.0, 5.4)		>0.05 ^d	

MDPH= Mean diffusivity normalized peak height; NBV= Normalised whole brain volume; WMH= White matter hyperintensity,

^aInterquartile range additionally computed for count data.

^b Analysis of covariance testing the difference between treatment groups at 24 months while adjusting for the baseline value and study site

^c MRI variable log10 transformed

^d Permutational analysis of covariance testing the difference between treatment groups at 24 months while adjusting by the baseline value and study site

^e Means, standard errors and 95% confidence intervals adjusted by the baseline value and study site

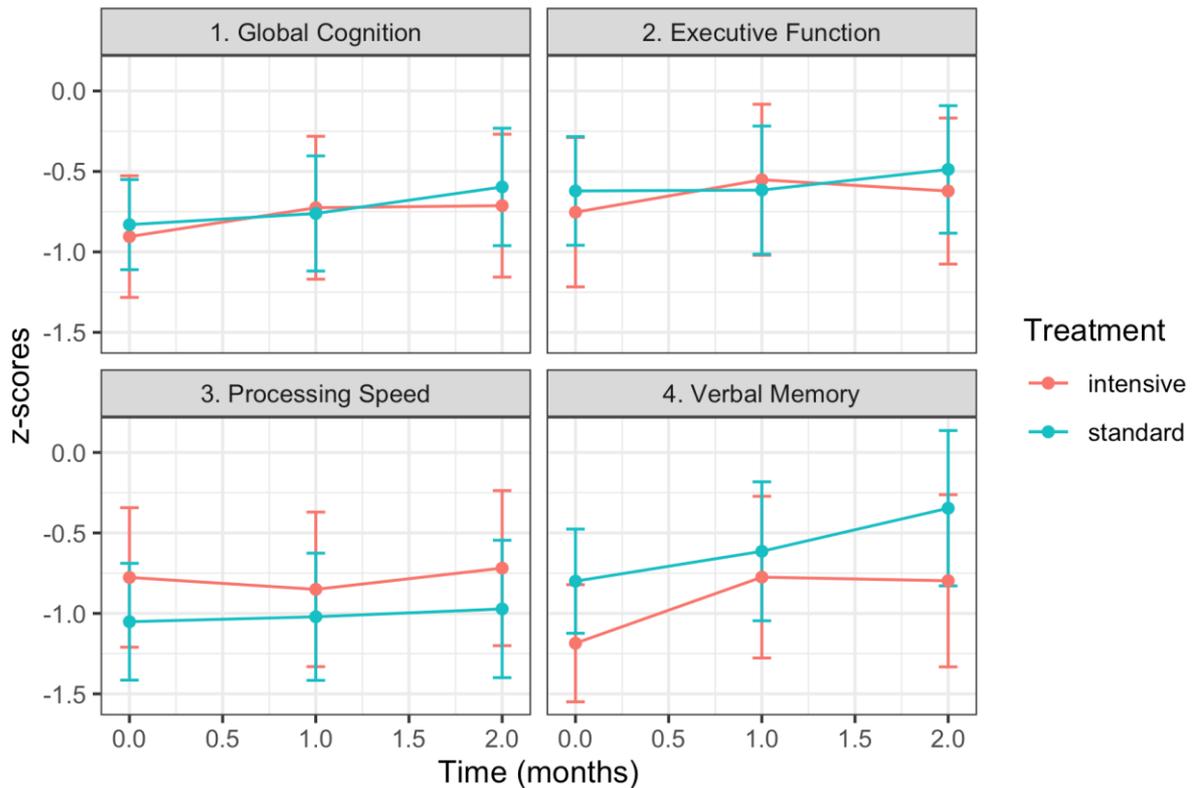


Figure 8: Cognition in the two treatment arms – per-protocol to treat analysis. Cognition in the two treatment arms- per-protocol analysis. The means and 95% confidence intervals in each treatment group are shown per time point for each normalised cognitive index score. Testing whether the change in cognition depends on treatment groups, a linear mixed model with a fixed effect interaction visit x treatment and patients as a random intercept was employed. Change in cognition did not depend on the treatment group for Global Cognition ($F_{(2, 142.85)} = 0.12$, $p = 0.89$), Executive Function ($F_{(2, 143.29)} = 0.33$, $p = 0.72$), processing speed ($F_{(2, 142.42)} = 0.21$, $p = 0.81$) or verbal memory ($F_{(2, 141.33)} = 0.62$, $p = 0.54$). Significant at $p < 0.05$.

3.5 Relationship between change in BP and change in imaging parameters

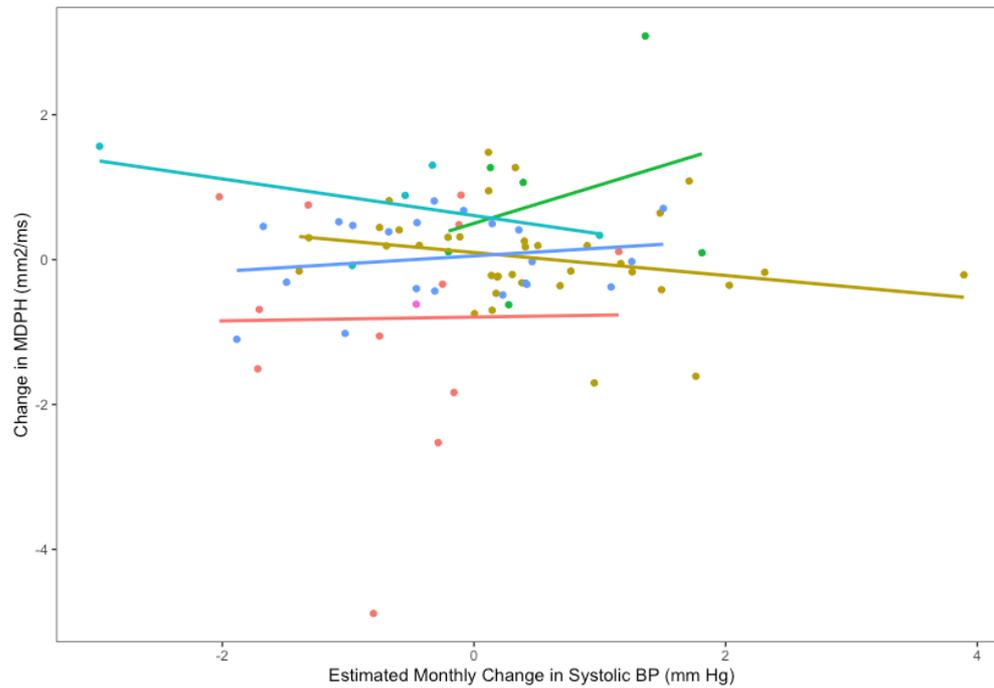
There was a reduction in the monthly systolic BP in the cohort (Mean (SD) = -0.56 (0.25)). Change in systolic BP was not associated with change in MDPH (Table 7). Lowering systolic BP was however related to less progression of WMH lesion load. ($\beta = 0.369$ (0.097), $p = 0.00027$) (Figure 9). No relationships were seen for NBV ($\beta = 0.177$ (0.106), $p = 0.10$), lacune count ($\beta = -0.082$ (0.084), $p = 0.33$) or CMB count ($\beta = 0.064$ (0.131), $p = 0.63$) (Table 7). There were 2 outliers in the regression analysis for NBV and WMH lesion load. A permutation test for linear models however confirmed previous evidence that change in systolic BP was associated with change in WMH lesion load ($p < 2e-16$) but not to change in NBV ($p = 0.15$).

TABLE 7. Association between change in systolic BP and change in DTI and MRI markers controlling for study site.
 Change in systolic BP was related to change in WMH lesion load. No association was found for the other imaging markers. Study site 6 served as a reference group in the dummy variable coding.

	Linear Regression						Poisson Regression			
	MDPH		NBV		WMH lesion load		Lacune		CMB	
	β (SE)	Adj. R	β (SE)	Adj. R	β (SE)	Adj. R	β (SE)	HL R ²	β (SE)	HL R ²
Study Site 1	0.900 (0.334), p= 0.01	0.14	0.350 (0.339), p= 0.31	0.07	-1.248 (0.308), p= 0.001	0.25	-0.603 (0.213), p= 0.005	0.23	-0.152 (0.399), p= 0.70	0.17
Study Site 2	0.866 (0.340), p= 0.01		0.426 (0.350), p= 0.23		-0.231 (0.317), p= 0.47		-1.110 (0.252), p= 1.08e-05		-0.522 (0.443), p= 0.24	
Study Site 3	1.722 (0.485), p= 0.001		0.735 (0.426), p= 0.09		-0.581 (0.382), p= 0.13		-0.712 (0.296), p= 0.02		-1.152 (0.675), p= 0.09	
Study Site 4	0.207 (0.969), p= 0.83		-1.485 (1.005), p= 0.14		1.436 (0.902), p= 0.12		1.471 (0.292), p= 4.71e-07		2.180 (0.486), p= 7.27e-06	
Study Site 5	0.866 (0.340), p= 0.01		0.836 (0.514), p= 0.11		-0.370 (0.461), p= 0.43		-1.150 (0.435), p= 0.01		-0.800 (0.782), p= 0.31	
Sys BP	-0.059 (0.106), p= 0.58		0.177 (0.106), p= 0.10		0.369 (0.097), p= 0.0003		-0.082 (0.084), p= 0.33		0.064 (0.131), p= 0.63	

Sys BP – change in systolic blood pressure, MDPH- normalised peak height mean diffusivity, NBV- normalised brain volume, WMH- white matter hyperintensity, CMB- cerebral microbleeds, Adj. R²- adjusted R², β - standardized regression coefficient, SE= standard error, HL R²- Hosmer and Lemeshow’s R2. Significant at p< 0.05

(A)



(B)

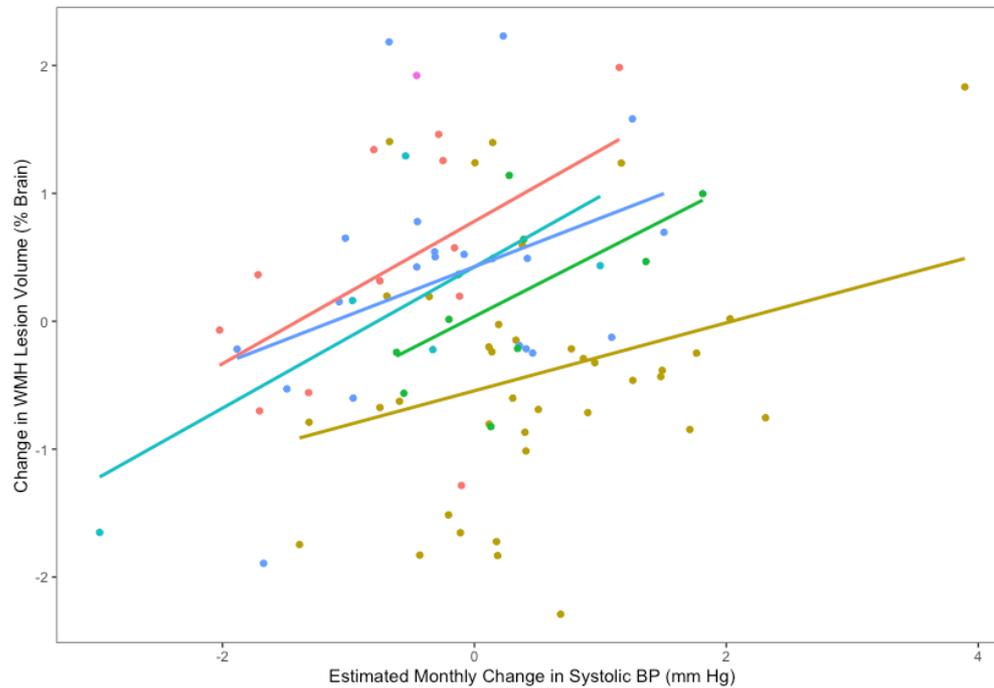


Figure 9: Relationship between change on systolic BP and change in MDPH (panel A) or WMH lesion load (panel B). Lines represent the relationships for individual centres; one centre, colored in pink, only recruited one patient and there is not represented by a line. In contrast to MDPH, there was a significant association between monthly estimated change in systolic BP and change in WMH lesion load.

3.6 Adverse events

The number of patients with any side effect was 45 in the intensive arm and 36 in the standard arm ($OR (CI) = 2.48 (0.96, 6.73)$, $p = 0.05$), and of any serious adverse events was intensive 13, standard 8 ($OR (CI) = 0.54 (0.18, 1.57)$, $p = 0.23$). There was no difference between groups in the number of falls (intensive 21 standard 14; $OR (CI) = 0.54 (0.22, 1.31)$, $p = 0.16$), or postural related dizziness (intensive 27 standard 22; $OR (CI) = 0.67 (0.30, 1.52)$, $p = 0.34$). During follow-up there were 3 strokes, and 1 death in the intensive, and 3 strokes, and 2 deaths in the standard arms

4. Discussion

PRESERVE is the first multicenter trial using a DTI measure as a primary endpoint in SVD. It confirmed data from a single centre study⁶⁶ that in a multicentre study a significant change in DTI could be detected over a two-years time period.

The unblinded results further showed that there was no difference in the rate of microstructural damage measured by DTI between the two treatment groups. This is consistent with intensive BP lowering not being associated with accelerated WM damage in patients with severe symptomatic SVD. Previous studies have shown that intensive BP treatment reduced cardiovascular events in primary prevention and stroke in secondary prevention^{195,197}.

Change in systolic BP was also not associated with the DTI marker measuring microstructural WM damage. Conversely, secondary analysis demonstrated that the degree of BP lowering correlated with less WMH lesion load progression. As WMH lesion load was a secondary endpoint, one needs to be cautious not to over interpret this finding. WMH as a clinical endpoint has been successfully employed in the SPRINT-MRI study¹⁰¹. The question really is why the DTI findings were not in line with the secondary WMH findings. One reason may be that DTI is more susceptible to between scanner differences than WMH lesion load, and this weakened its power to detect. Consistent with this, our results indicate that in contrast to WMH lesion load, change in DTI may be more influenced by the MRI scanner types (Siemens vs. Phillips). To our knowledge, this is the first study comparing the change in DTI between scanner types in SVD. Previous evidence in SVD indicated that scanner upgrades may impact DTI measures²¹² and that there may be a good intraclass reliability in DTI between 1.5T and 3T Siemens scanners¹¹³.

We were not able to detect any decline in cognition over the 2 years. Standardized scores of global cognition and verbal memory even increased over the time. There may be reasons why no change in cognition was seen. First, as shown in previous studies practice effects may have influenced the performance in neuropsychological tests²¹³. Second, patients in SVD show primary cognitive impairment specifically in domains such as executive function (EF) and processing speed (PS) but less in memory⁸⁶. But as shown previously, even more affected cognitive functions may decline slowly over many years in patients with severe SVD⁶⁴. Overall our results support the notion that cognition is not a sensitive surrogate endpoint for trials with SVD patients⁶⁶.

There are limitations in the study. First, the target number of participants to be enrolled in the study to be able to detect a treatment effect was not reached. Instead of 180 patients, 111 patients were enrolled in the imaging substudy. Due to a loss of follow-up and strict quality assessment of the images, only 81 patients were included in the ITT analysis. Second, the overall systolic BP was significantly reduced in the intensive group compared to the standard treatment group. However, BP target levels were not reached by every patient in the study. The per-protocol analysis's goal was to counter this limitation by only including patients who reached their BP level at 3 months. Similar results as in the ITT analysis was obtained. Second, it remains undetermined whether the significant differences in changes in DTI between scanner types is a systematic scanner problem. As patients were not scanned on different types of scanners twice, it is unclear how much the difference in DTI change is attributed to the different machines and how much it is due to systematic inter-group differences even when controlling for the imaging baseline measure.

To conclude, the PRESERVE multicenter trial provides some valuable lessons with regard to assessing the performance of imaging marker in a multicenter trial in SVD. First, in contrast to cognition, there was a significant overall change detectable in DTI and conventional MRI over time, even when using multiple MRI scanners. Second, changes in conventional MRI markers such as WMH and brain volume may have been more robust than in DTI markers for multicenter studies. In the next chapter the clinical relevance of the DTI and conventional MRI measures is tested and it is assessed whether DTI alone, conventional MRI markers alone or in combination with DTI may provide a strong prediction for dementia conversion across varying degrees of SVD severity.

Chapter 3

The Importance of WM Microstructural Damage in Patients with Varying Degrees of SVD Severity

1. Introduction

Conventional MRI shows characteristic features of SVD such as lacunes, WMH, CMB, PVS and brain atrophy⁶⁸. More advanced MRI techniques such as DTI further show abnormal structural WM integrity including areas outside of WMH¹⁷⁰. In the clinical context conventional markers and DTI markers have been demonstrated to be associated with impaired cognitive function and to predict dementia conversion over time^{44,86}. A sample size estimation in severe SVD further showed that DTI measures may be suitable as a surrogate markers and may serve as a clinical endpoint in a randomized clinical trial⁶⁶. However, prior to its use as a promising surrogate endpoint in a clinical trial in SVD, it first is critical to determine whether DTI actually consistently predicts cognitive impairment and dementia across populations characterized by different SVD severity. It is further important to assess the relative importance of the conventional MRI markers in the DTI prediction model as a combination of MRI markers may be needed to reflect the disease's heterogeneity beyond microstructural WM damage.

2. Methods

The OPTIMAL (OPTimising mulTImodal MRI markers for use as surrogate markers in trials of Vascular Cognitive Impairment due to cerebrAl small vesseL disease) collaboration was established to identify the most clinically relevant MRI markers in order to provide a combined MRI measure which can predict cognitive changes and progression of clinical endpoints (such as dementia) in a relatively short time period.

2.1 Participants

Six cohorts with differing degrees of SVD severity were included:

1. Severe symptomatic SVD (SCANS)
2. Severe symptomatic SVD (PRESERVE)
3. Moderate SVD (RUN-DMC)

4. Mild cognitive impairment (HARMONISATION)
5. Elderly stroke free population based cohort (ASPS-Fam)
6. Monogenic SVD (CADASIL)

All cohorts had approval by ethics committees of respective institutions. Informed consent was obtained from participants. Demographics for each cohort can be found in Table 1.

2.1.1 Severe symptomatic SVD Cohort (SCANS)

121 patients with symptomatic SVD, defined as a clinical lacunar stroke syndrome with MRI evidence of an anatomically corresponding lacunar infarct, and with confluent regions of WMH graded ≥ 2 on the modified Fazekas scale⁷³ came from 3 stroke services in South London (St George's Hospital, King's College Hospital and St Thomas' Hospital)²¹⁴. Patients were enrolled in the study at least 3 months post stroke. MRI scanning took place at baseline and over 3 yearly follow-up sessions. Images were obtained using a 1.5-T General Electric Signa HDxt MRI system. Acquisition parameters are presented in Table 2. Cognitive function was measured using well-established standardized tests, which are sensitive in detecting patterns of cognitive impairment in SVD⁴². Age-standardized test scores were used to form a measure of Global Cognition. A list of the neuropsychological test battery used can be found in Table 3.

A Dementia diagnosis was made with the Diagnostic and Statistical Manual of Mental Disorders V (DSM-V) by employing one of following criteria⁴⁴:

- A) The patient was diagnosed with dementia in a clinic or equivalent clinical service
- B) A neurologist and a clinical neuropsychologist saw all medical records and cognitive outcome measures, while being blinded to MRI and risk factor details, confirm that the clinical manifestations are in line with the DSM-5 criteria
- C) An Mini-Mental State Examination score²¹⁵ and Instrumental Activities of Daily Living (IADL) score²¹⁶ consistently below the score of 24 and below or equal to 7 respectively indicating cognitive impairment and reduced functional capabilities

Table 1: Clinical, imaging, cognition measures and sample sizes both at baseline and longitudinal in each cohort study

	Cohort					
	SCANS (n= 121)	RUN DMC (n= 503)	HARMONISATION (n= 127)	PRESERVE (n= 111)	ASPS-Fam (n= 382)	CADASIL (n= 58)
<i>Demographics</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>
Age (years)	70.01 (9.75)	65.62 (8.81)	72.23 (8.47)	68.07 (9.11)	65.43 (10.67)	47.90 (9.77)
Sex, male (%)	78 (0.65)	284 (0.57)	57 (0.45)	43 (0.39)	139 (0.40)	26 (0.45)
Included in cross-sectional analysis	yes	yes	yes	yes	yes	yes
Cohort sample size with complete DTI & MRI baseline	115	499	127	101	243	54
Included in longitudinal analysis	yes	yes	yes	no	no	no
Sample size in longitudinal analysis with complete repeated MRI & DTI	99	257	120	-	-	-
<i>Baseline complete DTI, MRI parameter</i>	<i>Mean (SD, IQR)</i>	<i>Mean (SD, IQR)</i>	<i>Mean (SD, IQR)</i>	<i>Mean (SD, IQR)</i>	<i>Mean (SD, IQR)</i>	<i>Mean (SD, IQR)</i>
MD Median (mm ² /s)	8.00e-04 (4.08-05)	8.30e-04 (3.71e-05)	8.82e-04 (6.08e-05)	7.87e-04 (4.39e-05)	7.69e-04 (3.03e-05)	8.89e-04 (1.30e-04)
Brain volume (ml)	1295.94 (92.37)	1060.82 (80.15)	1088.87 (128.83)	1349.19 (104.26)	1460.02 (144.38)	1171.88 (113.51)
WMH (IQR) (% brain or ml ³)	3.60 (2.99)*	0.87 (1.01)*	7.62 (9.45)**	3.50 (2.23)*	6.98 (5.72)**	109.52 (73.24)**
Baseline MRI count parameter	Median (range)	Median (range)	Median (range)	Median (range)	Median (range)	Median (range)
Lacune	2 (0- 27)	0 (0- 11)	0 (0- 8)	2 (0- 23)	0 (0- 7)	2.5 (0- 32)
CMB	0 (0- 144)	0 (0- 54)	0 (0- 95)	0 (0- 44)	0 (0- 9)	0 (0- 16)
<i>Baseline cognition with complete imaging</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>
Global cognition	-0.654 (0.833)	-0.015 (0.728)	-0.582 (0.892)	-0.789 (0.957)	0.033 (1.028)	-
TMT-B	-	-	-	-	-	-2.88 (4.08)

2.1.2 Multicentre severe symptomatic SVD cohort (PRESERVE)

As described in Chapter 2, 111 patients having a clinical lacunar stroke with an anatomically corresponding lacunar infarct on MRI together with confluent WMH graded as ≥ 2 on the modified Fazekas scale⁷³ were enrolled in a multicenter 2-years randomized clinical trial imaging substudy¹⁷⁶. MRI acquisition took place on eight 3-Tesla MRI scanners (3 Philips Achieva TX, 1 Philips Achieva, 1 Philips Ingenia, 1 Siemens Verio, 1 Siemens Prisma, 1 Siemens Magnetom Prisma fit). Acquisition sequence parameters across the scanners were as standardized as possible. Acquisition parameters of the multicenter study are shown in Chapter 2. 101 patients had complete MRI and DTI at baseline.

Neuropsychological test scores were age-standardized and used to create a cognitive Global Cognition index score. A list of the neuropsychological test battery used can be found in Table 3.

2.1.3 Moderate SVD Cohort (RUN DMC)

Patients with SVD, defined as the presence of lacunes and or WMH on neuroimaging, were enrolled in the Radboud University Nijmegen Diffusion Tensor and Magnetic Resonance Cohort (RUN DMC)²¹⁷. 503 patients with DTI were enrolled in the study at baseline in the year 2006. 499 patients had complete DTI and conventional MRI measures. 2 follow-up assessments took place in the years 2011 and 2015. Due to a scanner update between 2006 and 2011, only data from 2011 and 2015 were used for the longitudinal analysis. 257 patients had complete DTI and conventional MRI measures at 2011 and 2015.

MRI acquisition was based on 1.5-T Siemens Magnetom Avanto MRI machine. Acquisition parameters are listed in Table 2.

Cognitive function was assessed using validated cognitive tasks. Again, age-standardized scores were used to compute a Global Cognition index score. A list of the neuropsychological test battery used can be found in Table 3.

Dementia diagnosis was based on DSM-IV criteria²¹⁸. Probable Alzheimer's disease and vascular dementia was based on the National Institute on Aging and Alzheimer's Association criteria²¹⁹ and NINDS-AIREN criteria²²⁰ respectively.

2.1.4 Mild cognitive impairment cohort (HARMONISATION)

Patients were recruited from memory clinics at the National University Hospital and St. Luke's Hospital in Singapore²²¹. 127 MCI patients impaired in at least one

cognitive domain of a formal neuropsychological test battery, with or without a history of stroke, were included.

Imaging data were acquired on a 3T Siemens Magnetom Trio Tim system (Table 2). Baseline and 2 years follow-up DTI data and conventional MRI measures were available in 120 patients. Patients were tested on neuropsychological test batteries previously validated for elderly Singaporeans ²²². Test scores were standardized to the mean and standard deviation to form a measure of Global Cognition. More details regarding the test used can be found Table 3.

Dementia diagnosis for Alzheimer was based on National Institute on Aging and Alzheimer's Association criteria ²¹⁹. A diagnosis for vascular dementia was in accordance with the NINDS-AIREN criteria ²²⁰.

2.1.5 Elderly stroke free population based cohort (ASPS-Fam)

382 normal elderly people participated in a prospective single-center, community-based study on the cerebral effects of vascular risk factors in Graz, Austria ²²³. Inclusion criteria were being free of dementia and stroke as well as demonstrating normal neurological function.

Magnetic resonance acquisition was performed on a 3T Tim Trio whole body scanner (Table 2). 243 patients had complete DTI and MRI measures at baseline.

The cognitive index scores Global Cognition was created based on age-standardized test scores (Table 3) ²²⁴.

2.1.6 Monogenic cohort (CADASIL)

58 patients with a diagnosis of CADASIL confirmed by genetic testing or skin biopsy were enrolled ¹¹³. Previous events of transient ischemic attacks, stroke or gait disturbance were recorded. 54 had complete DTI and MRI measures and were included in analysis.

Imaging data were based on a 1.5-T GE Signa system in Munich (see Table 2 for acquisition parameters).

To measure cognitive function, the Trail-making test–B was used (Table 3). The main outcome score was normalized for age and education ²²⁵.

Table 2. MRI acquisition parameters in each single cohort study

		SCANS	RUN DMC	HARMONISATION	ASPS-FAM	CADASIL
Sequence						
T1	TR [ms]	11.5	22.50	23.00	1900	22
	TE [ms]	5	3.68	1.9	2.19	6
	Slice [mm]	1.1	1	1	1	1.2
FLAIR	TR [ms]	9000	9000	9000	1000	8402
	TE [ms]	130	84	82	69	151
	Slice [mm]	5	5	3	3	5
DTI	TR [ms]	15600	10200	6800	4900	8300
	TE [ms]	93.4	95	85	81	96
	Slice [mm]	2.5	2.5	3	3	5
	b-value [s/mm ²]	1000	900	1150	1000	1000
	Directions	25	61	61	12	41

TR= repetition time, TE= echo time, b-value= diffusion gradient strength

Table 3. Test scores used for computing Global Cognition or the trail-making test score (TMT-B)

Cohort	Cognitive Index	Task Name	Measure Description
SCANS	Global Cognition	TMT-B ²⁰⁸	Trail-making Test-B: alternating letters and numbers as quickly as possible while still maintaining accuracy
		SL-Verbal Fluency ²⁰⁹	Timed generation of words beginning with letter: FAS/ BHR
		mWCST ²²⁶	Card Sorting Test involving flexible shifting from learned dimensions
		BMIPB SOIP ⁴²	Speeded cancellation of second highest of five two-digit numbers
		Digit Symbol ²⁰⁵	Speeded transcoding task
		Grooved Pegboard ⁴²	Pick-up, rotation and placement of small pegs.
		Digit Span task ²²⁷	Immediate recall of digit strings (forwards & backwards)
		Logical Memory ²²⁷	Immediate and delayed recall of short stories
		Visual Reproduction ²²⁷	Immediate and delayed reproduction of line drawings
RUN DMC	Global Cognition	MST ²²⁸	1-letter Paper-and-Pencil Memory Scanning task: Reaction-time task on detecting memorised letters
		DSST ²²⁹	Letter–Digit Substitution Task involving match letters to numbers according to a key
		RAVLT ²¹⁰	Rey Auditory Verbal Learning Test involving verbal memory
		ROCF ^{210,230}	Rey Complex Figure Task involves reproducing a complicated line drawing, first by copying it freehand (recognition), and then drawing from memory (recall)
		Stroop ^{231,232}	Stroop Color Word Test (short form)

		VF ²³³	Verbal fluency about naming animals and professions
		VSAT ²³⁴	Verbal Series Attention Test include forward and reverse generation of arithmetic series, days of the week, and months of the year; number-letter sequencing; and auditory vigilance for a spoken target letter
PRESERVE	Global Cognition	TMT-A ²⁰⁸	Trail-making Test–A: connecting a set of 25 dots as quickly as possible while still maintaining accuracy
		TMT-B ²⁰⁸	Trail-making Test-B: alternating letters and numbers as quickly as possible while still maintaining accuracy
		WAIS-III ²²⁷	Wechsler Adult Intelligence Coding test involving coding numbers with characters according to a key
		FAS ²⁰⁹	Verbal fluency Letter subtask involving naming letters as soon as possible Verbal fluency Animals subtask involving naming animals as soon as possible
		RAVLT ²¹⁰	Rey Auditory Verbal Learning Test involving verbal memory
HARMONISATION	Global Cognition	FAB ²³⁵	Frontal Assessment Battery testing executive function
		Maze Task ²³⁶	Draw around the maze, keeping the pen tip within the maze
		Digit span task ²²⁷	Participant repeats numbers in the same order and later in the reverse order as read aloud by the examiner
		Visual memory span task ²²⁷	Patient is asked to redraw a list of stimuli presented to him
		Auditory detection task ²²²	Patients are asked to respond as quickly as possible to presented

			auditory signals
		BNT ²³⁷	Boston Naming Test. A test of confrontation naming where patients are asked to name objects presented visually as two dimensional line drawings in a booklet.
		VF ²³⁸	Verbal fluency task. Assesses spontaneous verbal production. Patients are asked to come up with as many words as possible about a predefined category in a fixed period of time
		SDMT ²³⁹	Symbol Digit Modality Test. patients are presented with rows of digits and are asked to substitute the corresponding from a key provided above
		Digit Cancellation task ²⁴⁰	The subject receives one of more digits he has to cross out from a presented list of values
		WMS-R Visual ²⁴¹	Wechsler Memory Scale— Revised (WMS-R) Visual Reproduction Copy task
		Clock Drawing task ²⁴²	Patient is asked to draw a clock
		WAIS-R Block task ²⁴³	The patient is asked to replicate a pattern of blocks that the test examiner presents to them
		Word List Recall task ²⁴⁴	List of 10 words is presented and immediate recall, delayed recall and delayed recognition is assessed
		Story Recall task ²²²	The subject is asked to recall details of a story that is read to him
		Picture Recall task ²²⁷	The subject is asked to recall details of one picture among a list of pictures that are shown to him
ASPS-Fam	Global Cognition	G-Factor ¹⁶⁶	A principal component measure involving figural and verbal memory of the Lern und Gedächtnis Test, Trail-making Test-B, Digit Span backward, Complex

			reaction time task and Purdue Pegboard Test
CADASIL	Executive function	TMT-B ²⁰⁸	Trail-making Test-B: alternating letters and numbers as quickly as possible while still maintaining accuracy

2.2 MRI processing

2.2.1 Brain Volume (BV) and WMH lesion volume

SCANS: Normalized BV was computed from native T1 images as an estimate of size of the brain relative to the skull size with SIENAX as part of the FMRIB software library ⁸⁶. WMH and brain tissues segmentations were carried out using the methods as described previously ²⁴⁵. Warped T1-weighted and FLAIR images were used to create population specific tissue probability maps (TPMs). TPMs were employed to segment native images creating tissue classes such as GM, WM, CSF and WMH. WMH volume was computed by binarizing the segmentations at a manually determined threshold. A measure of WMH, called SVDp, was calculated by taking the ratio of WMH volume to the total cerebral volume, which is composed of the sum of GM, WM and WMH ²⁴⁵.

RUN DMC: TPMs were calculated employing SPM 12 unified segmentation routine on the T1 MPRAGE images ⁹¹. The GM and WM volumes derived from the sum of the all voxel volumes were added to create a measure of total BV. WMH volumes were calculated by a semi-automatic WMH segmentation method ²⁴⁶ and the same WMH measure as in SCANS was used (SVDp) ²⁴⁵.

PRESERVE: T1W scans were segmented into GM, WM and CSF TPMs employing SPM12b (Statistical Parametric Mapping (SPM), <http://www.fil.ion.ucl.ac.uk/spm/>). Soft segmentations of the GM and WM TPMs were used to compute BV. Normalised BV (NBV) was computed from native T1 images as an estimate of size of the brain relative to the skull size with SIENAX as part of the FMRIB software library ¹⁷⁶. To obtain a measure of WMH a semi-automated program JIM (Xinapse Systems Limited; www.xinapse.com) was used to segment WMH regions. The whole brain lesion maps were used to compute a measure of lesion load as the percentage of WMH lesion volume against whole BV ¹⁷⁶.

HARMONISATION: Image preprocessing and the tissue classification algorithm have been described elsewhere ²⁴⁷. Briefly, a k-nearest-neighbor brain tissue technique was used to classify voxels into CSF, GM and normal appearing WM and volume (ml) was calculated from these measurements. WMH volumes were detected using an adapted threshold technique making use of the tissue segmentation method as described ²⁴⁸. Intracranial volume was the sum of the CSF, GM, normal WM and WMH.

ASPS-FAM: Intracranial volume was estimated by FreeSurfer 5.3 ²²³. A custom-written Interactive Data Language program was used to create WMH maps. Segmentation was performed on FLAIR images by 2 two raters ²²³.

CADASIL: Normalized BV was computed from native T1 images as an estimate of size of the brain relative to the skull size with SIENAX as part of the FMRIB software library ¹¹³. Subcortical lesions shown on FLAIR images were categorized as WMH. WMH segmentations were generated by a semiautomatic pipeline and corrected by trained raters ¹¹³.

2.2.2 Lacune count and CMB count

In all cohorts lacune and CMB were identified by trained raters blinded to all clinical data. In SCANS, RUN DMC, HARMONISATION and PRESERVE lacunes were defined as CSF filled cavity within the WM or subcortical areas between 3-15 *mm* in diameter on T1-weighted and FLAIR images ⁶⁸. In ASPS-Fam lacunes were graded as focal lesions with a maximum diameter of 10 *mm*. In CADASIL lacunes were detected on T1-weighted images with a signal identical to CSF, sharp delineation, and a diameter > 2 *mm*.

In SCANS, RUN DMC and PRESERVE CMB were identified on T2* weighted GRE (Gradient echo) images as focal spots up to 10 *mm* in diameter ^{86,132}. In HARMONISATION the rating of the CMB was done on susceptibility-weighted images employing the Brain Observer Microbleed Scale ^{116,249}. In ASPS-Fam CMB were defined as homogeneous rounded lesions with a diameter of 2-5 *mm*. In CADASIL CMB were graded on GRE sequences as rounded foci <5 *mm* in diameter.

2.2.3 DTI processing

Six commonly used WM histogram measures were computed with the DTI method already described in Chapter 2. The summary measures MD Median, MD normalised peak height (MDPH) and MD peak location (MD pkval), FA median, FA normalised peak height (FAPH) and FA peak location (FA pkval) were computed. In SCANS, PRESERVE, ASPS-Fam, HARMONY and CADASIL the eddy correct software from “FDT”, FMRIB’s Diffusion Toolbox, (<http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FDT>) was employed for DTI preprocessing. In RUN DMC diffusion data were preprocessed employing in-house developed iteratively re-weighted-least-squares algorithm named ‘PATCH’²⁵⁰. Susceptibility distortions were unwrapped by normalizing the images to the T1 images in the phase-encoding direction. For all datasets mean diffusivity (MD) and fractional anisotropy (FA) maps were created with ‘DTIFIT’. FLAIR to T1W and T1W to b0 registrations were performed and the affine transformation matrices were concatenated to produce a FLAIR-to-DTI transformation²⁰⁴. The TPMs were registered into DTI space using these transformations. A hard segmentation method was applied to generate maps of tissue classes. Histogram analysis was conducted on the MD and FA maps in all WM regions. The summary histogram measures were derived from normalized histograms with 1000 bins and an upper threshold of 2.6×10^{-3} .

2.3 Statistical analysis

2.3.1 Analyses of cross-sectional baseline data

The relationship between the DTI and cognition was tested using a linear regression controlling for clinical markers (age, gender and premorbid IQ or education). Using a model decomposition method²⁵¹, the amount of variance in cognition explained was determined. In the CADASIL sample, one patient was excluded as an outlier from the regression analysis and the dependent variable was power transformed to meet the statistical assumptions underlying the linear regression model.

To determine the best model fit, Akaike information criterion (*AIC*) together with the adjusted variance (*Adj R²*) was computed for each cohort. *AIC* assesses the model fit with penalization for the number of parameters. The *AIC* was then determined for the following models: clinical markers only model (clinical), clinical markers plus DTI (clinical-DTI), clinical markers plus conventional MRI markers (BV, WMH, lacune and CMB) (Clinical-MRI), and clinical markers with both conventional MRI markers and

DTI (clinical-MRI-DTI). Prior to the analysis, WMH, lacune and CMB counts were log10 transformed to normalize their distributions. To account for multicollinearity, the multiple regression models were checked with the variance inflation factor (*VIF*) statistic. All predictors in the models were below a *VIF* value threshold ($VIF < 10$) that should cause concern ²⁵².

To estimate the relative importance of each predictor in the clinical-MRI-DTI model in terms of explaining the overall variance in cognitive function, the model decomposition method was employed ²⁵¹. Again multicollinearity was checked with *VIF*.

2.3.2 Analyses of longitudinal data

The association between baseline DTI and dementia conversion was tested using a Cox regression model in SCANS, RUN DMC and HARMONISATION. The variables age, gender, and premorbid IQ or years of education were added as confounders. To assess the models' explained variance, the Nagelkerke R^2 measure was calculated. Receiver operating characteristic curve (*ROC*) and the Area under the curve (*AUC*) were computed ²⁵³.

Change in DTI, BV and WMH over time was determined. In studies with two MRI assessments (RUN DMC and HARMONISATION), the difference between the two time points was tested using paired t-tests. In SCANS, which had up to 4 MRI time points, change was estimated by a linear mixed model ²¹¹. The intercept and slope of each participant's linear trajectory were allowed to vary with both fixed and random effects. Fixed effect variation was accounted for by time, and random effect variation allowed for remaining inter-individual differences. The average fixed effects slope represent the average annualized change rate for a given measure. WMH was log 10 transformed prior to the linear mixed model computation. Change in lacunes and CMB were categorized as dichotomous in terms of change vs. no change in counts over time. Follow-up observations post dementia diagnoses were removed prior to the mixed model computation.

Employing a Cox regression in SCANS, or logistic regression in RUN DMC and HARMONISATION, the association between the change in DTI and dementia conversion was tested while accounting for the clinical markers. Using *AIC* the DTI model was compared to the clinical model, to the clinical-MRI model and to clinical-MRI-DTI model. The *AUC* was furthermore computed for each model and compared within cohorts using the pROC library ²⁵³.

3. Results

The cohorts' demographic characteristics are shown in Table 1.

3.1 Cross-sectional association

The DTI marker MD Median showed consistent associations, being related to cognition in all cohorts (Table 4), and was taken forward as the DTI marker used in further analyses. The DTI-clinical model explained 0.454- 0.515 of the model's variance in the single-centre non-monogenic cohorts, but less variance in the multicentre cohort ($Adj R^2 = 0.373$) (Table 5). The adjusted explained variance in the CADASIL cohort was 0.235 due the low variance explained by the clinical model in this cohort. DTI's normalized R^2 contribution was highest in CADASIL.

The variance in cognition explained by the different models is shown in Table 6. The additional variance explained by adding conventional MRI markers to the DTI-clinical model varied but was greatest in SCANS and CADASIL, while little or no additional variance was added in HARMONISATION and ASPS-Fam (Fig 1). Relative contributions of the different imaging markers of SVD to cognition in the clinical -DTI-MRI model were determined (Table 7). This varied according to the SVD's severity. In single-centre sporadic SVD (SCANS; RUN DMC) and CADASIL cohorts, brain volume (BV) explained most of the variance among all imaging markers. The importance of lacune count as factor varied according to the vascular disease severity with being high in SCANS, PRESERVE and CADASIL and low in RUN DMC, HARMONISATION and ASPS-Fam (Table 7). Of the clinical markers, age was the most important predictor in the ASPS-Fam, whereas education or premorbid IQ explained most variance in the other non-monogenic cohorts.

Table 4. Cross-sectional analysis between DTI measures and Global Cognition or TMT-B (CADASIL). Values show standardised regression coefficients: β (SE) for predictor variables in linear regression models of Global Cognition or TMT-B. Adjusted R² refers to the overall explained variance adjusted by the number of predictors in the model. DTI's normalized R² contribution to the overall model together with the 95% confidence interval (CI) is also shown.

	SCANS	RUN DMC	HARMONISATION	PRESERVE	ASPS=Fam	CADASIL
MD Median	-0.232 (0.068), p= 0.001 Adj. R ² = 0.478, Norm. Cont R ² (95% CI)= 0.170 (0.030, 0.330)	-0.218 (0.040), p= 8.76e-08 Adj. R ² = 0.454, Norm. Cont R ² (95% CI)= 0.237 (0.167, 0.317)	-0.344 (0.077), p= 1.98e-05 Adj. R ² = 0.456, Norm. Cont R ² (95% CI)= 0.192 (0.076, 0.336)	-0.410 (0.087), p= 9.19e-06 Adj. R ² = 0.373, Norm. Cont R ² (95% CI)= 0.297 (0.105, 0.464)	-0.149 (0.055), p= 0.008, Adj. R ² = 0.516, Norm. Cont R ² (95% CI)= 0.173 (0.093, 0.273)	-0.516 (0.177), p= 0.0055 Adj. R ² = 0.235, Norm. Cont R ² (95% CI)= 0.650 (0.235, 0.812)
FA Median	0.267 (0.069), p= 0.0002 Adj. R ² = 0.492, Norm. Cont R ² (95% CI)= 0.200 (0.066, 0.342)	0.098 (0.034), p= 0.004 Adj. R ² = 0.431, Norm. Cont R ² (95% CI)= 0.025 (0.002, 0.072)	0.268 (0.072), p= 0.0003 Adj. R ² = 0.433, Norm. Cont R ² (95% CI)= 0.171 (0.057, 0.324)	0.400 (0.082), p= 4.08e-06 Adj. R ² = 0.383, Norm. Cont R ² (95% CI)= 0.394 (0.156, 0.562)	0.119 (0.051), p= 0.020, Adj. R ² = 0.512, Norm. Cont R ² (95% CI)= 0.100 (0.037, 0.203)	0.515 (0.162), p= 0.0027 Adj. R ² = 0.257, Norm. Cont R ² (95% CI)= 0.677 (0.299, 0.789)
MDPH	0.315 (0.067), p= 7.17e-06 Adj. R ² = 0.520, Norm. Cont R ² (95% CI)= 0.231 (0.091, 0.389)	0.210 (0.043), p= 1.25e-06 Adj. R ² = 0.448, Norm. Cont R ² (95% CI)= 0.244 (0.175, 0.314)	0.338 (0.077), p= 2.36e-05 Adj. R ² = 0.455, Norm. Cont R ² (95% CI)= 0.188 (0.082, 0.337)	0.316 (0.090), p= 0.001 Adj. R ² = 0.314, Norm. Cont R ² (95% CI)= 0.195 (0.032, 0.371)	0.009 (0.050), p= 0.865 Adj. R ² = 0.500, Norm. Cont R ² (95% CI)= 0.026 (0.008, 0.075)	0.451 (0.179), p= 0.015 Adj. R ² = 0.202, Norm. Cont R ² (95% CI)= 0.609 (0.207, 0.747)
FAPH	-0.249 (0.069), p= 0.001 Adj. R ² = 0.484, Norm. Cont R ² (95% CI)= 0.122 (0.028, 0.263)	-0.093 (0.034), p= 0.006 Adj. R ² = 0.430, Norm. Cont R ² (95% CI)= 0.012 (0.002, 0.044)	-0.239 (0.076), p= 0.002 Adj. R ² = 0.416, Norm. Cont R ² (95% CI)= 0.156 (0.053, 0.304)	-0.374 (0.083), p= 1.72e-05 Adj. R ² = 0.364, Norm. Cont R ² (95% CI)= 0.357 (0.126, 0.545)	0.022 (0.050), p= 0.664 Adj. R ² = 0.500, Norm. Cont R ² (95% CI)= 0.020 (0.006, 0.072)	-0.454 (0.154), p= 0.005 Adj. R ² = 0.237, Norm. Cont R ² (95% CI)= 0.630 (0.209, 0.797)
MD pkval	-0.178 (0.069), p= 0.011 Adj. R ² = 0.456, Norm. Cont R ² (95% CI)= 0.110 (0.010, 0.280)	-0.146 (0.040), p= 2.69e-04 Adj. R ² = 0.437, Norm. Cont R ² (95% CI)= 0.174 (0.112, 0.249)	-0.099 (0.073), p= 0.177 Adj. R ² = 0.377, Norm. Cont R ² (95% CI)= 0.027 (0.003, 0.152)	-0.365 (0.089), p= 8.53e-05 Adj. R ² = 0.343, Norm. Cont R ² (95% CI)= 0.254 (0.076, 0.444)	-0.109 (0.054), p= 0.044 Adj. R ² = 0.509, Norm. Cont R ² (95% CI)= 0.130 (0.065, 0.216)	-0.610 (0.173), p= 0.001 Adj. R ² = 0.241, Norm. Cont R ² (95% CI)= 0.757 (0.214, 0.904)
FA pkval	0.228 (0.069), p= 0.001 Adj. R ² = 0.476, Norm. Cont R ² (95% CI)= 0.053 (0.009, 0.123)	0.006 (0.037), p= 0.864 Adj. R ² = 0.421, Norm. Cont R ² (95% CI)= 0.037 (0.017, 0.078)	0.095 (0.077), p= 0.216 Adj. R ² = 0.375, Norm. Cont R ² (95% CI)= 0.021 (0.005, 0.123)	0.257 (0.088), p= 0.004 Adj. R ² = 0.288, Norm. Cont R ² (95% CI)= 0.197 (0.016, 0.427)	0.086 (0.048), p= 0.072 Adj. R ² = 0.507, Norm. Cont R ² (95% CI)= 0.057 (0.010, 0.142)	0.159 (0.144), p= 0.274 Adj. R ² = 0.0655, Norm. Cont R ² (95% CI)= 0.305 (0.028, 0.656)

Table 5. Cross-sectional regression between MD Median together with the clinical markers and cognitive function. Values show the standardised regression coefficients: β (SE) for predictor variables in linear regression models of Global Cognition or of TMT-B.. Adjusted R^2 refers to the overall explained variance adjusted for the number of predictors in the model. Median MD's R^2 contribution to the overall model together with the 95% confidence interval (CI) is also shown. The contribution was significantly higher in CADASIL than in any other cohort. MD Median explained a significant proportion of variance in all cohorts.

	SCANS	RUN DMC	HARMONISATION	PRESERVE	ASPS-Fam	CADASIL
Age	-0.121 (0.070) p= 0.088	-0.326 (0.040), p= 4.10e-15	-0.107 (0.075), p= 0.158	0.362 (0.084), p= 4.33e-05	-0.459 (0.052), p < 2e-16	-0.001 (0.140), p= 0.993
Sex	-0.272 (0.147) p= 0.068	-0.148 (0.068), p= 0.030	0.639 (0.144), p= 2.00e-05	0.146 (0.172), p= 0.396	0.008 (0.101), p= 0.941	-0.010 (0.239), p= 0.969
NART/ Education	0.577 (0.067) p= 7.61e-14	0.414 (0.034), p < 2e-16	0.449 (0.069), p= 2.20e-09	0.341 (0.084), p= 1.06e-04	0.350 (0.052), p= 1.45e-10	0.199 (0.137), p= 0.153
MD Median	-0.232 (0.068), p= 0.001	-0.218 (0.040), p= 8.76e-08	-0.344 (0.077), p= 1.98e-05	-0.410 (0.087), p= 9.19e-06	-0.149 (0.055), p= 0.008 ,	-0.516 (0.177), p= 0.006
Model: Overall Adj. R^2	0.478,	0.454	0.456,	0.373	0.516,	0.235
DTI's normalized contribution (%) (95% CI) to the overall model's R^2 variance	0.170 (0.030- 0.330)	0.237 (0.167- 0.317)	0.192 (0.076- 0.336)	0.297 (0.105- 0.464)	0.173 (0.093- 0.273)	0.650 (0.235- 0.812)

MD Median= median of the mean diffusivity histogram measure, NART= premorbid IQ score, CI (95%)= 95% confidence interval, R^2 = explained variance

Table 6 Predictive model comparisons at baseline. The Adj. R² was the highest in the Clinical-MRI-DTI model in RUN DMC, PRESERVE and ASPS-Fam. The AIC of the Clinical-MRI-DTI was lowest (best model fit) for SCANS, RUN DMC and CADASIL. Adding conventional MRI measures to the clinical or clinical-DTI models significantly increased the overall explained variance (Adj R²) in SCANS and CADASIL.

Model summary measure	Model	SCANS	RUN DMC	HARMONISATION	PRESERVE	ASPS-Fam	CADASIL
AIC	Clinical	283.45	1148.17	304.77	261.09	496.22	131.10
	Clinical-DTI	255.59	1121.22	287.73	243.55	503.45	124.44
	Clinical-MRI	204.45	1119.72	303.27	247.49	493.74	123.15
	Clinical-MRI-DTI	199.54	1113.17	293.52	246.02	502.68	122.64
Adj R²	Clinical	0.420	0.422	0.373	0.229	0.502	0.107
	Clinical-DTI	0.478	0.454	0.456	0.373	0.516	0.235
	Clinical-MRI	0.570	0.459	0.398	0.365	0.513	0.292
	Clinical-MRI-DTI	0.564	0.467	0.447	0.391	0.523	0.311

AIC= Akaike information criterion, Adj R²= adjusted explained variance of the model, Clinical= age, sex, premorbid IQ or education; Clinical-DTI= age, sex, premorbid IQ or education and MD Median, Clinical-MRI= age, sex, premorbid IQ or education, brain volume, white matter hyperintensity volume, lacune count and cerebral microbleeds, Clinical-MRI-DTI= age, sex, premorbid IQ or education, brain volume, white matter hyperintensity volume, lacune count, cerebral microbleeds and MD Median

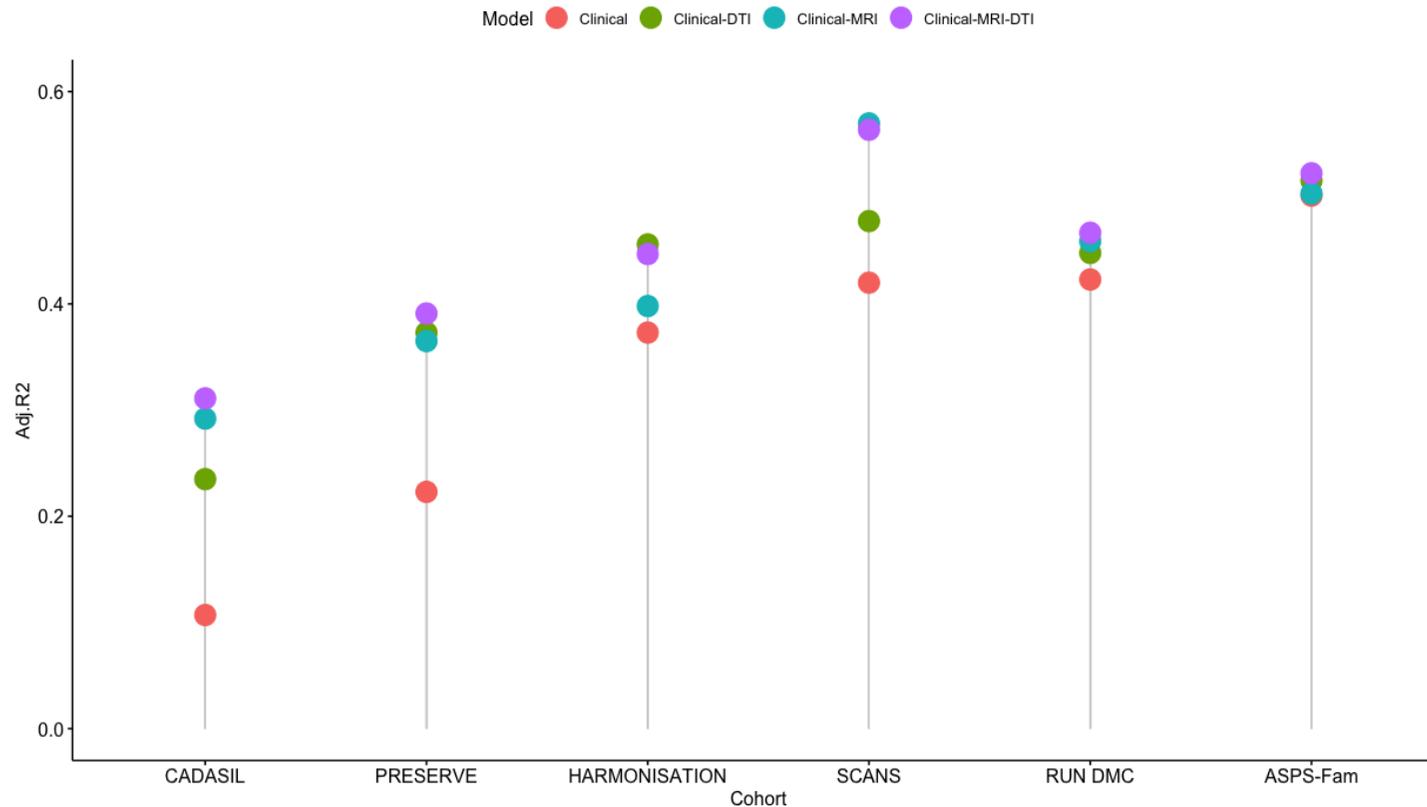


Figure 1. Adjusted R² variance for clinical risk model, DTI model and multimodal imaging model predicting baseline cognition. The adjusted variance was highest in SCANS and lowest in CADASIL. In the single-center cohorts adjusted R² of the Clinical-MRI-DTI models were significantly higher than Clinical-DTI models in SCANS and CADASIL compared to RUN DMC, HARMONISATION and ASPS-Fam.

Clinical= age, sex, premorbid IQ or education; Clinical-DTI= age, sex, premorbid IQ or education and MD Median, Clinical-MRI= age, sex, premorbid IQ or education, brain volume, white matter hyperintensity volume, lacune count and cerebral microbleeds, Clinical-MRI-DTI= age, sex, premorbid IQ or education, brain volume, white matter hyperintensity volume, lacune count, cerebral microbleeds and MD Median

Table 7 Contribution of the individual marker to the variance in global cognition in the clinical-MRI-DTI model. In single-center SVD (SCANS; RUN DMC) and CADASIL cohorts, brain volume (BV) explained most of the variance among all imaging markers. The importance of lacune count as factor varied according to the vascular disease severity with being high in SCANS, PRESERVE and CADASIL and low in RUN DMC, HARMONISATION and ASPS-Fam. In all but ASPS-Fam premorbid IQ explained most of the variance among the clinical markers. The importance of the clinical markers was low in CADASIL.

Clinical-MRI-DTI model- Contribution of individual marker to the variance in cognition												
(R², Bootstrap CI (95%) Metrics are normalized to sum to 100%)												
	<i>Cohorts</i>											
	<i>SCANS</i>		<i>RUN DMC</i>		<i>HARMONISATION</i>		<i>PRESERVE</i>		<i>ASPS-Fam</i>		<i>CADASIL</i>	
Multiple regression	R2 variance estimate	CI (95%)										
<i>MD Median</i>	0.062	0.015, 0.123	0.130	0.085, 0.183	0.151	0.062, 0.267	0.129	0.037, 0.257	0.135	0.067, 0.217	0.256	0.083, 0.435
<i>BV</i>	0.172	0.069, 0.319	0.166	0.115, 0.224	0.034	0.006, 0.120	0.029	0.016, 0.081	0.051	0.010, 0.109	0.291	0.072, 0.480
<i>WMH</i>	0.007	0.004, 0.045	0.078	0.053, 0.111	0.061	0.014, 0.158	0.040	0.014, 0.135	0.073	0.037, 0.127	0.039	0.017, 0.121
<i>Lacune count</i>	0.085	0.022, 0.169	0.016	0.005, 0.040	0.007	0.004, 0.044	0.120	0.018, 0.275	0.033	0.010, 0.073	0.207	0.042, 0.376
<i>CMB count</i>	0.017	0.005, 0.067	0.022	0.006, 0.054	0.008	0.004, 0.057	0.096	0.013, 0.253	-	-	0.027	0.012, 0.154
<i>Age</i>	0.027	0.006, 0.098	0.212	0.150, 0.273	0.083	0.018, 0.194	0.222	0.067, 0.358	0.430	0.314, 0.542	0.055	0.016, 0.288
<i>Sex, male</i>	0.011	0.002, 0.079	0.004	0.003, 0.019	0.179	0.063, 0.314	0.047	0.010, 0.171	0.012	0.005, 0.041	0.033	0.014, 0.152
<i>NART/ Education</i>	0.620	0.405, 0.756	0.372	0.283, 0.456	0.478	0.294, 0.612	0.257	0.095, 0.382	0.268	0.154, 0.379	0.092	0.013, 0.233

MD Median= median of the mean diffusivity histogram measure, BV= brain volume, WMH= white matter hyperintensity volume, CMB= cerebral microbleeds, NART= premorbid IQ score, CI (95%)= 95% confidence interval, R²= explained variance of the model

3.2 Longitudinal association with baseline imaging marker

Dementia incidence together with DTI was available for 3 studies; SCANS, 5 year follow-up, 20 (17%) dementia cases, all vascular (VD); RUN DMC, 9 years follow-up, 65 (13%) dementia cases, (21 Vascular dementia (VD), 30 AD, 10 AD/VD, 1 frontotemporal, 1 Lewy body, 1 progressive supranuclear palsy, 1 unknown) and HARMONISATION, 2 years follow-up, 23 (18 %) dementia cases, (3 VD, 20 AD).

In all 3 cohorts higher MD Median was associated with higher risk of dementia after controlling for the clinical markers (SCANS: HR (95% CI)= 2.048 (1.438, 2.918), p = 7.1e-05, AUC = 0.794; RUN DMC: HR (95% CI)= 1.364 (1.060, 1.755), p = 0.016, AUC = 0.825; HARMONISATION: HR (95% CI)= 1.784 (1.085, 2.935), p = 0.023, AUC = 0.757) (Table 8). MD Median and MDPH were the only two histogram markers predicting dementia across all three cohorts (Table 8).

3.3 Longitudinal association with change in imaging marker

Complete longitudinal data with repeat conventional MRI and DTI measures were available for SCANS (N = 99), RUN DMC (N =257) and HARMONISATION (N = 120) with dementia cases: SCANS 18; RUN DMC 12; HARMONISATION 21. There was a change in DTI in all cohorts (Table 9).

Change in DTI over 3 years predicted dementia conversion over 5 years in SCANS (HR (95% CI)= 2.588 (1.663, 4.027), AUC = 0.785) but not in RUN DMC (OR (95% CI)= 0.935 (0.498, 1.667), p = 0.825, AUC = 0.891) or in HARMONISATION (OR (95% CI)= 1.573 (0.998, 2.597), p = 0.109, AUC = 0.738) (Table 10). Although adding DTI increased the AUC from 0.684 to 0.738 in HARMONISATION, this increase was not significant. There was no increase in prediction in RUN DMC (Table 11). In SCANS, adding the conventional MRI measures added prediction over the clinical-DTI model with the AUC increasing from 0.785 to 0.872 (p = 0.05). The AIC indicated that the clinical-MRI-DTI model had the best fit. In the other 2 cohorts adding conventional MRI markers added little on top of the less complex models (Table 11).

Table 8. Longitudinal analysis between DTI baseline measures and dementia. Values show the standardized regression coefficients β (SE) for the baseline predictor variables in a Cox regression or logistic regression models of dementia conversion. The hazard ratio (HR) or odds ratio (OR) together with the confidence interval (CI) are shown. The model parameters Nagelkerke's R^2 ($Ng R^2$) or Hosmer and Lemeshow's R^2 (R^2_L) give an estimated amount of variation in the dependent variable explained by the model. The Area under the curve (AUC) evaluates how well the model classifies dementia conversion and no-dementia conversion at all possible cutoffs respectively. All regression models controlled for the effects of age, gender and NART-IQ or education.

	Baseline marker predicting dementia conversion					
Cohort	DTI markers	β (SE)	P-value	HR (95% CI)	Ng R^2 / R^2_L	AUC
SCANS	MD Median	0.717 (0.181)	7.1e-05	2.048 (1.438, 2.918)	0.206	0.794
	FA Median	-0.750 (0.251)	0.0028	0.472 (0.289, 0.773)	0.171	0.791
	MDPH	-0.959 (0.256)	0.0002	0.384 (0.232, 0.634)	0.212	0.811
	FAPH	0.782 (0.237)	0.001	2.186 (1.374, 3.478)	0.179	0.754
	MD pkval	0.596 (0.192)	0.002	1.816 (1.248, 2.644)	0.174	0.790
	FA pkval	-0.751 (0.263)	0.004	0.472 (0.282, 0.789)	0.168	0.771
	RUN DMC	MD Median	0.310 (0.136)	0.016	1.364 (1.060, 1.755)	0.165
FA Median		0.182 (0.112)	0.105	0.834 (0.670, 1.038)	0.160	0.817
MDPH		-0.812 (0.172)	2.41e-06	0.444 (0.317, 0.622)	0.193	0.847
FAPH		0.192 (0.121)	0.112	1.212 (0.956, 1.536)	0.160	0.818

	MD pkval	-0.080 (0.135)	0.554	0.923 (0.708, 1.203)	0.156	0.811
	FA pkval	-0.194 (0.121)	0.108	0.824 (0.650, 1.044)	0.160	0.814
HARMONISATI ON	MD Median	0.579 (0.253)	0.023	1.784 (1.085, 2.935)	0.096	0.761
	FA Median	-0.259 (0.253)	0.307	0.772 (0.470, 1.268)	0.067	0.722
	MDPH	-0.622 (0.283)	0.028	0.537 (0.308, 0.935)	0.095	0.761
	FAPH	0.309 (0.240)	0.198	1.362 (0.851, 2.180)	0.072	0.738
	MD pkval	0.045 (0.221)	0.840	1.046 (0.678, 1.612)	0.060	0.709
	FA pkval	0.187 (0.214)	0.382	1.206 (0.792, 1.836)	0.065	0.690

FA= fractional anisotropy, MD= mean diffusivity, pkval= peak value of the histogram distribution, PH= normalised peak height of the histogram distribution, β = standardised regression coefficient, SE= standard error of the regression coefficient, HR= hazard ratio, Ng R^2 = Nagelkerke's R^2 , R^2_L = Hosmer and Lemeshow's R^2 , AUC= area under the curve, 95% CI= 95% confidence interval

Table 9. DTI all Histogram and MRI change over time. Values show the change in MD Median and MRI markers over time. Change in the imaging markers were estimated employing a linear mixed model in SCANS. In RUN DMC and HARMONISATION paired t-tests were used to test the absolute difference in imaging measures between two time points. The *p*-value below 0.05 indicates that there was a significant change over time. The number of patients showing in an increase in lacune count and CMB count over time is also shown.

	Cohort					
	SCANS N= 99		RUN DMC N= 257		HARMONISATION N= 120	
MRI parameter change	Est. annual mean change (SE)	p-value	Abs. mean change (SD)	p-value	Abs. mean change (SD)	p-value
MD Median (mm ² /s)	5.37e-06 (5.42e-07)	2.97e-06	3.34e-06 (1.26e-05)	3.15e-05	2.30e-05 (3.69e-05)	4.04e-10
FA Median (mm ² /s)	-2.05e-03 (4.26e-04)	8.16e-06	0.002 (0.021)	0.19	-5.62e-03 (1.72e-02)	4.90e-04
MDPH (mm ² /s)	-3.87e-04 (3.32e-05)	<2e-16	-1.14e-04 (1.17e-03)	0.12	-6.05e-04 (1.27e-03)	8.11e-07
FAPH (mm ² /s)	-1.44e-06 (5.00e-06)	0.77	9.62e-05 (2.77e-04)	6.60e-08	2.00e-04 (5.77e-04)	2.25e-04
MD pkval (mm ² /s)	2.77e-06 (6.59e-07)	4.95e-05	-2.67e-06 (2.54e-05)	0.09	8.67e-06 (3.07e-05)	2.47e-03
FA pkval (mm ² /s)	-6.44e-03 (1.69e-03)	2.76e-04	-3.00e-03 (7.56e-02)	0.52	-8.00e-04 (1.66e-02)	0.60
BV (ml)	-8.88 (0.86)	<2e-16	-22.39 (19.95)	< 2.2e-16	-16.00 (132.86)	0.19
WMH	0.08 (0.01)	<2e-16	0.32 (0.46)	< 2.2e-16	2.56 (9.53)	3.85e-03
MRI incidence	Number of patients (Proportion %)		Number of patients (Proportion %)		Number of patients (Proportion %)	
Lacune count	27 (0.27)		30 (0.12)		10 (0.08)	
CMB count	36 (0.36)		32 (0.12)		24 (0.20)	

FA= fractional anisotropy, MD= mean diffusivity, pkval= peak value of the histogram distribution, PH= normalised peak height of the histogram distribution, BV= brain volume, WMH= white matter hyperintensity, CMB= cerebral microbleeds, Est. annual mean change = annual mean change estimated by the linear mixed model, Abs. mean change= absolute mean change between the 2 time points.

Table 10. Longitudinal analysis between DTI change measures and dementia. Values show the standardized regression coefficients β (SE) for the baseline predictor variables in a Cox regression or logistic regression models of dementia conversion. The odds ratio (OR) together with the confidence interval (CI) are shown. The model parameters Hosmer and Lemeshow's R^2 (R^2_L) give an estimated amount of variation in the dependent variable explained by the model. The Area under the curve (AUC) evaluates how well the model classifies dementia conversion and no-dementia conversion at all possible cutoffs respectively. All regression models control for the effects of age, gender and NART-IQ or education.

	Change marker predicting dementia conversion					
Cohort	DTI change markers	β (SE)	P-value	HR/ OR (95% CI)	Ng R^2 / R^2_L	AUC
SCANS	MDPH	-0.674 (0.243)	0.006	0.510 (0.317, 0.820)	0.119	0.750
	FAPH	0.652 (0.235)	0.006	1.919 (1.211, 3.041)	0.112	0.715
	MD Median	0.951 (0.226)	2.49e-05	2.588 (1.663, 4.027)	0.202	0.785
	FA Median	-0.388 (0.205)	0.059	0.679 (0.454, 1.014)	0.078	0.713
	MD pkval	0.388 (0.212)	0.067	1.474 (0.973, 2.234)	0.077	0.678
	FA pkval	-0.686 (0.215)	0.001	0.504 (0.331, 0.767)	0.131	0.769
	RUN DMC	MDPH	-0.409 (0.417)	0.327	0.665 (0.289, 1.482)	0.257
FAPH		0.005 (0.356)	0.990	0.995 (0.483, 1.910)	0.247	0.895
MD Median		-0.068 (0.305)	0.825	0.935 (0.498, 1.667)	0.248	0.891

	FA Median	-0.043 (0.356)	0.905	0.958 (0.496, 1.934)	0.247	0.896
	MD pkval	-0.175 (0.277)	0.528	0.839 (0.475, 1.428)	0.251	0.894
	FA pkval	0.015 (0.369)	0.967	1.015 (0.485, 2.069)	0.247	0.895
HARMONISATION	MDPH	-0.059 (0.251)	0.816	0.943 (0.570, 1.545)	0.076	0.682
	FAPH	0.375 (0.242)	0.120	1.455 (0.911, 2.389)	0.097	0.704
	MD Median	0.453 (0.237)	0.056	1.573 (0.998, 2.597)	0.109	0.738
	FA Median	-0.134 (0.245)	0.585	0.875 (0.538, 1.420)	0.078	0.693
	MD pkval	0.309 (0.245)	0.207	1.362 (0.847, 2.234)	0.090	0.704
	FA pkval	0.005 (0.267)	0.986	1.005 (0.599, 1.722)	0.075	0.685

FA= fractional anisotropy, MD= mean diffusivity, pkval= peak value of the histogram distribution, PH= normalised peak height of the histogram distribution, β = standardised regression coefficient, SE= standard error of the regression coefficient, HR= hazard ratio, OR= Odd's ratio, Ng R²= Nagelkerke's R², R²_L = Hosmer and Lemeshow's R², AUC= area under the curve, 95% CI= 95% confidence interval

Table 11. Change in multimodal imaging resulted in the best model fit and the largest AUC only in severe SVD. Values show the area under the curve (*AUC*) and Akaike information criterion (*AIC*) for the Cox regression and logistic regression models: a) clinical markers alone, b) change in DTI combined with clinical markers c) MRI markers combined with clinical markers, d) All markers included. *AUC* evaluates how well the model classifies dementia conversion and no-dementia conversion at all possible cutoffs respectively. Differences in *AUC* between the models were tested. *AIC* estimates the quality of each model relative to the other model within a cohort in terms of the trade-off between goodness of fit and the simplicity of the model.

Predictive Models	Cohorts					
	SCANS		RUN DMC		HARMONISATION	
	AUC	AIC	AUC	AIC	AUC	AIC
Clinical	0.653	158.77	0.895	81.01	0.684	110.94
Clinical-DTI	0.785	143.22	0.891	82.96	0.738	109.14
Clinical-MRI	0.853	144.66	0.916	83.13	0.727	116.87
Clinical-MRI-DTI	0.872	134.89	0.911	84.23	0.769	114.43
Model comparisons (DeLong's test for two correlated ROC curves)						
Clinical vs. Clinical-DTI	Z = -1.731, p= 0.083		Z = 0.788, p= 0.431		Z = -1.210, p= 0.226	
Clinical vs. Clinical-MRI	Z = -2.965, p= 0.003		Z = -0.723, p= 0.470		Z = -1.346, p= 0.179	
Clinical vs. Clinical-MRI-DTI	Z = -2.729, p= 0.006		Z = -0.622, p= 0.534		Z = -1.557, p= 0.119	
Clinical-DTI vs. Clinical-MRI-DTI	Z = -1.930, p= 0.054		Z = -0.874, p= 0.382		Z = -0.976, p= 0.329	

Clinical= age, sex, premorbid IQ or education; Clinical-DTI= age, sex, premorbid IQ or education and MD Median, Clinical-MRI= age, sex, premorbid IQ or education, brain volume, white matter hyperintensity volume, lacune count and cerebral microbleeds, Clinical-MRI-DTI= age, sex, premorbid IQ or education, brain volume, white matter hyperintensity volume, lacune count, cerebral microbleeds and MD Median, AUC= area under the curve, AIC= Akaike information criterion

4. Discussion

The findings across different SVD cohorts underline the importance of microstructural WM damage across different degrees of SVD as a major determinant of cognitive impairment and dementia conversion. In all prospective dementia cohorts ranging from severe SVD to amnesic MCI, baseline DTI predicted dementia conversion independently of the clinical markers. Whereas in severe SVD the primary dementia subtype was of vascular outcome, in amnesic MCI Alzheimer was the predominant diagnosis. Across all 6 cohorts, ranging from a stroke-free community-based cohort to a monogenic SVD cohort, DTI was significantly related to impaired cognitive function. The DTI measure explained most of the model's variance in the CADASIL cohort. The relative importance of conventional imaging markers varied depending on the SVD severity. While BV was an important predictor in all single-centre sporadic and monogenic SVD cohorts (SCANS, RUN DMC and CADASIL), lacune count was only a strong predictor in the severe and monogenic SVD groups (SCANS, CADASIL). None of the conventional MRI measures explained much in cognition in ASPS-Fam and HARMONISATION. The findings further underline the importance of clinical markers in explaining cognition across the different populations. Whereas in the stroke-free community cohort the predictor age had the highest estimate, in all other non-monogenic cohorts a measure of premorbid IQ or education was most important. This may be explained by the cognitive reserve hypothesis, where patients with a high cognitive reserve experiencing brain damage may be able to complete tasks better and may maintain greater degrees of neuronal damage before showing cognitive decline ²⁵⁴. In CADASIL clinical markers explained little in cognition, which is consistent with a monogenic condition setting with an earlier age of onset ²⁵⁵. A model reflecting the chapter's findings was proposed that describes the importance of microstructural WM damage, of conventional MRI markers and clinical markers (Figure 2). As disease severity increases, the importance of clinical markers decreases and the importance of DTI and conventional markers increases in terms of predicting impaired cognitive function and dementia conversion.

These results are overall in line with the hypothesis that WM damage results in disruption of WM pathways and therefore reduced brain network connectivity and impaired cognition. Network analysis based on DTI has shown that network summary measures are associated with cognitive function and predict dementia conversion in SVD ^{256–258}. Mediation analysis demonstrated that conventional markers of SVD such

as WM damage, lacunes, and CMB impair cognition through disruption of network integrity²⁵⁶. Our results support the importance of this mechanism across a wide range of patient groups, not only those with prominent SVD, but also those with MCI in whom the predominant pathology is likely to be AD. Previous studies have shown the importance of WM damage contributing to cognitive impairment in AD-like dementia^{259,260}.

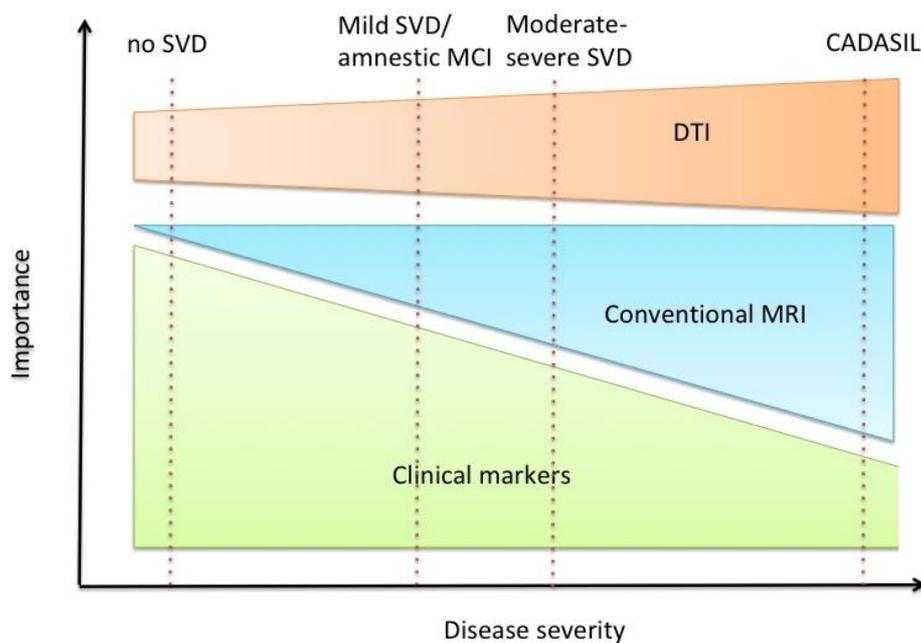


Figure 2. The importance of different SVD markers depending of the SVD severity related to clinical outcome measures. As disease severity increases in SVD, the importance of clinical markers decreases, and the importance of conventional MRI and DTI markers increases in terms of predicting impaired cognitive function and dementia conversion.

Whereas DTI at baseline was consistently associated with cognitive impairment and predicted dementia conversion in all cohorts, the clinical significance of change in DTI was inconsistent. MD Median as the primary marker of choice significantly changed in all prospective cohorts with dementia. The DTI marker's change was, however, only significantly associated with dementia in the severe SVD but not in the mild SVD or amnestic MCI cohort. To be employed as a surrogate marker serving as clinical endpoint in a phase 2 clinical trials, one important criterion is that the marker's change is associated with later relevant clinical outcome measures⁶⁶. The AUC further increased in severe SVD when adding conventional MRI measures to

the predictive model and resulted in a better model fit. This evidence altogether suggests that trials using DTI together with conventional MRI markers as surrogate markers are likely to be more successful in more severe cohorts of SVD where the vascular factor is the primary determinant for dementia conversion. On the other hand, less severe cohorts may require much larger sample sizes in a phase 2 clinical trial. In clinical trials with severe SVD conventional imaging markers may additionally be considered potentially further increasing the predictive power of dementia conversion.

The computation of the all WM histogram measure is time consuming and labor intensive with different MRI sequences being used. New methods such as the peak width skeletonized mean diffusivity (PSMD) ¹¹³ and the diffusion tensor image segmentation technique (DSEG) ²⁶¹ have the advantage of not only relying on the DTI-sequence alone but also being fully- or semiautomatic. This makes the implementation of DTI markers in large clinical trials much more feasible. These new imaging markers need to be compared to the more conventional histogram markers, described in this chapter, in terms of its utility as surrogate endpoints for a clinical trial across different SVD populations.

The major strength of the study is the validation of different imaging measures across various populations. There are also a number of limitations that need to be considered when interpreting the findings: Although these cohort studies were set up to analyze the association between imaging and clinical outcome measures, the underlying study design and image analysis differed between studies. First, the different cohorts had different length of follow-up, and number of follow-up MR scans. In RUN DMC we could only include time points 2011 and 2015 in the longitudinal DTI-change analysis due to scanner upgrades between 2006 and 2011. Second, different MRI scanners and different field strengths were employed (1.5 vs 3 Tesla). Third, different MRI analyses were used in quantifying WMH, BV and DTI median in the different studies. Fourth, the computation of cognitive scores as well as the criteria for dementia diagnosis was different across cohorts, which however may be strength as it shows the generalizability of the associations across a variety of methods.

To conclude, these findings emphasize the central role of WM microstructural damage in cognitive function and dementia conversion across all cohorts. Combining conventional markers especially lacune count and brain atrophy with DTI may

provide additional value in cohorts with more severe SVD. After finding convincing evidence of the importance of WM microstructural damage in all cohorts, it is now essential to determine which DTI marker best to employ in a phase 2 clinical trial and to test whether the advanced DTI markers such as PSMD and DSEG also should be best used in more severe SVD patients when testing new treatments.

Chapter 4

Which of the ways to analyze diffusion data results in the optimal surrogate marker for a future phase II clinical trial in SVD?

1. Introduction

The results in Chapter 3 demonstrated that across all cohorts baseline DTI was associated with cognitive function and predicted dementia conversion. Change in DTI was, however, only related to dementia conversion in the severe SVD cohort SCANS but not in the mild SVD or MCI cohort. As new ways of analyzing DTI data have been developed over the recent years, it is important to determine which DTI measure should be best used for a future phase II clinical trial. The aim of this chapter is to compare the conventional DTI marker employed as a measure in Chapter 3 to more recently developed DTI measures, which do not require conventional MRI sequences such as T1 or FLAIR in the marker's computation and which are further characterized by a semi- or fully automated computational pipeline. This makes the marker significantly more user friendly and more applicable for large clinical trial studies than the conventional DTI markers.

One of the most widely studied recent markers in SVD is the peak width skeletonized mean diffusivity (PSMD)¹¹³. It is an automated imaging marker that has been developed for patients with sporadic and monogenic SVD and is calculated based on the skeletonization of the WM using tract-based spatial statistics (TBSS)²⁶². The results of the original article showed that the marker was able to distinguish skeletonized WM structures in SVD from AD and healthy control cases and explained more variance in cognition than the conventional DTI marker whole brain MD peak height. PSMD further demonstrated sensitivity to change over 18 months in monogenic patients and required a lower sample size for a clinical trial than whole brain MD peak height, WMH volume, lacune volume and cognitive measures. Other studies showed that the PSMD is robustly associated with conventional MRI markers and outperformed these MRI markers as well as whole brain DTI measures in terms of predicting cognition²⁶³. Recent evidence further showed that age impacts PSMD differently than other DTI measures with a steady increase throughout adult life, including post adolescence and with a sharper increase from age 60 emphasizing the marker's potential usage as an indicator for the general ageing process²⁶⁴.

Changes in the marker over 3 years were associated with cognitive decline in a stroke-free and dementia-free elderly cohort ²⁶⁵. The PSMD marker has also been shown to be higher in CAA, where it was related to impaired cognitive function, as well in multiple sclerosis, where it was related to the overall WM lesion volume ^{266,267}.

Another emerging marker is diffusion tensor segmentation (DSEG) marker. This is based on anisotropic and isotropic diffusion metrics and is a semi-automatic marker ²⁶¹. The rationale behind the marker's development was to reduce the scanning time by only using a single instead of multimodal MRI metrics, to simplify the post-processing imaging steps and to be applicable as a marker in multi-center trials involving different scanner types. Results showed that baseline DSEG significantly predicted cognitive decline and later dementia conversion in severe SVD with a high discrimination ²⁶⁸. The *AUC* further increased when including both baseline and change in DSEG together with the clinical markers age, sex and premorbid IQ.

Markers of structural network efficiency have also been used in SVD cohort studies ^{256–258,269,270}. It has been shown that whole brain connectivity disruption is associated with conventional MRI measures in SVD such as WMH, lacune and CMB count ^{256,271}. Further analysis showed that the association between SVD imaging markers and impaired cognitive function is partly mediated by network disruption ^{256,271}. Further in-depth network analysis additionally demonstrated that the overall lower network connectivity in SVD may particularly be due to lower rich club connectivity, i.e. nodes with a high number of connections, which are connected to other highly connected nodes, which mediated the relationship between WMH and cognitive function ²⁶⁹. Recent evidence further showed that decline in structural network measures predicted later conversion to dementia and also mediated the link between conventional MRI markers of SVD and impaired cognitive function or dementia conversion ²⁵⁸.

To determine the most suitable surrogate imaging marker for a future phase 2 clinical trial, PSMD, DSEG and the global network efficiency measure (Geff), is compared to the MD Median, described in the previous chapter, and to a newly developed measure. The new metric is a summary measure calculated by a principal component analysis (PCA) which relies on existing DTI measures of WM microstructural damage and takes into account characteristics of diffusion from four diffusivity maps: mean diffusivity (MD), fractional anisotropy (FA), axial diffusivity (AD) and radial diffusivity (RD). Previously, it has been shown that the histogram

measures based on these diffusivity maps were related to impaired cognitive function in SVD and demonstrated significant change over time as SVD progressed over 3 years^{44,86}. Whereas MD and FA give information about the extent of WM microstructural damage and the directionality of WM diffusion, it has been suggested that additional information may be obtained by the axial and radial tensor measures with regard to the extent of axonal damage and of demyelination respectively^{272,273}. PCA was previously applied to all WM DTI histogram measures at baseline and component scores for the follow-up time points was computed employing the component coefficients and standardized DTI values based on the baseline values²⁰¹. Measures of AD and RD were excluded from this PCA and principal components were retained if eigenvalues were greater than 1. The results showed that the component scores accounted for 80.6% of the variance in the PCA. Their sample size estimation of a hypothetical clinical trial however showed that their PC scores would not require a lower minimum sample size estimate for a hypothetical clinical trial than the underlying single all WM measures of the PCA such as MDPH or MD Median.

The PCA in this chapter is computed differently by taking into account the summary measures of all 4 diffusivity maps separately both at baseline and longitudinally. By comparing the first principal component (PC1) to existing imaging measures, it will be determined which measure may be best employed as a surrogate marker for a future phase II clinical trial. The WM imaging measures will be tested based on the following criteria across the all cohorts:

- 1) the markers' baseline association with impaired cognitive function
- 2) the markers' baseline prediction for dementia conversion
- 3) the marker change over time in a multicenter trial setting
- 4) the markers' change over time related to dementia conversion
- 5) the markers' minimum sample size required for a phase II clinical trial in SVD

2. Methods

2.1 Participants

The patients included came from the six cohorts described in Chapter 3:

1. Severe symptomatic SVD (SCANS)
2. Severe symptomatic SVD (PRESERVE)

3. Moderate SVD (RUN-DMC)
4. Mild cognitive impairment (HARMONISATION)
5. Elderly stroke free population based cohort (ASPS-Fam)
6. Monogenic SVD (CADASIL)

In this chapter the RUN DMC cohort data were restricted to the two follow-up time points 2011 and 2015 due to a lack of all imaging measures at 2006. All cohorts used for analysis had approval by the ethics committee of the respective institutions. Informed consent was obtained from all participants.

2.2 DTI imaging measures

2.2.1 Peak width of skeletonized mean diffusivity (PSMD)

The PSMD is a fully automatically computed imaging marker ¹¹³ and is divided into two main modules: 1) tract-based spatial statistics ²⁶², 2) histogram analysis. The first step was to align the patient's FA data to a 1x1x1 standard space by employing the FSL software package called FNIRT and taking the standard space FMRIB 1mm FA template as a pre-defined reference image. The individual's FA data were subsequently projected onto the skeleton derived from the standard space template thresholded at an FA value of 0.2. Following that, the MD images were projected onto the skeleton by employing the FA-derived projection parameter. The skeletonized MD data were additionally masked with the template skeleton that is thresholded at an FA value of 0.3. This last step ensured that the skeleton was not contaminated by CSF partial volume effects. Areas directly next to the ventricles were removed from the analysis by a custom-made mask. The voxel-based MD values on the skeleton, which represent the main WM tracts in the brain, were then plotted into a histogram showing a normalized distribution of the skeletonized MD values. To obtain the PSMD value, the difference between the 5th and 95th percentiles of the histogram distribution was computed which reflects the peak width of the MD histogram.

2.2.2 Diffusion tensor image segmentation

The DSEG is a semiautomatic DTI marker and assigns an unique diffusion profile to each voxel that is based on the magnitudes of isotropic and anisotropic diffusion metrics ²⁶¹. The processing pipeline is composed of 4 main modules:

- 1) Pre-processing pipeline
- 2) Diffusion tensor decomposition

- 3) DTI segmentation algorithm
- 4) Creating the summary measure

The preprocessing pipeline comprised of the following parts: First, b0 images were skull stripped with the FSL brain extraction software tool BET ²⁷⁴. Second, the cerebellum was excluded employing an automated pipeline ²⁷⁵.

Employing FSL DTIFIT diffusion tensors were created for each voxel. The diffusion tensors were then decomposed into isotropic (p) and anisotropic (q) segments using the following formulas ²⁷⁶:

$$p = \sqrt{3MD}$$

$$q = \sqrt{(\lambda_1 - MD)^2 + (\lambda_2 - MD)^2 + (\lambda_3 - MD)^2}$$

where

$$MD = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3}$$

where λ_n is the eigenvalue of the n th eigenvector.

The isotropic measure refers to a measure of MD , whereas the anisotropic measure represents the deviation of the diffusivities from isotropy.

The DTI segmentation (DSEG) algorithm was then run. Using a k-median cluster analysis the diffusion data of each voxel was then categorized into 16 segments that reflect the distinct magnitudes of anisotropy (q) and isotropy (p) microstructural diffusion properties. The analysis allows comparison of the segments and to find brain regions with similar diffusion profiles.

The number of cerebrum voxels per segment could then be quantified as percentage in order to compose the diffusion characteristics of the entire brain. The DSEG spectra were created for each patient by determining the percentage of each segment to the total cerebral volume. The respective spectral diffusion profile consisting of 16-dimensional vectors can be compared between patients by a summary metric in the form of a scalar product, θ , which reflects differences in whole brain diffusion in patients with respect to a reference brain. The scalar product was computed as follows:

$$\theta(\text{radians}) = \cos^{-1} \left(\frac{A * B}{\|A\| \|B\|} \right)$$

where A and B were the spectral vectors of the reference brain and of the dataset respectively.

2.2.3 Global Network Efficiency Measure (Geff)

To model the brain based on network analysis, graph theory has become a vital tool²⁷⁷. A graph, G , represents a fairly elementary model by which the neuronal system is reduced to a collection of nodes and edges.

Edges refer to interconnections between any pair of nodes. These connections are often captured in a two-dimensional matrix, called a connectivity matrix, where each column and each row correspond to a different node. The information regarding the connections between nodes is found at the intersection of i th row and j th column. The value of each component in the matrix represents the 'edge' or connectivity between the two nodes.

In the current brain network analysis nodes referred to different brain regions defined by the Automated Anatomical Labeling (AAL) atlas consisting of a total of 116 GM ROIs. 90 regions (80 cortical and 10 subcortical), defined from the Desikan-Killiany parcellation of the cerebral cortex and subcortical nuclei²⁷⁸, were used as nodes in this study, excluding those in the cerebellum in the AAL atlas. Diffusion-weighted images were pre-processed and diffusion tensor was fit as described in Chapter 2²⁷⁹.

The standard space template FMRIB58_FA_1mm was registered to the patients FA maps in ANTs (stnava.github.io/ANTs/). The AAL atlas was then transformed to the patient's FA image using the nearest-neighbor interpolation method and employing the transformation matrix created by the registration.

Employing whole-brain deterministic tractography WM connectivity was modeled by relying on the directional information within each diffusion tensor²⁸⁰. A continuous tensor field was built with the eigenvector corresponding to the largest eigenvalue in each voxel. Streamlines (max 4 per voxel) of a length between 20- 250 nm were launched and spread in orthograde and retrograde directions by trilinear interpolation

of the tensor field with a vector step size of 0.5 mm. Streamlines ended in regions where FA < 0.15 or the angle turn between consecutive vectors was $\theta > 45$.

Two nodes, i and j , were linked by an edge under the condition that the terminals of a streamline, created by the deterministic tractography, was located between both regions. Edges, w_{ij} , were weighted based on the length, l , of the streamlines in mm:

$$w_{ij} = \frac{1}{2} \sum_{m=0}^M \frac{1}{l_m}$$

where M is finite group of streamlines linking the 2 nodes. To correct for the number of seeds per millimeter, the connectivity distance was computed as follows:

$$d_{ij} = \frac{1}{w_{ij}}$$

With the aim of avoiding false-positive rates, the edges were thresholded at $w_{ij} \geq 1$.

The network efficiency between brain regions was estimated employing the Brain Connectivity Toolbox (www.brain-connectivity-toolbox.net) and the whole-brain summary measure called weighted global network efficiency (Geff) was computed. Geff is a measure of network integration. It shows how well connected all nodes of the brain network are in comparison to an idealized network where each node is connected with every other one²⁸¹. It is defined as:

$$Geff = \frac{1}{L'} = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{l_{ij}}$$

where L' is the average shortest path length between all possible pairs using the harmonic mean²⁸² and where l_{ij} is the shortest path length from node j to node i .

2.2.4 MD Median

The image processing pipeline has been described in Chapter 2 and is summarized in Figure 2 in Chapter 2.

2.2.5 Principal component measure

The image processing pipeline has been described in Chapter 2 and is summarized in Figure 2 in Chapter 2. Five summary measures characterizing the histogram were computed on 4 WM diffusion maps: (i) mean diffusivity (MD), (ii) fractional anisotropy (FA), (iii) axial diffusion (AD) and (iii) radial diffusion (RD):

- Median
- Normalised peak height (PH)
- Peak location (pkval)
- Skew
- Kurtosis

A PCA was computed based on all 20 baseline histogram summary measures. A separate PCA was furthermore calculated using all change scores over time in the DTI measures. PCA is an unsupervised learning method and geometrically projects multi-dimensional data structures into constructs called principal components (*PCs*) which effectively reduces the dimensionality of the data while retaining maximal amount of information of the variables. The aim is to get a good description of the high-dimensional data while employing a limited count of *PCs*.

The data was first centered on the means of each variable's dimension. This ensures that the multidimensional data set was centered on the *PCs* origins. The first *PC* (*PC1*) captures the direction where most variation in the multidimensional data is found. Subsequent *PCs* are geometrically orthogonal and capture increasingly less variation in the data. *PCs* was computed by minimizing the sum of all distances between centered data points and their projection on the *PC* and by inversely maximizing the total sum of squared distances from the projected data points to the origin. The squared distances from the projected data points to the origin is the eigenvalue for the *PC*. The eigenvalue was then converted into the variation around the origin by dividing the eigenvalue by the sample size minus 1.

The association of the original variables and *PC* can be expressed in component loadings. Loading scores (*a*) are coefficients of the linear combination of the original variables (X_p) from which the *PC* is created:

$$PC_1 = a_{11}X_1 + a_{12}X_2 + \dots + a_{1p}X_p$$

where:

$$a_{11}^2 + a_{12}^2 + \dots + a_{1p}^2 = 1$$

It describes how much each variable contributes to the *PC* and whether the association is negatively or positively correlated. Mathematically, the component loadings are equal to the variables' coordinates divided by the square root of the *PC's* eigenvalue. The patient's *PC* scores (*E*) can further be calculated by taking the

patient's standardized score (S) for each variable and by multiplying them by the PC loading (a). The products are then summed up to create the PC component score for each patient:

$$E_{PC1} = S_{11} a_{11} + S_{12} a_{12} + \dots S_{1p} a_{1p}$$

2.3 Cognition and Dementia

The cognitive index measures Global cognition (Global) and the Trail-making test B (TMT-B), described in Chapter 3, were included in the cross-sectional analysis. For the longitudinal analysis, the binary clinical outcome of dementia conversion as described in Chapter 3 was used.

2.4 Statistical analysis

2.4.1 PCA on baseline conventional DTI

Heatmaps were created reflecting the magnitude of Pearson correlation between the conventional 20 DTI markers in each cohort. PCA was subsequently performed using all 20 DTI conventional measures. The scores of the first principal component (PC1) were used as a predictive DTI summary measure for the subsequent cross-sectional analysis.

2.4.2 Cross-sectional analysis

The relationship between the DTI measures and cognition was tested using a linear regression while adjusting for age, gender and premorbid IQ or years of education completed. In the multicenter study one center recruited only a single patient and therefore this participant was excluded. Study site was incorporated as an additional confounder into the linear model. In the CADASIL cohort one patient outlier observation was removed from the analysis for violating the model's assumption of the linear regression analysis. The linear model's underlying assumptions such as the normality of the residuals, the homoscedasticity and linearity of the relationship were met.

2.4.3 Longitudinal analysis

2.4.3.1 Baseline imaging dementia analysis

The association between the baseline DTI value and dementia conversion was tested using a Cox regression model in SCANS, RUN DMC and HARMONISATION. The clinical markers age, gender and premorbid IQ or years of education were added as covariates. The proportional hazard assumption was held in all time-to event predictive models tested by the Global Schoenfeld Test. The hazard ratio was computed to estimate the effect of the relationship. Receiver operating characteristic curve (ROC) were computed to show the diagnostic discriminatory ability of a binary classifier system with varying thresholds (pRoc) ²⁵³.

2.4.3.2 Change in single imaging markers

To assess the DTI's sensitivity to change over the follow-up period, a linear mixed model was fitted for MD Median, PSMD, DSEG and Geff in SCANS. Fixed effect variation was explained by time, and random effect variation allowed for remaining inter-individual differences. The intercept and slope of each patient's linear trajectory were allowed to vary with both fixed and random effects. The average fixed effects slopes of time are interpreted as the average annualized change rate in a given imaging measure per additional year of follow-up. The statistical significance of change in DTI was determined with a paired t-test for RUN DMC, HARMONISATION, CADASIL and PRESERVE. In ASPSP-Fam the patient's longitudinal sample size with DTI was too low to reliably estimate any meaningful changes ($N= 69$).

2.4.3.3 PCA on change in conventional DTI in PRESERVE

PCA was performed on the 20 conventional DTI change scores in PRESERVE. Differences in PC1 and all other DTI change measures between scanner sites while accounting the marker's baseline measure were tested using an ANCOVA model with permutation.

2.4.3.4 PCA on change in conventional DTI

To create a longitudinal DTI compound measure for the subsequent dementia analysis, a PCA was applied to the changes of 19 DTI markers used at baseline in SCANS and of all 20 DTI markers in RUN DMC and HARMONISATION. In SCANS no individual trajectories of AD Median could be estimated due to limited variability in the data. Again, the scores of PC1 were used as a predictive DTI measure for further analysis

2.4.3.5 Change in DTI and dementia conversion

Running a Cox regression in SCANS or logistic regression models in RUN DMC and HARMONISATION, the predictive relationship between the change in DTI and dementia conversion was tested. ROC curves were modeled to estimate the diagnostic discriminatory ability of each model ²⁵³.

2.4.3.6 Sample size estimation for hypothetical clinical trial

In the SCANS data set sample size estimation for a hypothetical clinical trial with the imaging marker was performed using the longpower R package ²⁸³ and by varying the treatment effect sizes such as 10%, 20%, 30%, with a statistical power of 0.80 and two-tailed type I error rate of 0.05. Sample size was estimated in RUN DMC, HARMONISATION and CADASIL for the imaging marker using the following formula ²⁸⁴ :

$$n = \frac{8(CV)^2}{(PC)^2} [1 + (1 - PC)^2]$$

where PC is the proportion change in means defined as

$$PC = \frac{\mu_1 - \mu_2}{\mu_1}$$

and CV is the coefficient of variation which is

$$CV = \frac{\sigma_1}{\mu_1} = \frac{\sigma_2}{\mu_2}.$$

The sample size for the PCA scores in all cohorts was computed using the pwr package ²⁸⁵. The effect size was computed by

$$d = \frac{SD * k}{SD}$$

where SD is the standard deviation and k the respective treatment effects (10%, 20%, 30%).

3. Results

3.1 PCA baseline

There was a strong correlation between the 20 baseline DTI markers in all SVD cohorts and in the MCI cohort. In the stroke-free ASPS-Fam cohort the associations between DTI markers was significantly lower. This finding is also reflected in the PCA where the percentage of explained variance of PC1 was highest in the CADASIL (84.4%) group and lowest in ASPS-Fam (40%). In the sporadic SVD cohorts the PC1's explained variance ranged between 60.3- 71.5%. The PC1 explained 71.2% of the variance in the MCI cohort.

3.2 Cross-sectional results

There was a significant association between all imaging markers and impaired cognitive function in all SVD and in the MCI cohorts (Table 2, Table 3). The adjusted explained variance (*Adj. R²*) in each single-center sporadic SVD cohorts was similar across DTI markers (Table 2). The standardized regression coefficients were overall higher in SCANS than in RUN DMC. In the multicenter study PRESERVE the *Adj. R²* was lower in DSEG and Geff than in PC1, PSMD or MD Median. In CADASIL the PSMD model explained most of the model's variance. The overall model fit measured by the *AIC* was best for PSMD in HARMONISATION, PRESERVE and CADASIL. On the other hand, the Geff model had the best model fit for SCANS, RUN DMC and ASPS-Fam. Not one marker was consistently more strongly associated with cognitive impairment across cohorts.

Table 1 Overview over the cohorts. Clinical markers, imaging markers and sample sizes both at baseline and longitudinal are shown.

	Cohort					
	SCANS (n= 121)	RUN DMC (n= 332)	HARMONISATION (n= 127)	PRESERVE (n= 111)	ASPS-Fam (n= 382)	CADASIL (n= 58)
<i>Demographics</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>
Age (SD)	70.01 (9.75)	68.91 (8.28)	72.23 (8.47)	68.07 (9.11)	65.43 (10.67)	47.90 (9.77)
Sex, male (%)	78 (0.65)	194 (0.58)	57 (0.45)	43 (0.39)	139 (0.40)	26 (0.45)
Included in cross-sectional analysis	yes	yes	yes	yes	yes	yes
Cohort sample size with complete DTI in cross-sectional analysis	113	332	127	101	256	54
Included in longitudinal analysis	yes	yes	yes	yes	no	yes
Sample size in longitudinal analysis with complete repeated DTI	97	268	127	81	-	53
<i>Baseline complete DTI parameter</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>
MD Median (mm ² /s)	8.01e-04 (4.09e-05)	8.03e-04 (3.86e-05)	8.82e-04 (6.08e-05)	7.87e-04 (4.28e-05)	7.69e-04 (3.04e-05)	8.89e-04 (1.30e-04)
PSMD (mm ² /s)	3.80e-04 (1.14e-04)	3.20e-04 (9.35e-05)	3.60e-04 (7.71e-05)	3.93e-04 (9.77e-05)	2.97e-04 (5.31e-05)	5.63e-04 (1.88e-04)
DSEG (mm ² /s)	22.04 (9.62)	20.32 (8.04)	32.31 (8.24)	47.65 (4.01)	49.95 (8.13)	22.39 (12.36)
Geff (mm ² /s)	2.26e-03 (1.01)	10.02 (2.52)	0.41 (0.22)	0.17 (0.10)	4.36 (1.65)	2.16 (1.16)

Figure 1. Heatmaps and percentage of variance explained by PC1. The heatmaps indicate the strength (intensity) of the positive (red) and negative (blue) Pearson correlation between DTI measures. The percentage of explained variance was between 69- 71% in all 3 cohorts.

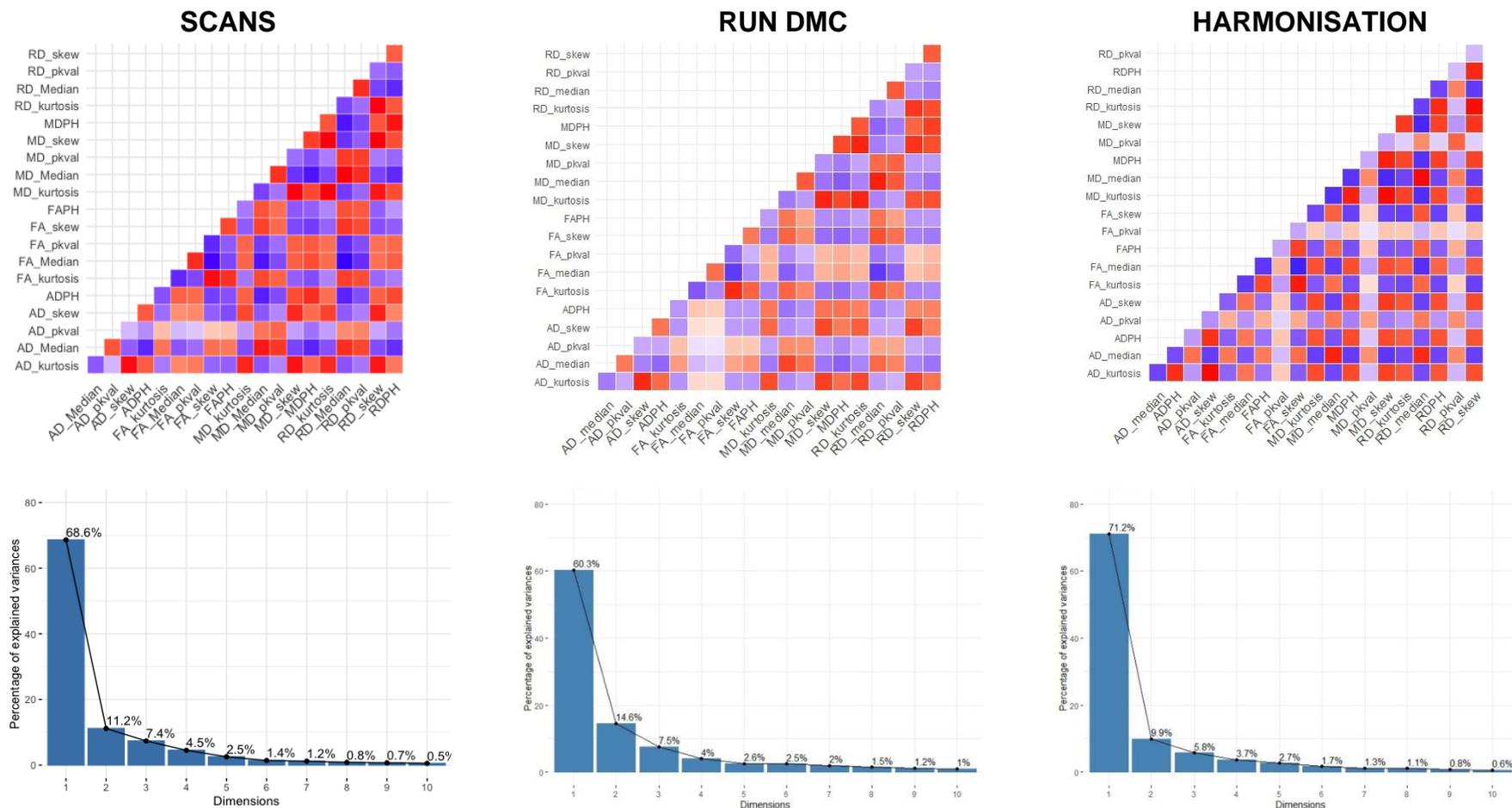


Figure 2. Heatmaps and percentage of variance explained by PC1. The heatmaps indicate the strength (intensity) of the positive (red) and negative (blue) Pearson correlation between DTI measures. The percentage of explained variance was between 40- 84.4% in all 3 cohorts.

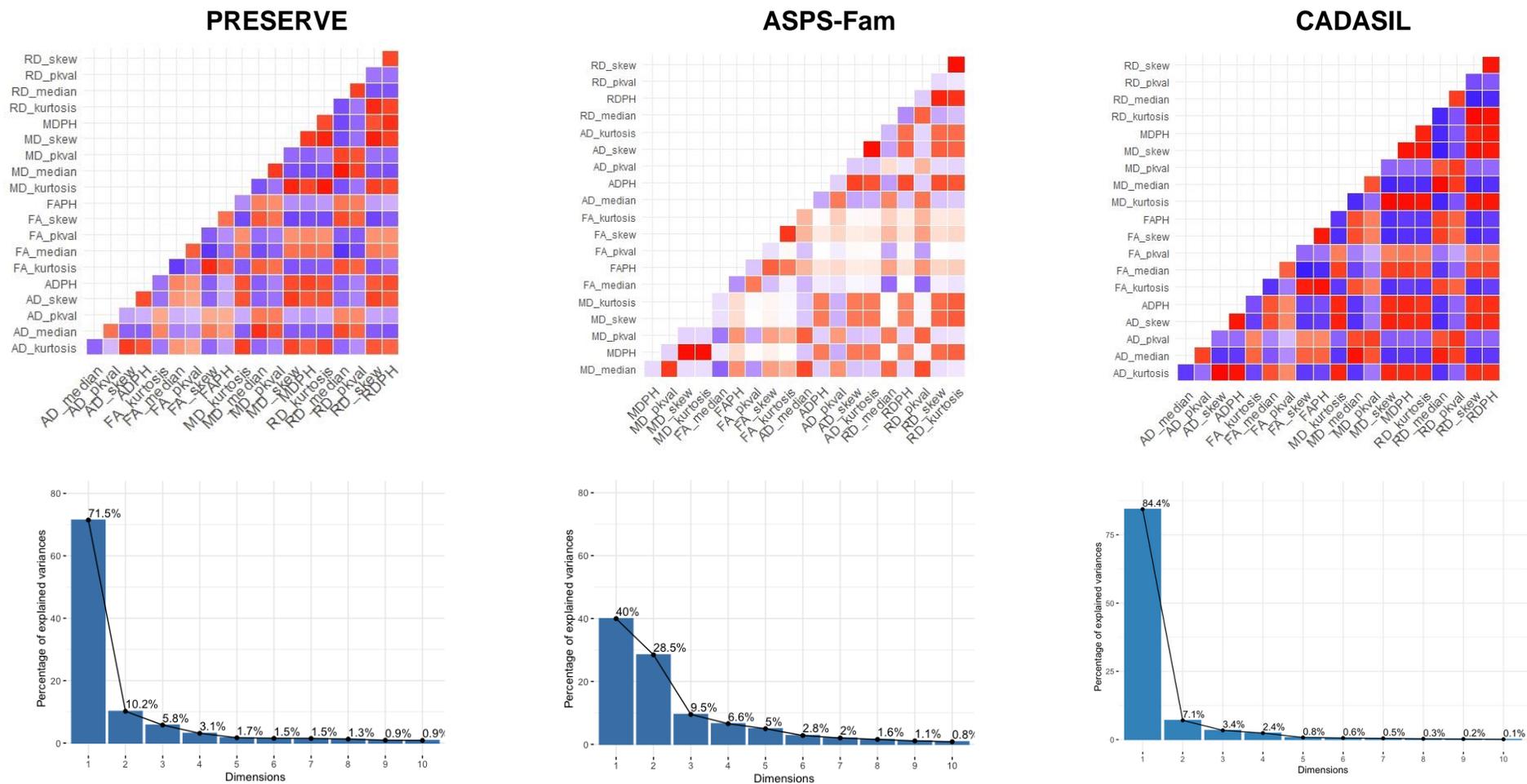


Table 2. Cross-sectional association between imaging measures and cognition. The cross-sectional associations between DTI measure and impaired cognitive function for SCANS, RUN DMC and HARMONISATION. All markers were related to cognition in all 3 cohorts. Adjusted R² was overall highest in SCANS and smallest in HARMONISATION. AIC was lowest for Geff in SCANS and in RUN DMC and for PSMD in HARMONISATION.

	Global Cognition											
	SCANS				RUN DMC				HARMONISATION			
Baseline Markers	β (95% CI)	P-Value	Adjusted R ²	AIC	β (95% CI)	P-Value	Adjusted R ²	AIC	β (95% CI)	P-Value	Adjusted R ²	AIC
MD Median	-0.25 (-0.38, -0.12)	<0.001	0.518	245.19	-0.18 (-0.28, -0.08)	< 0.001	0.464	634.85	-0.35 (-0.50, -0.19)	<0.001	0.456	287.73
PC1	-0.30 (-0.43, -0.18)	<0.001	0.546	238.33	-0.19 (-0.30, -0.09)	< 0.001	0.464	634.56	-0.36 (-0.51, -0.20)	<0.001	0.460	286.88
PSMD	-0.30 (-0.42, -0.17)	<0.001	0.545	238.67	-0.25 (-0.37, -0.14)	< 0.001	0.474	629.37	-0.37 (-0.51, -0.23)	<0.001	0.483	281.37
DSEG	-0.38 (-0.53, -0.23)	<0.001	0.556	235.75	-0.23 (-0.35, -0.10)	< 0.001	0.461	636.14	-0.44 (-0.62, -0.25)	<0.001	0.466	285.48
Geff	0.35 (0.22, -0.48)	<0.001	0.569	232.40	0.28 (0.18, 0.38)	< 0.001	0.492	619.30	0.21 (0.07, 0.35)	<0.001	0.409	298.11

MD Median= mean diffusivity median of the all WM histogram, PC1= scores of the first principal component, PSMD= peak width of skeletonized mean diffusivity, DSEG= diffusion tensor image segmentation, Geff= global efficiency network measure, β = standardized regression coefficient, 95% CI= 95% confidence interval, AIC= Akaike information criterion, P-Value= statistical value of significance with $p < 0.05$.

Table 3. Cross-sectional association between imaging measures and cognition. The cross-sectional associations between imaging and impaired cognitive function for PRESERVE, ASPS-Fam and CADASIL. All markers were related to cognition in PRESERVE and CADASIL. In ASPS-Fam MD Median, DSEG and Geff were associated with impaired cognitive function. AIC was lowest for PSMD in PRESERVE and CADASIL and lowest for Geff in ASPS-Fam.

Baseline Markers	Global Cognition/ TMT-B											
	PRESERVE				ASPS-Fam				CADASIL			
	β (95% CI)	P-Value	Adjusted R ²	AIC	β (95% CI)	P-Value	Adjusted R ²	AIC	β (95% CI)	P-Value	Adjusted R ²	AIC
MD Median	-0.40 (-0.57, -0.23)	<0.001	0.389	242.69	0.14 (-0.24, -0.04)	0.01	0.516	488.48	-0.50 (-0.82, -0.18)	<0.001	0.227	133.15
PC1	-0.37 (-0.55, -0.20)	<0.001	0.375	245.01	-0.04 (-0.14, 0.06)	0.38	0.503	494.77	0.50 (0.15, 0.85)	0.01	0.206	134.50
PSMD	-0.41 (-0.58, -0.25)	<0.001	0.409	239.53	-0.03 (-0.13, 0.08)	0.64	0.501	495.32	-0.70 (-1.01, -0.39)	<0.001	0.357	124.11
DSEG	-0.19 (-0.37, -0.01)	0.04	0.282	258.78	-0.14 (-0.23, -0.05)	<0.001	0.521	486.35	-0.50 (-0.80, -0.20)	<0.001	0.251	131.65
Geff	-0.19 (-0.37, -0.01)	0.04	0.281	258.84	0.18 (0.09, 0.28)	<0.001	0.530	482.02	0.45 (0.09, 0.82)	0.02	0.243	136.43

MD Median= mean diffusivity median of the all WM histogram, PC1= scores of the first principal component, PSMD= peak width of skeletonized mean diffusivity, DSEG= diffusion tensor image segmentation, Geff= global efficiency network measure, β = standardized regression coefficient, 95% CI= 95% confidence interval, AIC= Akaike information criterion, P-Value= statistical value of significance with $p < 0.05$

3.3 Longitudinal results

3.3.1 Baseline DTI predicts dementia conversion

In SCANS complete baseline DTI measures were available in 113 patients of whom 18 converted to dementia. In RUN DMC 284 patients had complete baseline imaging data with 13 patients converting to dementia over a 4 years period. In HARMONISATION 127 patients had complete DTI data at baseline of with 23 patients converting to dementia. Baseline DTI predicted dementia conversion independently of the clinical markers in SCANS, RUN DMC and HARMONISATION (Table 4). Overall, the AUC of the predictive models were higher in the sporadic SVD cohorts SCANS and RUN DMC than in HARMONISATION.

3.3.2 Change in DTI

There was a significant change in the MD Median and DSEG measures in all cohorts over time (Table 5). Geff also significant changed in SCANS, RUN DMC, HARMONISATION but not in PRESERVE or CADASIL. There was a significant change in PSMD in all single-center cohort studies but not in PRESERVE.

3.3.3 Change in DTI measures across scanner sites in PRESERVE

PC1 explained 31.9% of the variance when running a PCA on the 20 conventional DTI measures in PRESERVE. There was no difference between scanner sites for any change in DTI measures while accounting for its respective baseline value (PC1: $p= 0.18$; PSMD: $p= 0.96$; Geff: $p= 0.11$, DSEG: $p= 0.51$; MD Median: $p= 1$). Overall, PC1 was less affected by outlier values compared to the other measures (Figure 3).

3.3.4 PCA on change in conventional DTI

Changes in all DTI measures in SCANS, RUN DMC and HARMONISATION are shown in Tables 7-9. Overall, most DTI markers significant changed over time. PC1 explained significantly less variation of the data in the sporadic SVD than in the MCI cohort. PC1 accounted for 36.3% of the variance in SCANS, 32.6% in RUN DMC and 57% in HARMONISATION (Figure 4).

3.3.5 Change in DTI was associated with dementia conversion

For the dementia analysis in SCANS complete imaging measures with at least 2 time points were available for 97 patients with 17 patients converting to dementia. In RUN DMC 268 patients were available of whom 12 of them converted to dementia. In HARMONISATION 126 were included of who 23 converted to dementia. Before conducting a Cox regression in SCANS, observations post dementia diagnosis were removed and the linear mixed model was rerun. As before, there were significant changes in all DTI measures (Table 10). Change over time in all imaging measures was associated with dementia conversion in SCANS, but not in RUN DMC. In HARMONISATION change in MD Median was significantly associated with dementia conversion. In SCANS the association and model fit was strongest in the DSEG model and also resulted in the highest *AUC* with 0.924 (Table 11). Due to the small dementia incidence in RUN DMC relative to the overall cohort size, a Firth's Bias-reduced logistic regression was additionally run on the RUN DMC data. In line with previous evidence, no significant association between any changes in DTI and dementia conversion was found (Table 12).

Table 4. Baseline imaging markers predicting dementia conversion. Except for Geff in HARMONISATION, all imaging markers predicted dementia conversion independently of the clinical markers age, sex and premorbid IQ or education. The model fit was best for DSEG in SCANS, PC1 in RUN DMC and PSMD in HARMONISATION.

Baseline Markers	Dementia conversion											
	SCANS				RUN DMC				HARMONISATION			
	HR (95% CI)	P-Value	AIC	AUC	HR (95% CI)	P-Value	AIC	AUC	HR (95% CI)	P-Value	AIC	AUC
MD Median	2.19 (1.51, 3.16)	<0.001	138.37	0.832	2.56 (1.77, 3.70)	< 0.001	102.88	0.881	1.78 (1.08, 2.93)	0.02	173.51	0.761
PC1	2.28 (1.51, 3.44)	<0.001	139.60	0.825	3.06 (2.00, 4.69)	< 0.001	101.87	0.900	1.74 (1.03, 2.92)	0.04	174.17	0.765
PSMD	1.74 (1.29, 2.34)	<0.001	143.70	0.804	2.59 (1.39, 4.81)	0.003	113.04	0.875	1.73 (1.13, 2.65)	0.01	172.72	0.765
DSEG	3.52 (2.09, 5.92)	<0.001	128.21	0.908	4.69 (1.78, 12.36)	0.002	111.05	0.892	1.94 (1.06, 3.53)	0.03	173.67	0.753
Geff	0.37 (0.23, 0.61)	<0.001	139.39	0.842	0.27 (0.13, 0.59)	< 0.001	109.65	0.879	0.79 (0.49, 1.30)	0.36	177.58	0.702

MD Median= mean diffusivity median of the all WM histogram, PC1= scores of the first principal component, PSMD= peak width of skeletonized mean diffusivity, DSEG= diffusion tensor image segmentation, Geff= global efficiency network measure, AIC= Akaike information criterion, HR= hazard ratio, AUC= area under the curve, 95% CI= 95% confidence interval

Table 5. Change in DTI measures across cohort studies. There was a significant change in all markers in all cohorts except for PRESERVE and CADASIL. In PRESERVE there was a significant change DSEG and MD Median but not for PSMD and Geff. In CADASIL there was only a marginally significant change for Geff.

Markers	Change in DTI over time											
	PSMD			DSEG			Geff			MD Median		
	Baseline	Change	P-Value	Baseline	Change	P-value	Baseline	Change	P-Value	Baseline	Change	P-Value
SCANS ^a	3.78e-04 (3.58e-04, 3.98e-04)	1.40e-05 (1.03e-05, 1.77e-05)	<0.001	21.69 (19.90, 23.49)	1.61 (1.42, 1.80)	<0.001	8.12 (7.68, 8.57)	-0.18 (-0.23, -0.13)	<0.001	7.98e-04 (7.90e-04, 8.06e-04)	5.63e-06 (4.48e-06, 6.78e-06)	<0.001
RUN DMC ^b	3.05e-04 (8.33e-05)	2.70e-05 (2.69e-05)	<0.001	19.10 (7.56)	2.46 (2.27)	<0.001	10.37 (2.45)	-0.13 (0.94)	0.02	7.98e-04 (3.44e-05)	3.434e-06 (1.25e-05)	<0.001
HARMONISATION ^b	3.60e-04 (7.71e-05)	2.76e-05 (4.04e-05)	<0.001	32.31 (8.24)	2.70 (3.33)	<0.001	0.41 (0.22)	-0.05 (0.12)	<0.001	8.82e-04 (6.081e-05)	2.237e-05 (3.62e-05)	<0.001
PRESERVE ^b	3.94e-04 (1.05e-04)	-8.77e-06 (5.59e-05)	0.16	47.83 (4.12)	8.54 (7.91)	<0.001	0.17 (0.11)	0.003 (0.137)	0.87	7.88e-04 (4.61e-05)	8.31e-06 (1.70e-05)	<0.001
CADASIL ^b	5.53e-04 (1.74e-04)	6.44e-05 (6.05e-05)	<0.001	22.00 (12.14)	9.09 (6.12)	<0.001	2.19 (2.04)	-0.15 (0.53)	0.05	8.82e-04 (1.21e-04)	6.73e-05 (4.78e-05)	<0.001

^a linear mixed model in SCANS with the output: baseline intercept (95% confidence interval), estimated annual mean change (95% confidence interval) and p-value for each single imaging measure

^b paired t-test in RUN DMC, HARMONISATION, PRESERVE and CADASIL with the output: baseline mean (standard deviation), absolute mean change (standard deviation) and p-value for each single imaging measure

MD Median= mean diffusivity median of the all WM histogram, PC1= scores of the first principal component, PSMD= peak width of skeletonized mean diffusivity, DSEG= diffusion tensor image segmentation, Geff= global efficiency network measure

Table 6: Change in conventional DTI measures in PRESERVE. Except for MD pkval, RD pkval and RD skew the markers significantly changed over the 2 years in the multicenter trial.

DTI all WM marker	Mean baseline value (SD)	Mean 5 year change (SD)	P-Value
MD pkval	7.64e-04 (4.12e-05)	3.70e-06 (2.31e-05)	0.15
MD PH	1.33e-02 (2.45e-03)	-8.29e-04 (1.22e-03)	3.39e-08
MD Median	7.88e-04 (4.61e-05)	8.31e-06 (1.70e-05)	3.35e-05
MD kurtosis	10.00 (4.83)	-1.41 (1.76)	3.04e-10
MD skew	2.37 (0.58)	-0.23 (0.28)	3.67e-10
FA pkval	0.31 (4.84e-02)	-3.49e-03 (3.97e-02)	4.31e-01
FA PH	3.22e-03 (2.38e-04)	-3.48e-05 (1.77e-04)	8.03e-02
FA Median	3.34e-01 (2.81e-02)	-3.46e-03 (1.29e-02)	1.83e-02
FA kurtosis	0.52 (0.34)	4.83e-02 (0.18)	1.51e-02
FA skew	0.67 (0.15)	2.27e-02 (6.93e-02)	4.22e-03
AD pkval	1.03e-03 (5.51e-05)	1.40e-05 (5.14e-05)	0.02
AD PH	7.50e-03 (8.46e-04)	-2.78e-04 (5.00e-04)	3.7e-06
AD Median	1.10e-03 (5.17e-05)	1.11e-05 (2.31e-05)	4.8e-05
AD kurtosis	3.73 (1.43)	-1.66 (0.80)	< 2e-16
AD skew	1.38 (0.26)	-0.36 (0.19)	< 2e-16
RD pkval	6.12e-04 (4.62e-05)	4.70e-06 (2.80e-05)	0.14
RD PH	1.18e-02 (1.60e-03)	-4.94e-04 (8.03e-04)	4.52e-07
RD Median	6.33e-04 (4.70e-05)	7.90e-06 (1.69e-05)	7.74e-05
RD kurtosis	8.50 (3.56)	-0.28 (1.41)	8.26e-02
RD skew	2.08 (0.43)	-3.87e-02 (0.23)	0.14

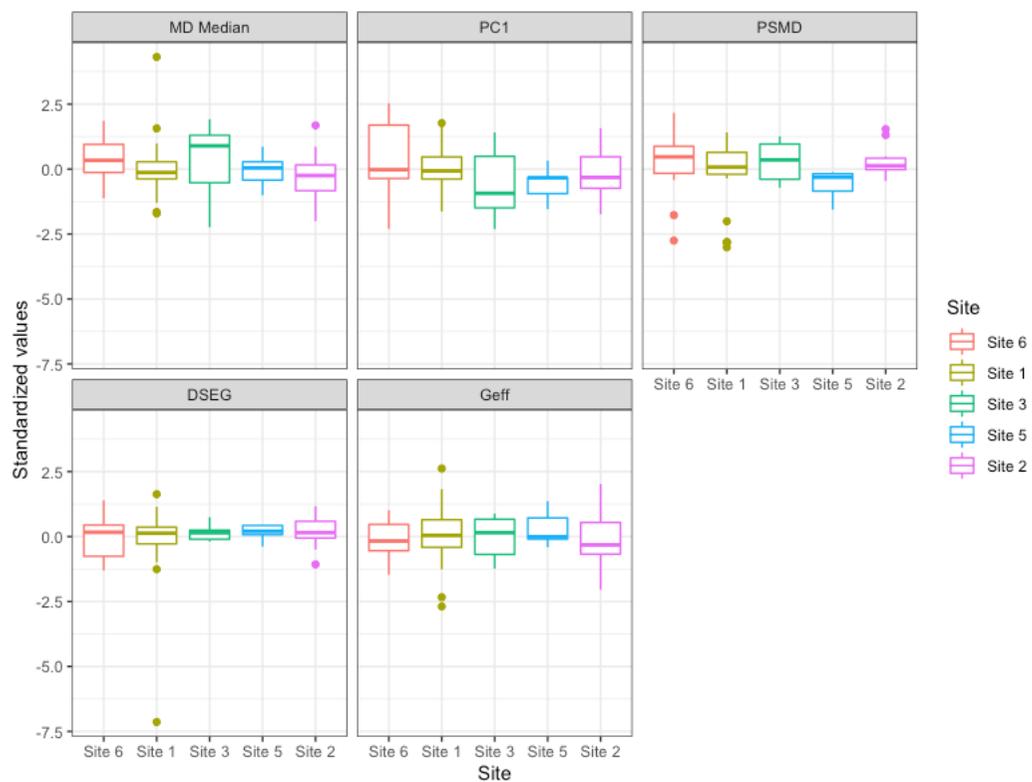


Figure 3. Change in DTI & PC1 across scanner types in PRESERVE. There was no significant difference in DTI between sites for any DTI measure. There were less outlier values in PC1 compared to the other measures.

Table 7 Change in conventional DTI measures over time in SCANS. Except for MD skew, FAPH and RD kurtosis the markers significantly changed over the 3 years

DTI all WM marker	Estimated mean baseline value (SE)	Estimated mean annual change (SE)	Wald-test	P-Value
MD pkval	7.70e-04 (3.18e-06)	2.94e-06 (6.50e-07)	4.52	1.39e-05
MD PH	1.52e-02 (2.84e-04)	-3.84e-04 (3.28e-05)	-11.70	< 0.001
MD Median	7.98e-04 (4.07e-06)	5.43e-06 (6.06e-07)	8.96	< 0.001
MD kurtosis	15.22 (0.66)	-0.32 (0.13)	-2.36	0.02
MD skew	2.75 (0.06)	0.02 (0.01)	1.44	0.15
FA pkval	0.27 (0.01)	-6.18e-03 (1.80e-03)	-3.43	< 0.001
FA PH	3.05e-03 (2.17e-05)	1.32e-06 (5.42e-06)	0.24	0.807
FA Median	0.29 (2.86e-03)	-2.16e-03 (4.41e-04)	-4.90	< 0.001
FA kurtosis	0.52 (0.03)	0.02 (6.62e-03)	2.23	0.03
FA skew	0.68 (0.01)	0.01 (2.39e-03)	4.54	< 0.001
AD pkval	1.01e-03 (3.62e-06)	3.71e-06 (1.16e-06)	3.19	0.002
AD PH	8.44e-03 (8.90e-05)	-1.28e-04 (1.20e-05)	-10.63	< 0.001
AD kurtosis	5.20 (0.15)	0.25 (0.04)	6.56	< 0.001
AD skew	1.51 (2.43e-02)	6.32e-02 (7.05e-03)	8.96	< 0.001
RD pkval	6.42e-04 (3.84e-06)	4.67e-06 (1.39e-06)	3.37	0.001
RD PH	1.26e-02 (1.81e-04)	-2.77e-04 (2.44e-05)	-11.34	< 0.001
RD Median	6.65e-04 (4.57e-06)	6.30e-06 (1.31e-06)	4.80	< 0.001
RD kurtosis	12.06 (0.46)	-0.11 (0.11)	-1.10	0.27
RD skew	2.37 (0.04)	0.04 (0.01)	3.70	< 0.001

Table 8 Change in conventional DTI measures over time in RUN DMC. 12

markers showed a significant change over the 4 years.

DTI all WM histogram marker	Mean baseline value (SD)	Mean 5 year change (SD)	Paired t-test	P-Value
MD pkval	7.81e-04 (3.43e-05)	-2.12e-06 (2.55e-05)	1.37	0.17
MD PH	1.36e-02 (2.20e-03)	-1.49e-04 (1.20e-03)	2.04	0.04
MD Median	7.98e-04 (3.44e-05)	3.43e-06 (1.25e-05)	-4.50	1.01e-05
MD kurtosis	18.53 (6.39)	6.72e-03 (4.49)	-0.025	0.98
MD skew	3.10 (0.61)	7.21e-02 (0.43)	-2.78	0.01
FA pkval	0.33 (6.70e-02)	-3.10e-03 (7.46e-02)	0.68	0.50
FA PH	3.36e-03 (2.97e-04)	9.60e-05 (2.76e-04)	-5.72	2.89e-08
FA Median	0.34 (2.87e-02)	1.85e-03 (2.09e-02)	-1.45	0.15
FA kurtosis	0.43 (0.34)	4.84e-02 (0.21)	-3.80	1.81e-04
FA skew	0.67 (0.14)	1.48e-02 (8.60e-02)	-2.82	5.17e-03
AD pkval	1.06e-03 (5.07e-05)	3.42e-06 (5.04e-05)	-1.12	0.27
AD PH	8.80e-03 (8.80e-04)	4.13e-05 (8.29e-04)	-0.82	0.41
AD Median	1.12e-03 (3.87e-05)	7.96e-06 (2.00e-05)	-6.54	3.08e-10
AD kurtosis	9.02 (2.42)	0.29 (2.06)	-2.30	0.02
AD skew	2.04 (0.33)	9.25e-02 (0.29)	-5.23	3.45e-07
RD pkval	6.29e-04 (4.35e-05)	3.16e-07 (3.31e-05)	-0.16	0.88
RD PH	1.02e-02 (1.16e-03)	1.40e-05 (8.68e-04)	-0.27	0.79
RD Median	6.42e-04 (3.79e-05)	2.04e-06 (1.57e-05)	-2.14	0.03
RD kurtosis	1.28 (3.92)	0.70 (3.47)	-3.29	0.01
RD skew	2.38 (0.41)	0.14 (0.38)	-6.05	4.79e-09

Table 9 Change in conventional DTI measures over time in HARMONISATION.
 Except for FA location there was a significant change for all other DTI markers over the 2 years

DTI all WM marker	Mean baseline value (SD)	Mean 2 year change (SD)	Paired t-test	P-Value
MD pkval	7.68e-04 (3.76e-05)	8.25e-06 (3.02e-05)	-3.08	2.58e-03
MD PH	0.012 (2.13e-03)	-6.30e-04 (1.26e-03)	5.65	1.00e-07
MD Median	8.82e-04 (6.08e-05)	2.24e-05 (3.62e-05)	-6.96	1.65e-10
MD kurtosis	8.71 (2.52)	-0.69 (1.33)	5.80	4.97e-08
MD skew	2.94 (0.41)	-0.12 (0.22)	6.22	6.76e-09
FA pkval	0.07 (0.02)	-1.10e-03 (0.02)	0.75	0.45
FA PH	5.32e-03 (6.53e-04)	1.76e-04 (5.83e-04)	-3.39	9.37e-04
FA Median	0.19 (0.02)	-5.33e-03 (0.02)	3.53	5.77e-04
FA kurtosis	0.99 (0.82)	0.13 (0.63)	-2.36	0.02
FA skew	1.28 (0.26)	0.06 (0.19)	-3.22	1.64e-03
AD pkval	1.03e-03 (5.30e-05)	1.10e-05 (5.14e-05)	-2.41	0.02
AD PH	8.22e-03 (1.00e-03)	-2.52e-04 (7.83e-04)	3.63	4.08e04
AD Median	1.13e-03 (5.49e-05)	2.01e-05 (2.86e-05)	-7.93	1.00e-12
AD kurtosis	3.72 (0.924)	-0.27 (0.60)	5.13	1.09e-06
AD skew	2.16 (0.22)	-0.07 (0.14)	5.78	5.66e-08
RD pkval	6.49e-04 (4.78e-05)	1.40e-05 (4.15e-05)	-3.80	2.26e-04
RD PH	8.94e-03 (1.38e-03)	-4.05e-04 (8.74e-04)	5.23	7.00e-07
RD Median	7.80e-04 (6.57e-05)	2.34e-05 (3.99e-05)	-6.62	9.50e-10
RD kurtosis	5.09 (1.51)	-0.37 (0.82)	5.05	1.53e-06
RD skew	2.35 (0.31)	-0.09 (0.17)	5.61	1.25e-07

Figure 4. Heatmap showing the magnitude of positive and negative Pearson’s correlation coefficients between all DTI measures and the variance explained by the principal components capturing the variance of the DTI measures. The red and blue tiles refer to positive and negative correlation respectively. The overall correlation and the percentage of explained variance were lower in both SVD cohorts than in the MCI cohort.

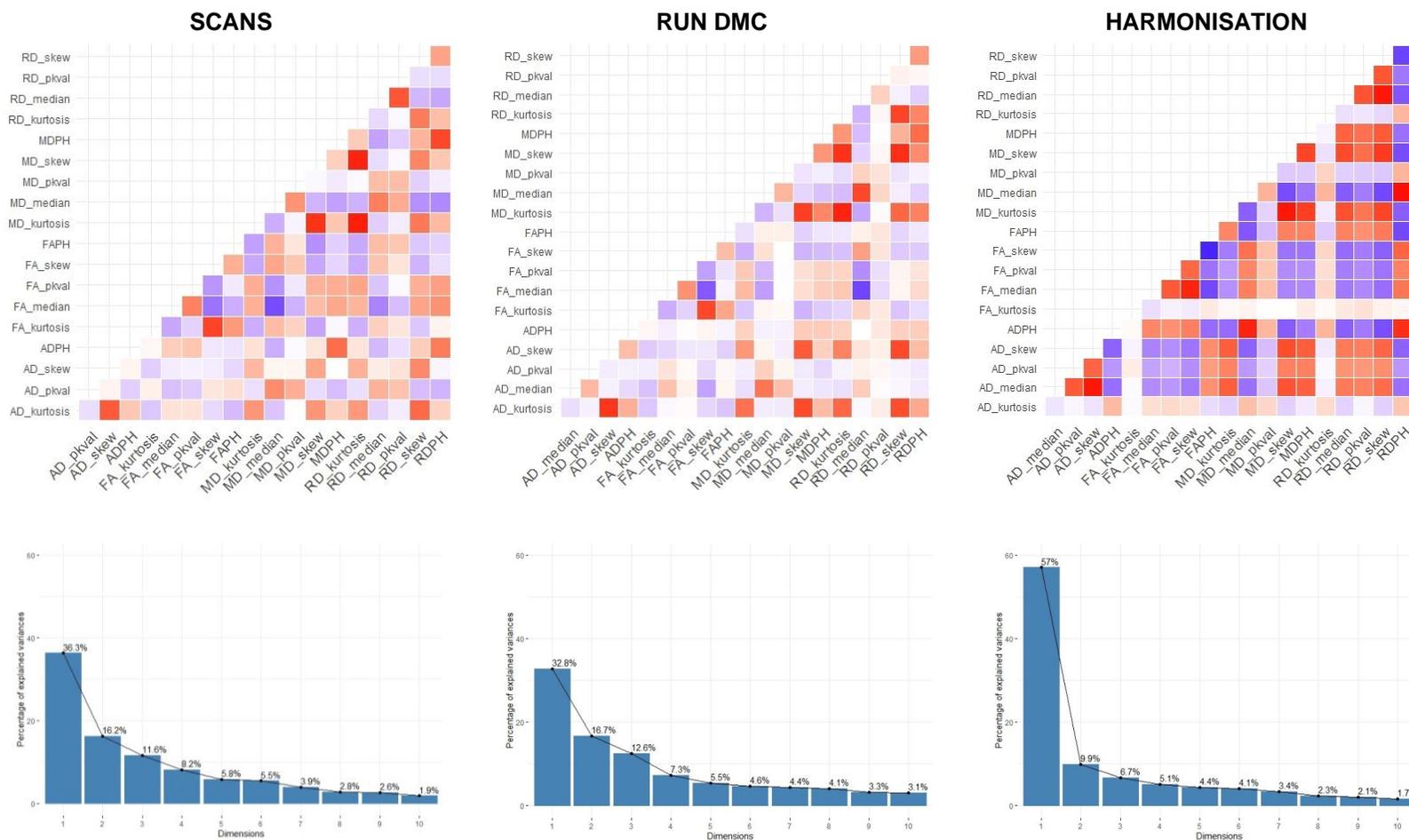


Table 10: Change in DTI markers after removal of observations post dementia diagnosis in SCANS. All markers significantly changed over the 3 years.

	Estimated mean baseline value (SE)	Estimated mean annual change (SE)	Wald-test	P-Value
MD Median	7.98e-04 (4.07e-06)	5.43e-06 (6.06e-07)	8.96	< 0.001
PSMD	3.78e-04 (1.05e-05)	1.36e-05 (1.89e-06)	7.21	<0.001
DSEG	20.19 (0.84)	1.17 (0.09)	13.42	<0.001
Geff	8.12 (0.23)	-0.18 (0.03)	-6.68	<0.001

Table 11. Change in DTI and dementia conversion. Change in DTI was consistently associated with dementia conversion only in severe SVD but not in mild SVD or MCI. The AUC was highest and AIC lowest for DSEG. Change in MD Median was associated with dementia conversion in HARMONISATION.

	Dementia conversion											
	SCANS				RUN DMC				HARMONISATION			
<i>Change Markers</i>	<i>HR (95% CI)</i>	<i>P-Value</i>	<i>AIC</i>	<i>AUC</i>	<i>OR (95% CI)</i>	<i>P-Value</i>	<i>AIC</i>	<i>AUC</i>	<i>OR (95% CI)</i>	<i>P-Value</i>	<i>AIC</i>	<i>AUC</i>
MD Median	2.46 (1.58, 3.83)	<0.001	134.67	0.789	0.90 (0.49, 1.56)	0.71	83.73	0.890	1.60 (1.02, 2.63)	0.04	113.52	0.755
PC1	2.47 (1.52, 4.01)	<0.001	137.21	0.786	1.25 (0.58, 2.48)	0.55	83.53	0.896	0.71 (0.44, 1.12)	0.14	115.58	0.733
PSMD	2.34 (1.60, 3.43)	<0.001	134.21	0.813	1.09 (0.63, 1.73)	0.73	83.76	0.895	1.43 (0.91, 2.30)	0.11	115.28	0.743
DSEG	4.01 (2.21, 7.29)	<0.001	122.87	0.924	1.53 (0.81, 2.91)	0.19	82.16	0.903	1.42 (0.87, 2.35)	0.16	115.77	0.723
Geff	0.49 (0.32, 0.75)	<0.001	140.24	0.793	0.74 (0.41, 1.37)	0.33	82.96	0.897	1.06 (0.64, 1.75)	0.83	117.72	0.708

MD Median= Mean diffusivity Median of the all WM histogram, PC1= Scores of the first principal component, PSMD= Peak width of skeletonized mean diffusivity, DSEG= Diffusion tensor image segmentation, Geff= Global efficiency network measure, AIC= Akaike information criterion, HR= hazard ratio, OR= Odd's ratio, AUC= area under the curve, 95% CI= 95% confidence interval

Table 12. Firth's Bias-Reduced Logistic Regression in RUN DMC. Change in DTI was also not associated with dementia conversion in RUN DMC when running a Firth's Bias-Reduced Logistic Regression

	Dementia conversion		
	<i>RUN DMC</i>		
<i>Change Markers</i>	β (95% CI)	<i>Chi-square</i>	<i>P-Value</i>
MD Median	-0.09 (-0.67, 0.44)	0.11	0.75
PC1	0.26 (-0.52, 0.91)	0.45	0.51
PSMD	0.12 (-0.42, 0.54)	0.23	0.64
DSEG	0.40 (-0.23, 1.01)	1.51	0.22
Geff	-0.28 (-0.85, 0.33)	0.90	0.34

MD Median= Mean diffusivity median of the all WM histogram, PC1= Scores of the first principal component, PSMD= Peak width of skeletonized mean diffusivity, DSEG= Diffusion tensor image segmentation, Geff= Global efficiency network measure, β = standardized regression coefficient

3.3.6 Sample size estimation for a hypothetical clinical trial

Estimating the sample size of a hypothetical clinical trial with varying durations and varying treatment effect sizes in SCANS, RUN DMC, HARMONISATION and CADASIL showed that the change in DSEG required the lowest sample size estimate in all SVD cohort studies (Table 13). Whereas PSMD required the second lowest sample size in all sporadic SVD cohorts, MD Median was second in the CADASIL cohort. In HARMONISATION the required sample sizes were overall significantly higher than in most other cohort studies.

Table 13. Sample size estimation per treatment arm for a hypothetical clinical trial across cohorts with varying treatment effect sizes. DSEG required the lowest minimum sample size in SCANS, RUN DMC and CADASIL. The minimum sample size was higher in the MCI cohort HARMONISATION.

	SCANS			RUN DMC			HARMONISATION			CADASIL		
Duration of RCT	3 years (3 follow-up measurements)			4 years (1 follow-up measurement)			2 years (1 follow-up measurement)			1.5 years (1 follow-up measurement)		
Treatment effect	10%	20%	30%	10%	20%	30%	10%	20%	30%	10%	20%	30%
MD Median	1201	301	134	20215	4580	1849	3795	860	348	732	166	67
PC1	1571	394	176	1571	394	176	1571	394	176	1571	394	176
PSMD	1115	279	124	1443	327	132	3101	703	284	1276	289	117
DSEG	420	105	47	1175	267	108	2200	499	202	656	149	60
Geff	2378	595	265	74077	16780	6776	8383	1899	767	19068	4320	1745

MD Median= mean diffusivity median of the all WM histogram, PC1= scores of the first principal component, PSMD= peak width of skeletonized mean diffusivity, DSEG= diffusion tensor image segmentation, Geff= global efficiency network measure, RCT= randomized controlled trial

4. Discussion

Five promising imaging markers were tested across six different cohorts varying in SVD severity. The results showed that all markers were associated with impaired cognitive function across all SVD and MCI cohorts. The association was stronger in the severe SVD, MCI and CADASIL group compared to the mild SVD cohort in RUN DMC. All baseline markers further predicted later dementia conversion in all SVD cohorts. In MCI baseline Geff did not predict dementia conversion. DSEG and MD Median significantly changed over time across all cohorts. In contrast, PSMD and Geff did not change over time in the multicenter study. Changes in the 4 parameters and PC1 only predicted dementia conversion in SCANS but not in RUN DMC. Change in MD Median was associated with dementia conversion in HARMONISATION. The sample size estimation for a clinical trial demonstrated that DSEG and PSMD would require the lowest minimum sample size estimates across both sporadic SVD cohorts.

These findings show that recently developed markers such as DSEG and PSMD may be promising surrogate markers for a future phase II clinical trial in SVD. More conventional markers such as the MD Median or the PC1 requiring multimodal MRI sequences and more complex post-processing image analysis showed similar associations with the clinical outcomes but required a larger minimum sample size for a hypothetical clinical trial in sporadic SVD. DSEG may further be suitable as a clinical endpoint in a multicenter trial involving different scanners as it also significantly changed over time in PRESERVE. Results from the other severe SVD cohort, SCANS, further showed that significant changes in PSMD and DSEG predicted dementia conversion. The minimum sample size estimates for Geff varied greatly between the cohort studies with very high estimates in the RUN DMC and CADASIL study. Geff may therefore not be a good choice as a surrogate clinical endpoint in a trial with SVD but may be more suitable as measure for understanding the mechanistic processes underlying SVD progression.

There may be reasons why the parameters provided different results. Differences in markers' reliability may have influenced the longitudinal results. Markers with higher reproducibility may require a lower minimum sample size for a hypothetical clinical trial. As PSMD uses the center of the tracts and reduces the effect of more noisy voxels, it may have a higher reliability over time. When it comes to more global measures such as DSEG and Geff, reproducibility should also be fairly high as noise

will have less of an effect. Differences in reproducibility may however only be half of the story when it comes to differences in minimum sample sizes. The other aspect is the marker's sensitivity to change. As shown in this Chapter in the multicentre study, no change could be detected in Geff, in contrast to DSEG or MD Median. It may therefore be that also in single-center studies such as SCANS, RUN DMC or CADASIL the sensitivity to change may be easily impacted by certain noise artifacts resulting in a significant increase in minimum sample size.

The study has limitations. First different duration and number of follow-up time points were used which may have affected the findings across cohorts with varying SVD severity. Consistent associations between changes in DTI and dementia conversion was not only found in the cohort characterized by a more severe SVD progression but also by the one having multiple follow-up imaging sessions. Annual follow-up time points in RUN DMC may have resulted in similar findings as in SCANS. Second, different MRI acquisition parameters were employed for the different cohorts potentially affecting the results. However, as shown in PRESERVE marked by multiple MRI acquisition parameters across sites certain markers such as DSEG or MD Median still demonstrated consistent findings. It still though remains to be determined whether these consistent findings also reflect high reproducibility of these markers.

To conclude, recently developed DTI markers such as DSEG and PSMD may be suitable as surrogate markers in a clinical trial in SVD. Strong associations with impaired cognitive function, marked changes over time also in multicentre studies for DSEG and a significant prediction between the markers' changes and dementia conversion makes them potentially useful to be surrogate clinical endpoints in patients with severe SVD.

Chapter 5

How reproducible are the DTI markers?

1. Introduction

Results from the previous chapters showed that recently developed DTI markers are promising as surrogate markers for a clinical trial in SVD. It was further demonstrated that more than 100 patients per treatment arm would still be required to detect a treatment effect of 20% in any SVD group. Given this sample size, it is unlikely that a clinical trial design would be single-centre but rather multi-centre possibly with different scanner types. As shown in Chapter 2, however, DTI markers may depend on the scanner types therefore impacting the precision of measurements and their reliability as surrogate markers. The required sample size may have to be considerably larger when accounting for multi-scanner variations. Not doing so, may compromise the statistical power in the trial leading to a low likelihood detection rate of finding a possible treatment effect. Determining reproducibility of the imaging markers not only within-scanners but also between scanners is therefore essential.

Previous evidence on the reproducibility of MRI markers in SVD has been summarized²⁸⁶. Most studies so far focused on brain volume (BV), which can be estimated with high precision using automated methods²⁸⁷ and which is a central marker in other neurological diseases such as AD²⁸⁸. BV showed a higher reproducibility within, than between, centres²⁸⁶. There were several factors that may affect measurement inter-center reliability such as effects of shifts in magnetic fields or coil effects²⁸⁶. WMH showed a high reproducibility within-centre, while evidence regarding between-center reproducibility is still lacking²⁸⁶. The SPRINT-MIND MRI study though demonstrated that WMH may be an important surrogate marker also in multi-scanner clinical trial settings¹⁰¹. Evidence regarding the reproducibility of other SVD markers such as lacunes, CMB or perivascular spaces has been sparse and inconclusive²⁸⁶.

Similarly to most conventional MRI markers, evidence regarding DTI's reproducibility in SVD has been sparse. In 7 CADASIL patients inter-scanner reproducibility of PSMD was high when employing 2 different scanners with various field strengths (1.5 vs 3T)¹¹³. In 10 sporadic SVD patients within-centre reproducibility was high for

all WM histogram measures such as median MD, MDPH with a mean percentage difference below 1% when being scanned twice within a few hours ⁸⁵.

This chapter aims to gather further evidence regarding the reproducibility of DTI markers, which have been developed for SVD. We determined the DTI markers' reproducibility across 3 different study designs. First, the reproducibility of DTI markers between scanner types (Siemens vs. Phillips) will be assessed in elderly participants being scanned twice within a few weeks period. Second, the reproducibility of DTI markers within the same session and between visits (few weeks apart) will be compared in young healthy controls when using the same scanner. Third, the markers' reproducibility between visits being a few weeks apart will be determined in SVD patients and compared to non-SVD participants when employing the same scanner type.

2. Method

2.1 Participants

2.1.1 Biocog

14 patients from the Biocog initiative (Biocog) were scanned twice within a few weeks apart employing 2 different MRI vendors located in Berlin and Utrecht. Patients were recruited to develop biomarkers that allow risk prediction of postoperative cognitive impairment in elderly patients ²⁸⁹. The inclusion criteria were being age 65 or older, eligible for surgical intervention, surgical operating time more than 1 h, hospital length of stay longer than 7 days.

2.1.2 WBIC DTI study

26 healthy participants with no history of neuropsychiatric disorder or substance abuse and a age range of 23-55 years were enrolled in the study ²⁹⁰. 22 of them were scanned twice at 2 visits in Cambridge with a mean range of 33 days and were included in this analysis. The study was designed to assess inter-subject variability and reproducibility of DTI data in a healthy population.

2.1.3 DTI-Network study

35 participants (26 SVD, 19 controls) were recruited for this study²⁹¹. SVD patients were age-matched with control subjects being 66 and 68 years of age on average respectively. Participants were recruited to assess the reproducibility of structural and functional network measures. Inclusion criteria for the SVD group were (1) history of clinical lacunar stroke syndrome with MRI evidence of an anatomically appropriate lacunar infarct, (2) presence of confluent WMH (Fazekas scale ≥ 2)⁷³. Exclusion criteria were any cause of stroke other than SVD. Repeat scan data was available for 25 participants (15 SVD, 10 controls), which were included in this analysis.

2.2 MRI acquisition

2.2.1 Biocog

Patients were imaged on a 3 T Magnetom Trio MRI system (Siemens AG, Erlangen, Germany) and on a Phillips Achieva 3 T MRI scanner employing 12-channel head coil. Multishell DTI data were obtained in both locations employing 30 non-collinear directions at b-values of 0, 1000, 2500 s/mm^2 on the Siemens scanner and 0, 1000, 3000 s/mm^2 on the Phillips scanner. The respective voxel size was 2.5 x 2.5 x 2.5 mm^3 , $TR= 6500ms$, $TE= 100 ms$ for the 3T Magnetom Trio MRI system and 0.96 x 1.19 x 4 mm^3 , $TR= 3294 ms$, $TE= 68 ms$ for the 3T Phillips Achieva.

2.2.2 WBIC DTI study

Participants were imaged using a 3T Siemens Verio MRI scanner (Siemens AG, Erlangen, Germany) with a 32-channel head coil. The structural sequences were 3D T1- weighted magnetization prepared rapid gradient echo (MPRAGE), fluid attenuated inversion recovery (FLAIR), gradient echo and dual spin echo (proton density/ T2-weighted). The DTI data were obtained employing 63 non-collinear directions, $b = 1000 s/mm^2$ with one volume acquired without diffusion weighting ($b = 0$), $TE= 106 ms$, $TR= 11700 ms$, 63 slices and 2 mm^3 isotropic voxels.

2.2.3 DTI-Network study

Participants underwent imaging on a 3 T Verio MRI system (Siemens AG, Erlangen, Germany) with a 32-channel head coil. The following structural sequences were acquired to compute the SVD markers: 1 mm volumetric T1 weighted MPRAGE,

0.9375 × 0.9375 × 2 mm T2 weighted FLAIR, 0.86 × 0.86 × 5 mm T2* weighted gradient echo. Axial single shot T2*-weighted EPI sequence with diffusion-weighted images ($b = 1000 \text{ s/mm}^2$) were acquired in 63 non-collinear directions with $TR= 11700 \text{ ms}$, $TE= 106 \text{ ms}$ and 63 contiguous 2 mm slices. Eight non-diffusion weighted images ($b = 0 \text{ s/mm}^2$) were acquired.

2.3 MRI analysis

The following markers being used in this chapter have been fully described in Chapter 4:

- Peak Width of Skeletonized Mean Diffusivity (PSMD)
- Diffusion Tensor Image Segmentation (DSEG)
- Global efficiency network measure (Geff)
- Median mean diffusivity all WM histogram (MD Median)

Mean Skeletonized Mean Diffusivity (MSMD) was additionally computed, as the recently developed measure was proposed to be more stable than the PSMD by the developers (<http://www.psm-d-marker.com/>). The MSMD also uses the tract-based spatial statistics (TBSS) pipeline as PSMD. But instead of taking the peak width, the mean of the MD histogram is calculated as the final measure.

The Biocog multi-shell data were reduced to single-shell data ($b= 1000 \text{ s/mm}^2$) in order to compute DSEG, PSMD, MSMD and Geff.

2.4 Statistical Analysis

To assess the reproducibility of markers between scanner types, within a MRI session and between visits, 2-way random-effects model was employed and the intra-class correlation coefficient (*ICC*) was calculated. This model was used as all subjects were rated by the same set of raters and as there was systematic variance both in the rates and in the raters²⁹². Selecting this model, the aim was to generalize the reliability evidence to any raters who have very similar characteristics like the raters chosen in this study²⁹³.

Further, the coefficient of variation in percentage (*CV (%)*) was employed in order to estimate the extent of variability in reference to the mean of the patient group. DTI images of outlier values were checked and it was determined whether the outlier was due to the processing pipeline.

Running a linear regression, it was also tested whether age was associated with the markers' between-visit reproducibility in each session of the WBIC-DTI study. The absolute difference of the markers between the visits was taken as the criterion for the linear model. The underlying assumptions of the regression model such as normality of residuals, homoscedasticity and linear relationship were met.

3. Results

3.1 Between-scanner reproducibility in Biocog

There was low to moderate reproducibility ($ICC < 0.75$) between scanner types in all imaging markers in the Biocog data (Table 1). ICC was lower in PSMD compared to the other markers. Follow-up analysis showed that taking out 2 patients significantly increased the ICC for PSMD and Geff. No quality DTI data issues could, however, be detected in these 2 cases and they therefore remained in the analysis. Analysis however showed that MD distributions of the whole brain when CSF is removed were significantly different between the DTI images for these patients, which may have likely affected PSMD and Geff but less DSEG (Figure 1). Similar discrepancies in histogram distributions were not found for the other participants.

3.2 Within- and between reproducibility in WBIC DTI study

3.2.1 Within-session reproducibility in the WBIC DTI study

Within-session reproducibility ranged between 0.726-0.915 with PSMD being the lowest and MD Median being the highest (Table 2). One visit involving 2 sessions was excluded due to data quality issues in one session on all markers (Figure 2). Reproducibility was significantly higher for MSMD ($ICC= 0.906$) and MD Median ($ICC= 0.915$) than PSMD ($ICC= 0.726$) and Geff ($ICC= 0.767$).

3.2.2 Between-visit reproducibility in the WBIC DTI study

Between-visit reproducibility ranged between 0.502-0.805 with DSEG ($ICC= 0.631$) and Geff ($ICC= 0.502$) being the lowest and MD Median ($ICC= 0.805$) and MSMD ($ICC= 0.716$) being the highest (Table 3).

3.3 Between-visit reproducibility in the DTI-Network study

The DTI-Network study overall demonstrated high reproducibility across most markers (Table 4, Table 5). In SVD and control patients the *ICC* was significantly lower in PSMD and MSMD. Subsequent analysis demonstrated that 2 PSMD outlier values in each subgroup were due to problems in the markers' preprocessing pipeline (Figure 4, Figure 5). PSMD's reproducibility significantly increased in the SVD group but not in the health control group after excluding 2 patients showing image-preprocessing issues (Table 4, Table 5).

3.4 Between-visit reproducibility and age

Age was not associated with any between-visit reproducibility for any marker in the WBIC DTI study (Table 6, Figure 6).

Table 1. Reproducibility of the markers between Siemens and Phillips scanners. Reproducibility was lower in PSMD compared to other markers. Removing 2 outlier cases significantly improved the ICC in PSMD, MSMD and Geff.

Comparing Scanner sites differences					
Marker (Number of pat.)	Scanner type	Mean (SD)	CV (%)	ICC (95% CI) N= 14	ICC (95% CI) N= 12
PSMD (N= 14)	Siemens	4.04 e-04 (8.85e-05)	21.88	0.314 (-0.116- 0.689)	0.528 (-0.07- 0.844)
	Phillips	3.40e-04 (5.18e-05)	15.24		
MSMD (N= 14)	Siemens	8.41e-04 (5.77e-05)	6.86	0.605 (0.093- 0.858)	0.859 (0.454- 0.961)
	Phillips	8.12e-04 (4.45e-05)	5.49		
DSEG (N= 14)	Siemens	20.59 (7.61)	36.95	0.642 (-0.09- 0.901)	0.715 (-0.072- 0.930)
	Phillips	26.13 (5.76)	22.04		
Geff (N= 14)	Siemens	1.63 (0.46)	28.00	0.553 (0.08- 0.828)	0.820 (-0.054- 0.964)
	Phillips	1.80 (0.68)	37.93		

PSMD= peak width skeletonized mean diffusivity, DSEG= diffusion tensor image segmentation, Geff= global efficiency network measure, MSMD= mean skeletonized mean diffusivity, MD Median= median of the mean diffusivity histogram, CV (%)=coefficient of variation in percentage, ICC (95% CI)= intraclass correlation coefficient with 95% CI included

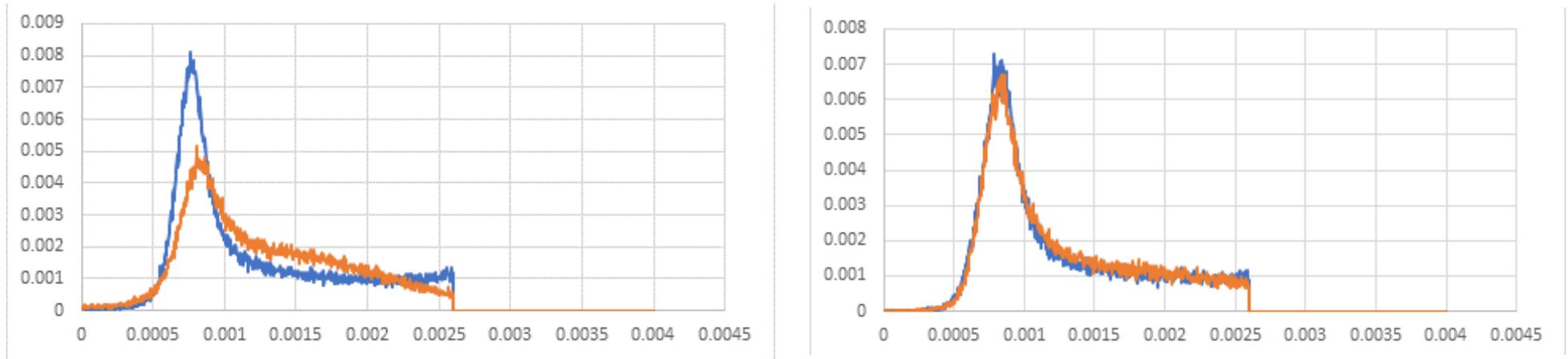
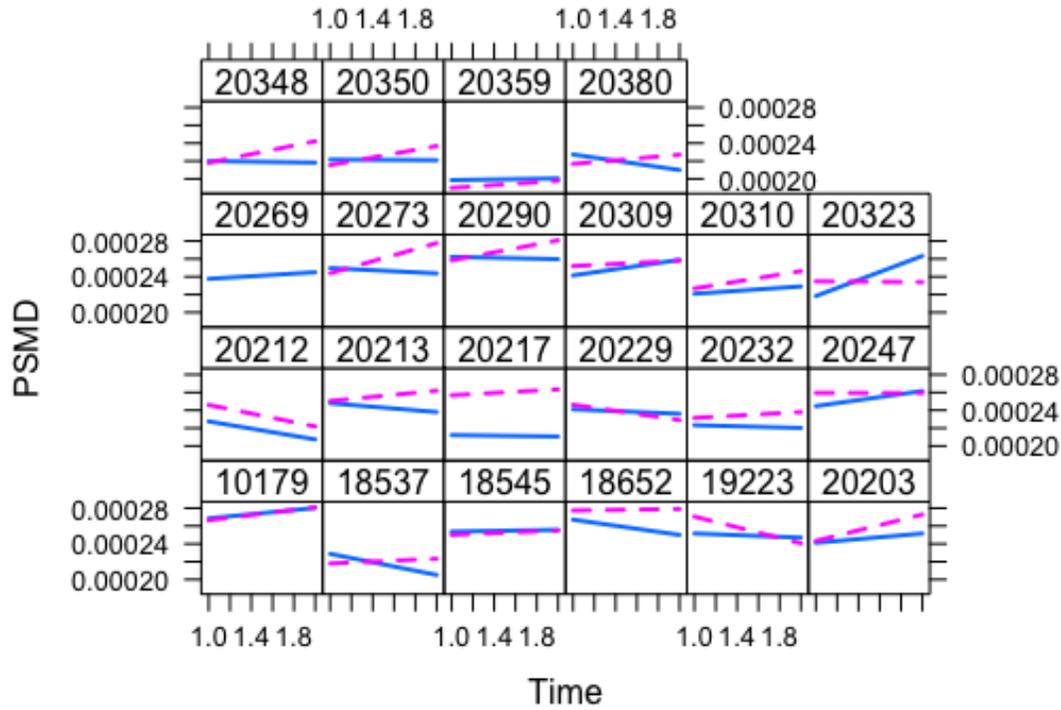


Figure 1. MD histogram distributions for the whole brain with the CSF thresholded out. The blue line is Philips and the orange line is Siemens. The left panel shows a patient having an outlier value in PSMD with a significant difference in the histograms. The right panel shows a patient with no difference in histograms and with no outlier PSMD values.

(A)



(B)

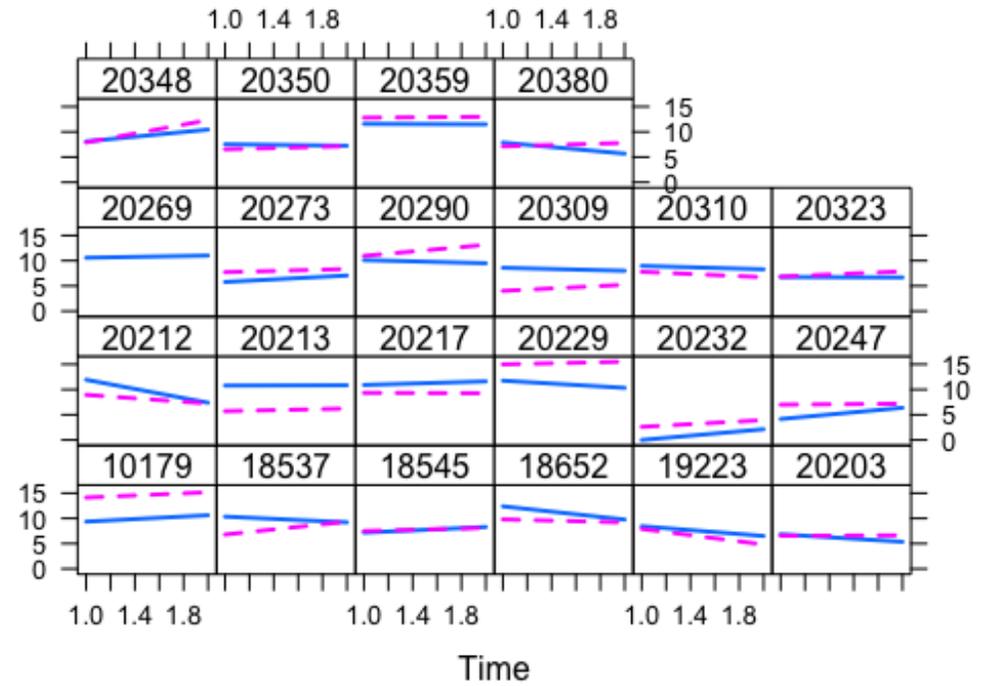


Figure 2. The automatic and semiautomatic markers PSMD (A) and DSEG (B) over time in each participant in the WBIC DTI study. Blue and pink line refers to the visit 1 and visit 2 respectively. The more horizontal the lines are, the more reproducible the measures are in a session.

Table 2. Reproducibility of the markers within the same session in the WBIC DTI study. Reproducibility varied across the markers with the highest reproducibility in MSMD and MD Median and the lowest in PSMD and Geff.

Reproducibility of markers within the same session in healthy people				
Marker (Number of patients)	Session	Mean (SD)	CV (%)	ICC (95% CI) (N= 22)
PSMD (N= 22)	Session 1	2.39e-04 (1.99e-05)	8.34	0.726 (0.548- 0.841)
	Session 2	2.43e-04 (2.32e-05)	9.57	
DSEG (N= 22)	Session 1	8.44 (2.93)	34.73	0.837 (0.719- 0.908)
	Session 2	8.56 (2.83)	33.02	
Geff (N= 22)	Session 1	5.63 (0.97)	17.15	0.767 (0.609- 0.867)
	Session 2	5.57 (1.10)	19.76	
MSMD (N= 22)	Session 1	7.53e-04 (1.46e-05)	1.94	0.906 (0.829- 0.948)
	Session 2	7.55e-04 (1.57e-05)	2.08	
MD Median (N= 22)	Session 1	7.88e-04 (1.88e-05)	2.39	0.915 (0.849- 0.953)
	Session 2	7.89e-04 (1.84e-05)	2.33	

PSMD= peak width skeletonized mean diffusivity, DSEG= diffusion tensor image segmentation, Geff= global efficiency network measure, MSMD= mean skeletonized mean diffusivity, MD Median= median of the mean diffusivity histogram, CV (%)=coefficient of variation in percentage, ICC (95% CI)= intraclass correlation coefficient with 95% CI included

Table 3. Reproducibility of the markers between visits in the WBIC DTI study. Reproducibility was low to moderate (ICC < 0.75) in all markers except for MD Median.

Reproducibility of markers between visits in health people				
Marker (Number of patients.)	Visit	Mean (SD)	CV (%)	ICC (95% CI) (N= 22)
PSMD	Visit 1	2.37e-04 (2.06e-05)	8.70	0.675 (0.389- 0.828)
	Visit 2	2.46e-04 (2.22e-05)	9.03	
DSEG	Visit 1	8.45 (2.64)	31.27	0.631 (0.409- 0.782)
	Visit 2	8.51 (3.08)	36.22	
Geff	Visit 1	5.55 (1.12)	20.24	0.502 (0.242- 0.696)
	Visit 2	5.68 (0.97)	17.11	
MSMD	Visit 1	7.55e-04 (1.45e-05)	1.92	0.716 (0.534- 0.835)
	Visit 2	7.52e-04 (1.57e-05)	2.09	
MD Median	Visit 1	7.88e-04 (1.79e-05)	2.28	0.805 (0.668- 0.889)
	Visit 2	7.89e-04 (1.93e-05)	2.44	

PSMD= peak width skeletonized mean diffusivity, DSEG= diffusion tensor image segmentation, Geff= global efficiency network measure, MSMD= mean skeletonized mean diffusivity, MD Median= median of the mean diffusivity histogram, CV (%)=coefficient of variation in percentage, ICC (95% CI)= intraclass correlation coefficient with 95% CI included

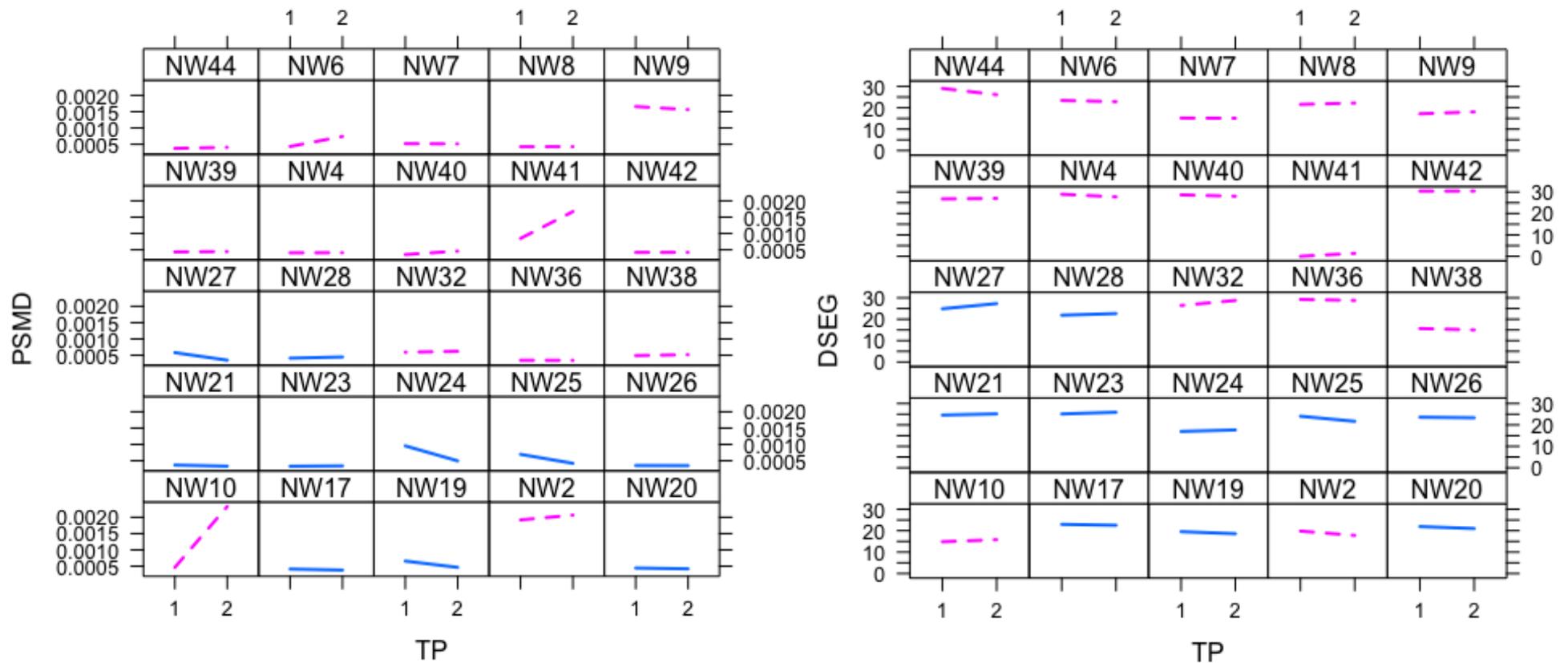


Figure 3. PSMD and DSEG over time in each participant in the DTI-Network study. Blue and pink lines refer to the control group and SVD group respectively. There were 2 significant PSMD outlier values in the SVD group (NW41, NW10). There were also 2 outlier values for the control group (NW24, NW25). In contrast, DSEG showed no outlier values.

Table 4. Reproducibility of imaging markers between visits in the DTI-Network study when including SVD patients only. Reproducibility was lower in the PSMD and MSMD than in the other markers. Reproducibility of the PSMD was significantly compromised due to 2 patients.

Reproducibility of markers over time SVD study					
Marker (Number of pat.)	TP	Mean (SD)	CV (%)	ICC (95% CI) N= 15)	ICC (95% CI) (N= 13)
PSMD	Visit 1	6.46e-04 (4.82e-04)	74.60	0.604 (0.179- 0.844)	0.980 (0.935- 0.994)
	Visit 2 (within 6 m)	8.63e-04 (6.79e-04)	78.66		
DSEG	Visit 1	21.81 (8.28)	38.00	0.987 (0.963- 0.996)	-
	Visit 2 (within 6 m)	21.68 (7.87)	36.31		
Geff	Visit 1	5.97 (2.76)	46.13	0.956 (0.877- 0.985)	-
	Visit 2 (within 6 m)	5.76 (2.86)	49.53		
MSMD	Visit 1	8.30e-04 (7.12e-05)	8.59	0.459 (-0.058- 0.780)	0.512 (-0.036- 0.821)
	Visit 2 (within 6 m)	8.44e-04 (1.02e-04)	12.05		
MD Median	Visit 1	9.26e-04 (6.80e-05)	7.34	0.984 (0.954- 0.995)	-
	Visit 2 (within 6 m)	9.22e-04 (6.57e-05)	7.13		

PSMD= peak width skeletonized mean diffusivity, DSEG= diffusion tensor image segmentation, Geff= global efficiency network measure, MSMD= mean skeletonized mean diffusivity, MD Median= median of the mean diffusivity histogram, CV (%)=coefficient of variation in percentage, ICC (95% CI)= intraclass correlation coefficient with 95% CI included

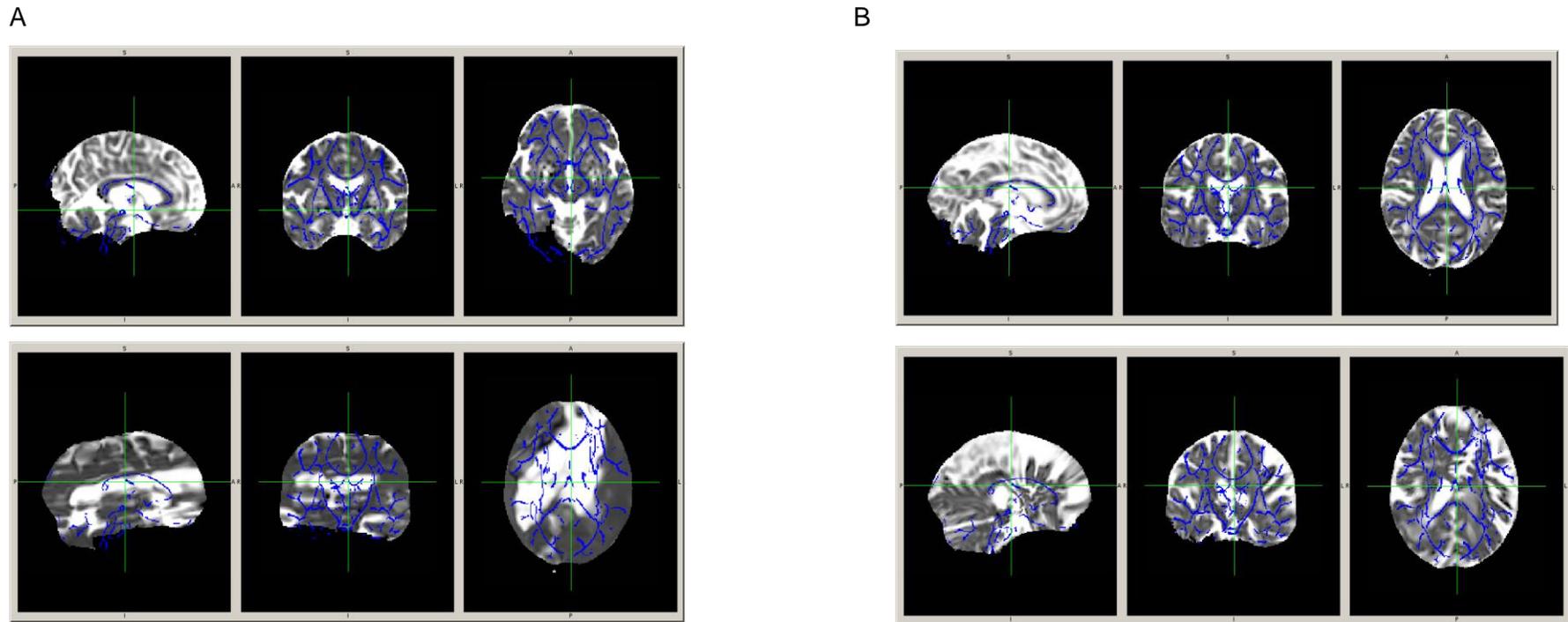


Figure 4. FA mean non-masked skeleton (blue) against the MD image underlying the 2 PSMD outlier cases in the DTI-Network study. In both patients (NW10, NW41), see panel A and panel B respectively, there were problems with segmentation and registration in the automatic PSMD pipeline.

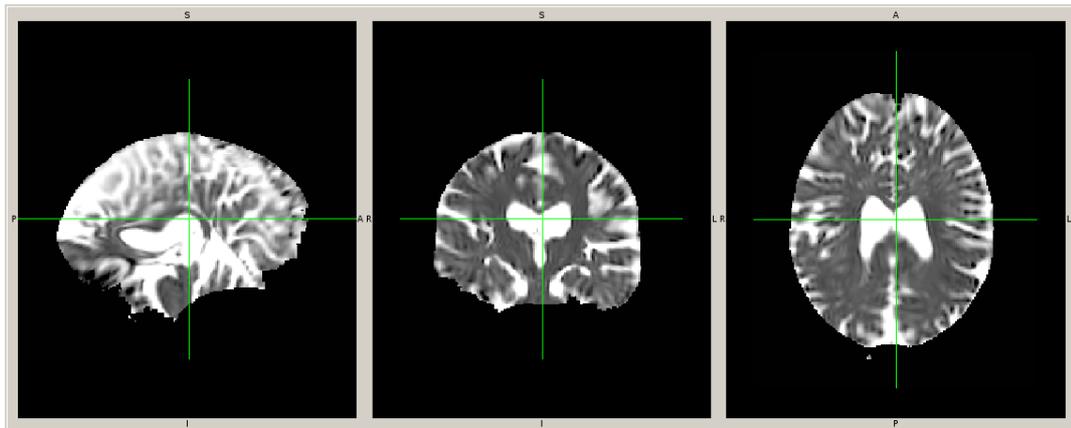
Table 5. Reproducibility of imaging markers between visits in the DTI-Network study when including healthy participants only.

Reproducibility was lower in the PSMD and MSMD than in the other markers. Reproducibility was increased for MSMD when excluding 2 outlier cases.

Reproducibility of markers over time in healthy participants					
Marker (Number of pat.)	TP	Mean (SD)	CV (%)	ICC (95% CI) (N= 10)	ICC (95% CI) (N= 8)
PSMD	Visit 1	5.22e-04 (2.00e-04)	38.26	0.306 (-0.180- 0.739)	0.345 (-0.25- 0.807)
	Visit 2 (within 6 m)	4.01e-04 (5.74e-05)	14.32		
DSEG	Visit 1	22.55 (2.60)	11.53	0.903 (0.654- 0.975)	
	Visit 2 (within 6 m)	22.58 (3.06)	13.53		
Geff	Visit 1	11.18 (2.77)	24.75	0.871 (0.567- 0.966)	
	Visit 2 (within 6 m)	22.58 (3.06)	13.53		
MSMD	Visit 1	7.28e-04 (2.29e-05)	3.15	0.327 (-0.354- 0.778)	0.643 (-0.104- 0.919)
	Visit 2 (within 6 m)	7.21e-04 (2.46e-05)	3.42		
MD Median	Visit 1	8.62e-04 (3.38e-05)	3.92	0.867 (0.548- 0.965)	
	Visit 2 (within 6 m)	8.61e-04 (3.53e-05)	4.10		

PSMD= peak width skeletonized mean diffusivity, DSEG= diffusion tensor image segmentation, Geff= global efficiency network measure, MSMD= mean skeletonized mean diffusivity, MD Median= median of the mean diffusivity histogram, CV (%)=coefficient of variation in percentage, ICC (95% CI)= intraclass correlation coefficient with 95% CI included

A



B

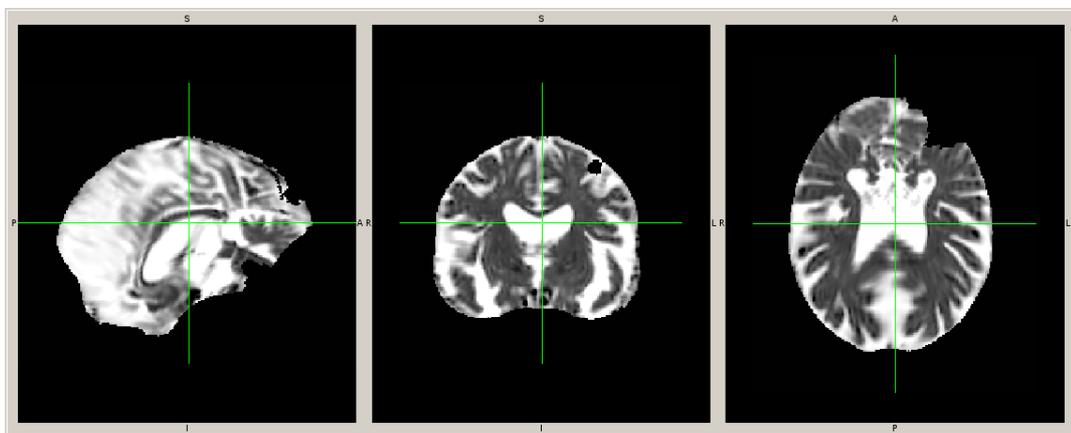


Figure 5. Problems in the segmentation of the DTI data associated with 2 PSMD outlier cases (NW24, NW25) in the DTI-Network study. In both subjects, see panel A and panel B respectively, there were segmentation problems in the automatic PSMD pipeline.

Table 6. Association between the markers' difference between visits and age in each session. There were no significant associations between the markers' difference between the visits and age.

Difference in markers between visits	Session 1	Session 2
PSMD	β (SE)= -0.322 (0.211), $p= 0.145$, Adj. $R^2= 0.062$	β (SE)= -0.105 (0.222), $p= 0.643$, Adj. $R^2= -0.039$
DSEG	β (SE)= 0.050 (0.224), $p= 0.827$, Adj. $R^2= -0.050$	β (SE)= 0.090 (0.223), $p= 0.691$, Adj. $R^2= -0.042$
Geff	β (SE)= -0.015 (0.027), $p= 0.603$, Adj. $R^2= -0.037$	β (SE)= -0.011 (0.027), $p= 0.691$, Adj. $R^2= -0.042$
MSMD	β (SE)= -0.047 (0.224), $p= 0.835$, Adj. $R^2= -0.050$	β (SE)= 0.164 (0.221), $p= 0.466$, Adj. $R^2= -0.022$
MD Median	β (SE)= 0.052 (0.224), $p= 0.819$, Adj. $R^2= -0.050$	β (SE)= 0.071 (0.223), $p= 0.754$, Adj. $R^2= -0.045$

β = standardized regression coefficient, SE= standard error, p = significance value, Adj. R^2 = adjusted explained R variance, PSMD= peak width skeletonized mean diffusivity, DSEG= diffusion tensor image segmentation, Geff= global efficiency network measure, MSMD= mean skeletonized mean diffusivity, MD Median= median of the mean diffusivity histogram

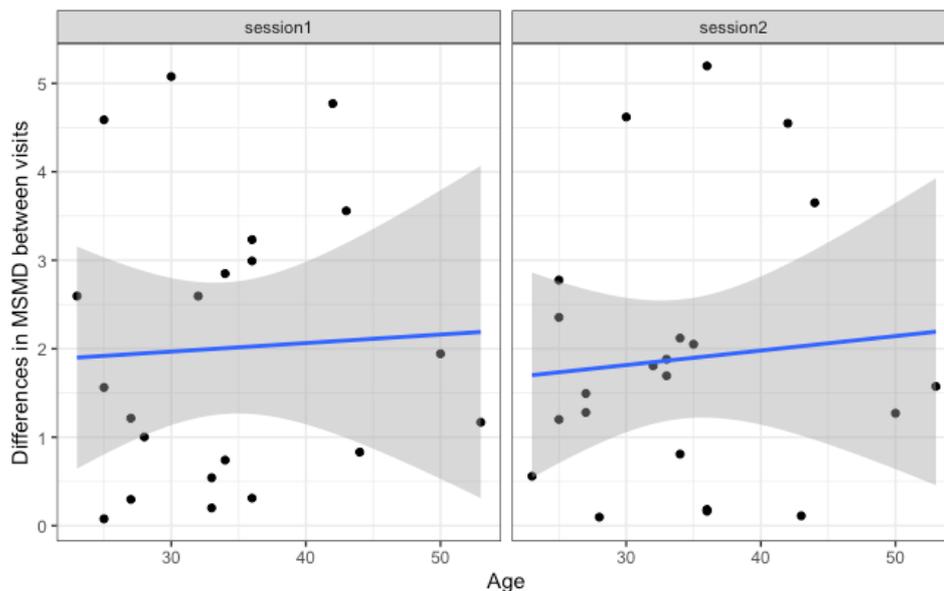


Figure 6. Association between age and difference in MSMD between visits. There was no association between age and the marker's difference in both sessions.

4. Discussion

The goal of this chapter was to assess the within-subject reproducibility of SVD-DTI markers. The reproducibility was assessed across a number of different designs and different populations: a) between scanner reproducibility in elderly patients, b) within-session and between-visit reproducibility in young healthy participants c) between-visit reproducibility in sporadic SVD patients and age-matched healthy controls.

Between-scanner reproducibility in the Biocog data was low to moderate ($ICC < 0.75$) with PSMD being the lowest. Follow-up analysis showed that the MD distribution was different in two patients between the scanner types. Excluding those patients significantly increased reproducibility in PSMD, MSMD and Geff. The MD distribution discrepancies are likely to be an issue with the scanner itself, rather than the measure as it only affects two subjects, but it was no apparent what the issue was and therefore it needs to be included in assessment of reproducibility.

Within-session reproducibility was higher than between-visit reproducibility in the WBIC-DTI data in healthy individuals aged 23-53 years of age. MSMD and MD Median showed the highest reproducibility in both designs. The higher within-session reproducibility may primarily be explained by subject specific factors. While the scanner type with the magnetic field strength and vendor, the coil with the number of channels, the gradient, the software employed and the post-processing steps taken were the same across all 4 sessions, the subject's motion, head positioning and physiological status may have varied.

In the DTI-Network study, which was marked by more advanced age, the between-visit reproducibility of the automatic marker PSMD and MSMD was significantly lower and produced more outliers due to image segmentation and registration issues than other markers such as the semi-automatic marker DSEG. These preprocessing issues in the automatic markers could have likely been manually corrected if needed. As PSMD and MSMD were designed to be automatic markers to be run in large cohorts, we used them in this automated way in our evaluation. The semi-automatic marker DSEG had the highest reproducibility in SVD as well as in matched control subjects.

It is unclear why the between-visit reproducibility is different for some of the markers when comparing WBIC-DTI to DTI-Network. One explanation may be that the factor

age affects reproducibility as the WBIC-DTI data had a younger age group included. The results in the WBIC-DTI data, however, showed that there was no association between age and the reproducibility for any of the imaging markers. An alternative explanation may be that these markers were primarily developed for an elderly patients group, which may have affected the reproducibility in the younger data cohort WBIC-DTI. While this may hold true for PSMD which is a marker originally developed for capturing the SVD burden mostly happening in older age, it may be not true for Geff which is a marker being employed across different age groups ^{294–296}. Comparing SVD to age-matched controls in DTI-Network though indicates that age may have little effect and the vascular burden is the primary factor for PSMD and MSMD. The reproducibility was significantly lower in the disease-free participants in comparison to the patients with SVD after the outlier cases were removed.

This is the first time that the reproducibility of the recently developed SVD-DTI measures has been compared. Most studies investigating reproducibility of DTI measures so far largely focused on participants from a young healthy population ^{297–300}. There is so far limited evidence showing how much reproducibility may be affected when including elderly participants. FA and MD overall had good reproducibility in healthy elderly participants ³⁰¹. While MD was overall more reproducible than FA when looking at the whole brain, FA was more reliable in subcortical WM regions. More evidence is, however, needed, also in clinical populations, as multi-scanner trials in elderly populations involving surrogate imaging measures are increasingly considered as a design of choice ^{101,302–304}. In such studies differences in DTI measures are interpreted as a progression of neurological disease or the biological responsiveness to a particular treatment. It is essential that studies do not only show potential treatment effects but also demonstrate reproducibility results across centers and studies employing different MRI acquisition protocols.

There are limitations. Biocog and Network-DTI study only had a small sample size. As a result, a few cases could have significantly influenced the reproducibility measures. Furthermore multi-shell DTI data had to be reduced into single-shell in Biocog, which may have further attenuated the overall reproducibility of the measures. These findings are preliminary and more studies designed to assess the reproducibility of markers in clinical populations are needed.

To conclude, between-scanner reproducibility was only low-to moderate when not examining raw DTI summary measures such as the whole-brain MD distributions. DTI markers such as MD Median and DSEG may be more reproducible than other markers such as PSMD. Despite its automatic nature, it is recommended that each step of the processing pipeline is checked in detail. It remains to be determined why the reproducibility was significantly higher in the DTI-Network study than in the WBIC-DTI study.

Chapter 6

Is serum neurofilament light chain a suitable surrogate marker for a clinical trial in SVD patients?

1. Introduction

The previous chapters showed that conventional DTI histogram markers and advanced DTI markers show promising features to be employed as surrogate markers in SVD. Particularly in severe SVD, baseline and change in all DTI measures predicted dementia conversion. The minimum sample size estimate for a hypothetical clinical trial was lowest in the sporadic SVD cohorts for PSMD and DSEG. It was furthermore demonstrated that in severe SVD the predictive power for dementia conversion increases when combining DTI with conventional MRI measures such as brain volume and lacune count.

Despite these valuable insights, there are still key limitations that should be considered when designing a clinical trial with imaging measures. First, to track the exact progression in the disease and to be able to limit the required patients' sample to a reasonable size, repeated imaging on multiple occasions is required which is expensive. In a 3-years follow-up hypothetical randomised clinical trial with severe SVD patients enrolled, a minimum sample size of 185 patients per treatment arm would be required when having annual repeated DTI imaging and a treatment effect size of 25%⁶⁶. Second, in multicentre clinical trials the issue of harmonization is an important aspect to consider. The findings of chapter 5 showed that imaging measures are only modestly reproducible when assessing them across different scanner sites.

Blood-based circulating biomarkers have 2 advantages that eliminate these 2 limitations. First, blood samples can be easily acquired on multiple occasions in the course of the clinical trial. Second, the challenge of harmonizing across sites can be prevented by processing the blood samples in one centralized laboratory.

Over the recent years, one circulating biomarker has caught increasing attention across the different research disciplines, which focuses on various neurological

diseases such as multiple sclerosis, AD or stroke ^{190,305}. Serum neurofilament light chain (NfL) is a marker released on neuronal axial injury into the CSF and blood ¹⁸⁹. Recent advancements in the technology allowed for reliable quantifications of low NfL concentrations levels in blood samples of healthy as well as diseased people making the marker more attractive ¹⁸⁹. When it comes to ischemic and lacunar stroke, NfL was higher at baseline and remained elevated at 3 months and 6 months follow-up before returning to normal levels at 15 months post stroke ^{306,307}. Higher NfL at baseline was associated with clinical measures such as disability as well as quantitative imaging measures such as WMH at baseline, new WM lesions, infarct volume and DTI markers ^{306,307}. These findings emphasize the potential usefulness of this marker in monitoring neuroaxonal injury in patients with ischemic strokes.

NfL may be a clinically useful marker in sporadic and monogenic SVD ¹⁹³. NfL levels were higher in SVD patients compared to controls and were associated with both established SVD imaging markers as well as clinical outcomes such as impaired cognitive function and disability. The longitudinal clinical utility of the marker is furthermore shown in CADASIL patients where baseline NfL predicted 7-year changes in disability and cognition as well as 17-year survival ³⁰⁸. The authors concluded that NfL may be a promising surrogate marker in clinical trials. A recent study subsequently showed that in mild SVD baseline NfL predicted incident lacunes as well as predicted executive function at follow-up independently of age, sex, education and depression ³⁰⁹. Baseline NfL did not, however, predict dementia conversion after adjusting for baseline age.

All these findings emphasize the need of investigating whether NfL may be a promising surrogate marker for a clinical trial in SVD. There are 3 criteria that a surrogate marker for a phase 2 clinical trial would need to fulfill: (a) able to predict clinical outcome, where changes triggered by a therapy on a surrogate marker correspond to changes in a clinically meaningful end point; and (b) change in the marker need be noticeable prospectively; (c) the minimum sample size required to show therapeutic efficacy should be feasible to reach as a target in a clinical trial ⁶⁶.

This chapter aims are three-fold. First, it will be determined whether baseline NfL is associated with cognitive impairment and disability in SVD patients with a lacunar infarct. Second, it will investigated whether baseline NfL predicts longitudinal clinical outcomes such as cognitive decline and dementia conversion and changes in SVD imaging markers in severe SVD patients. Third, it will be determined whether there is

a significant change in NfL over time and whether this change predicts dementia conversion.

2. Methods

2.1 Patients

Both baseline and follow-up data was used from the St George's Cognition and Neuroimaging in Stroke (SCANS) prospective study ²⁰¹. 121 patients with symptomatic SVD, defined as a clinical lacunar stroke syndrome with MRI evidence of an anatomically corresponding lacunar infarct, and with confluent regions of WMH graded ≥ 2 on the modified Fazekas scale ⁷³ were enrolled from 3 stroke services in South London (St George's Hospital, King's College Hospital and St Thomas' Hospital) at least 3 months post stroke. Exclusion criteria were: any cause of stroke other than SVD (cardio-embolic sources, large vessel disease with $>50\%$ stenosis in intra- or extra-cranial arteries), other central nervous system disorders, major psychiatric disorders and any cause of WM disease other than SVD. Written informed consent was obtained from all patients participating in the study. The study was granted ethical approved by Wandsworth Research Ethic Committee.

Subjects were followed up annually for 5 years with 6 incidences of strokes being recorded (4 lacunar strokes, 2 intracerebral haemorrhages). MRI was performed at baseline and after 1, 2 and 3 years. Information on dementia incidence was acquired annually. All blood samples had been collected at the time of MRI imaging. Of the 121 subjects recruited, blood was available for 113, and in 90 patients blood samples were available from at least at 2 time points. Details regarding the sample size for each marker per time point can be found in Figure 1. Due to a small sample size, observations coming from follow-up time point 4 and 5 were removed.

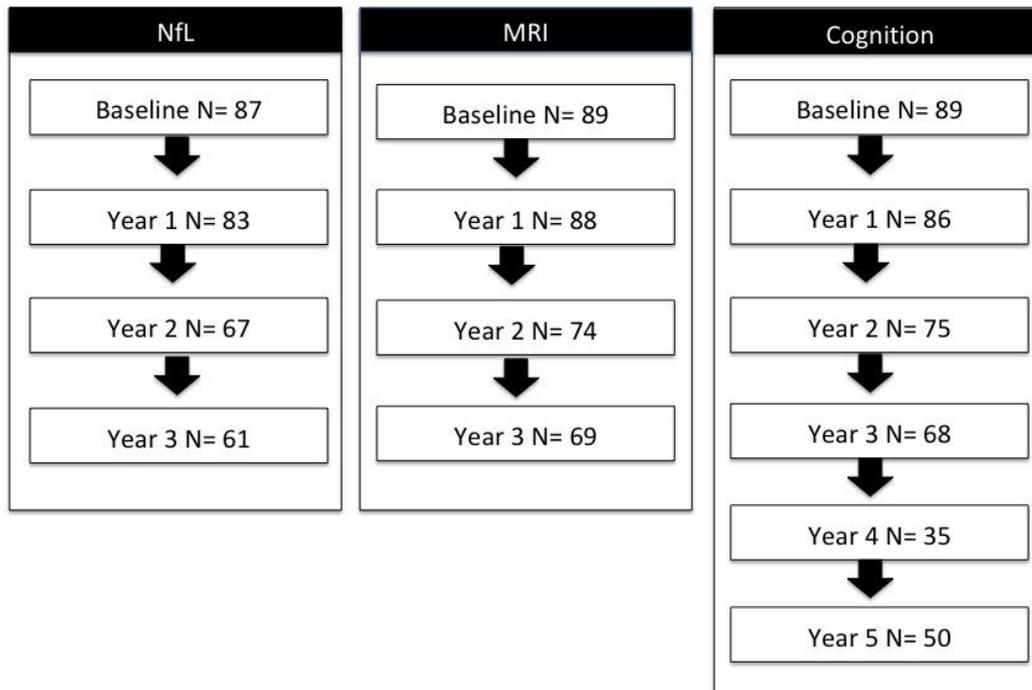


Figure 1. Longitudinal flow chart showing the number of patients per time point with serum neurofilament light chain (NfL) measures, imaging measures and cognitive measures.

NfL= neurofilament light chain, Imaging= imaging measures, Cognition= cognitive measures

2.2 Serum Neurofilament Light Chain

All assays were performed at the same time in 2019-20. All samples were analyzed on the same single-molecule array instrument (Simoa HD-1, Quanterix, Lexington, MA, USA) at the University of Basel by Prof. Nils Peters and Prof. Jens Kuhle. They employed the capture monoclonal antibody (mAb) 47:3 and the biotinylated detector mAb 2:1 (UmanDiagnostics, Umeå, Sweden)³¹⁰ transferred onto the Simoa platform. Bovine lyophilized NfL was received from UmanDiagnostics. Calibrators were between 0 to 2,000 *pg/mL*. Intra- and inter-assay variabilities were under 20%. The analytical sensitivity was 0.32 *pg/mL*. All samples had a signal over the assay's analytical sensitivity. The NfL analysis was performed blind to the dementia outcome measures.

2.3 MRI acquisition

Images were obtained using a 1.5-T General Electric Signa HDxt MRI system²⁰¹ employing the same image acquisition protocol at every time point. The details of MRI acquisition in SCANS were described in Chapter 3.

2.4 Conventional MRI markers

NBV, WMH lesion volume (WMH), lacune count and CMB count were computed as previously described in detail in Chapter 3.

2.5 Diffusion tensor imaging analysis

Two methods to assess WM ultrastructure on DTI were used:

2.5.1 All WM Histogram analysis measure

The analysis has been described previously in Chapter 2 and 3. The primary outcome measure was mean diffusivity peak height (MDPH) as this has previously been shown to be the most sensitive parameter to change in the SCANS study ^{44,201}.

2.5.2 Peak width of skeletonized mean diffusivity (PSMD)

The Peak width of skeletonized mean diffusivity (PSMD) is a fully automatically computed imaging marker being publicly available (<http://psmd-marker.com>) and has been fully described previously in Chapter 4. The computation is divided into four main modules: 1) DTI sequence on MRI, 2) WM tract skeletonization using tract-based statistics, 3) excluding CSF prone regions employing a custom mask, 4) histogram values of mean diffusion on the skeleton. The PSMD measure refers to the difference between the 5th and 95th percentiles of the histogram distribution.

2.6 Cognition and dementia diagnosis

A battery of well-established, standardized tasks sensitive to the cognitive impairments in SVD was carried out annually. Full details have been described in Chapter 3. Cognitive index scores were created by grouping the standardized measures into subdomain scores ⁶⁴: Global cognition score (Global), Executive function (EF) and Processing speed (PS).

Dementia was diagnosed using the DSM-5 definition of major neurocognitive disorder, as previously described in Chapter 3.

2.7 Disability

Disability was assessed using the modified Rankin Scale ^{311,312}, a 7-point scale measuring disability during stroke.

2.8 Statistical Analysis

NfL, WMH lesion load, lacune and CMB count were log 10 transformed due to the skewness of their data distribution. One outlier observation at baseline was excluded as being nearly 30 times higher than any of the patient's follow-up NfL values. This patient did not show any clinical aspects that could explain such a high value as being 63 years old, having been diagnosed with diabetes and having suffered a clinical stroke 7 years ago.

Associations between NfL and DTI-MRI variables were tested using Spearman's rank correlations. Relationships between NfL or imaging measures and cognition were tested using linear regression after controlling for the clinical markers age, gender and premorbid IQ (NART). Employing a poisson regression model the association between NfL and disability was tested while controlling for baseline age.

The effect of time on change in cognition for Global and PS over the 5 follow-up years was estimated using a linear mixed model (lme4)²¹¹. Fixed effect variation was explained by time, and random effect variation allowed for remaining inter-individual differences. The intercept and slope of each patient's linear trajectory were allowed to vary with both fixed and random effects. The average fixed effects slopes of time are the average annualized change rate for a given cognition. Using maximum likelihood estimation to fit the model, the missing slope for one participant was estimated by taking information from all other patients included in the model.

To determine the predictive association between baseline NfL or baseline DTI and cognitive decline, a linear regression model was employed. Baseline cognition together with the clinical markers age, gender and NART were added as confounders. Employing a logistic regression, it was furthermore tested whether baseline NfL predicted an increase vs. no increase in mRS as a binary outcome over 5 years while controlling for age and its baseline disability score.

A Cox regression model was used to test whether the baseline NfL or imaging marker predicted dementia conversion. The proportional hazards assumption was met on the basis of Schoenfeld residuals. To assess how well the predictive models' discriminatory ability classifies patients converting to dementia and those with no dementia at varying threshold, ROC curves were computed and the area under the curve (AUC) was estimated²⁵³.

The effect of time on change in the imaging marker and in NfL over the 3 follow-up years was estimated using a linear mixed model as previously described (Ime4) ²¹¹. Using maximum likelihood estimation to fit the model, the missing DTI slope for two participants was estimated by taking information from all other patients included in the model. For the dementia analysis, observations after the dementia diagnosis were omitted.

To determine the predictive association between the change in the markers and cognitive decline or dementia conversion, a linear regression or a Cox regression model adjusted by age, gender and premorbid IQ was employed. The proportional hazards assumption was met on the basis of Schoenfeld residuals and the AUC was estimated ²⁵³.

Finally it was also tested whether baseline and change in NfL predicted change in lacunes and CMB employing a logistic regression model controlling for the imaging marker's baseline measure and baseline age. The AUC was computed to assess diagnostic discriminatory ability in correctly classifying the binary outcome measures ²⁵³. The association between baseline and change in NfL and change in NBV and in WMH was tested using a linear regression model while controlling for the imaging marker's intercept value coming from the linear mixed model and baseline age.

3. Results

3.1 Cross-sectional analysis

Study began in December 2007 and was completed in October 2013 with a consecutive case series design. Recruitment started in December 2007 and finished in August 2010. NfL data was available for 113 subjects. Clinical characteristics, cognitive measures and SVD markers at baseline are shown in Table 1. NfL levels inversely associated with global cognitive function, the executive function and processing speed subdomain scores (Table 2).

Higher NfL levels were positively associated with lacune count, CMB count, WMH, and with DTI measures (higher PSMD and lower MDPH), and negatively with NBV (Figure 2). We also determined whether association with cognition persisted after controlling for DTI (MDPH) and for the clinical markers; the association remained significant for PS but not for Global or EF (Table 2). Serum NfL also negatively correlated with disability (Table 2, Figure 3).

3.2 Longitudinal analysis

3.2.1 Prediction by baseline NfL

There was a significant decline in both Global and Processing speed over the 5 years follow-up (Table 3). Higher NfL at baseline predicted lower function in Global ($\beta= -0.335$ ($SE= 0.094$), $p=0.001$) independently of the clinical markers and baseline cognition (Table 4). Higher NfL also predicted lower function in Global ($\beta= -0.303$ ($SE= 0.095$), $p= 0.002$), when controlling DTI to the existing regression model. In contrast, baseline NfL did not predict increases in mRS score ($\beta= 0.407$ ($SE= 0.864$), $p= 0.638$).

Table 1. Baseline characteristics referring to clinical, cognitive and SVD markers

Baseline characteristics (N= 113)	
Clinical Characteristics	
Age, <i>Mean (SD)</i> years	70.012 (9.911)
Sex, (% <i>male</i>)	74 (0.66)
NART, <i>Mean (SD)</i>	99.301 (15.544)
mRS score, <i>Median (range)</i>	1 (0- 4)
Cognition	
Global, <i>Mean (SD)</i>	-0.654 (0.850)
EF, <i>Mean (SD)</i>	-0.869 (1.092)
PS, <i>Mean (SD)</i>	-0.987 (0.910)
SVD Marker	
NfL (pg/mL), <i>Mean (range)</i>	36.505 (5.7 – 708.9)
MDPH (mm ² /s), <i>Mean (SD)</i>	0.015 (0.003)
PSMD (mm ² /s), <i>Mean (SD)</i>	3.864e-04 (1.113e-04)
NBV (ml), <i>Mean (SD)</i>	1292.199 (87.946)
WMH (% brain), <i>Mean (range)</i>	3.507 (0.29- 12.81)
Lacune (number), <i>Median (range)</i>	2 (0- 27)
CMB (numbr), <i>Median (range)</i>	0 (0- 144)

NART- Premorbid IQ, mRS score- Modified Rankin Scale, Global- Global Cognition, EF-Executive Function, and PS–Processing Speed, , NfL- Serum Neurofilament Light Chain, MDPH- Mean Diffusivity Normalised Peak Height, PSMD- Peak Width of Skeletonised Mean Diffusivity, NBV- Normalised Brain Volume, WMH- White Matter Hyperintensity Lesion Load, Lacune- Lacune count, CMB- Cerebral Microbleeds count

Table 2. Cross-sectional regression between DTI and/ or NfL and cognition or disability.

	Regression on cognition with single marker						Regression on cognition with multiple markers						Regression on disability	
	Global		EF		PS		Global		EF		PS		mRS score	
	β (SE)	Adj. R	β (SE)	Adj. R	β (SE)	Adj. R	β (SE)	Adj. R	β (SE)	Adj. R	β (SE)	Adj. R	β (SE)	HL R ²
NfL (pg/mL) (log 10)	-0.201 (0.072), p= 0.006	0.459	-0.155 (0.074), p= 0.038	0.429	-0.244 (0.082), p= 0.004	0.288	-0.118 (0.075), p= 0.119	0.510	-0.086 (0.080), p= 0.284	0.450	-0.187 (0.086), p= 0.031	0.366	0.279 (0.079), p= 0.0004	0.111
MDPH (mm ² /s)	0.290 (0.071), p= 8.42e-05	0.503	0.221 (0.075), p= 0.004	0.449	0.330 (0.082), p= 0.0001	0.342	0.255 (0.074), p= 0.001		0.1951 (0.079), p= 0.015		0.273 (0.084), p= 0.002			
PSMD (mm ² /s)	-0.277 (0.067), p= 7.23e-05	0.506	-0.227 (0.070), p= 0.002	0.462	-0.322 (0.078), p= 6.70e-05	0.344								

Values show standardized regression coefficients β and standard errors (SE) for predictor variables in regression models of: Global- Global Cognition, EF-Executive Function, and PS-Processing Speed, mRS score- Modified Rankin Scale, NfL- Serum Neurofilament Light Chain, MDPH- Mean Diffusivity Normalised Peak Height, PSMD- Peak Width of Skeletonised Mean Diffusivity, Adj R²- Adjusted explained variance. HL R² - Hosmer and Lemeshow's R². Significant at p <0.05

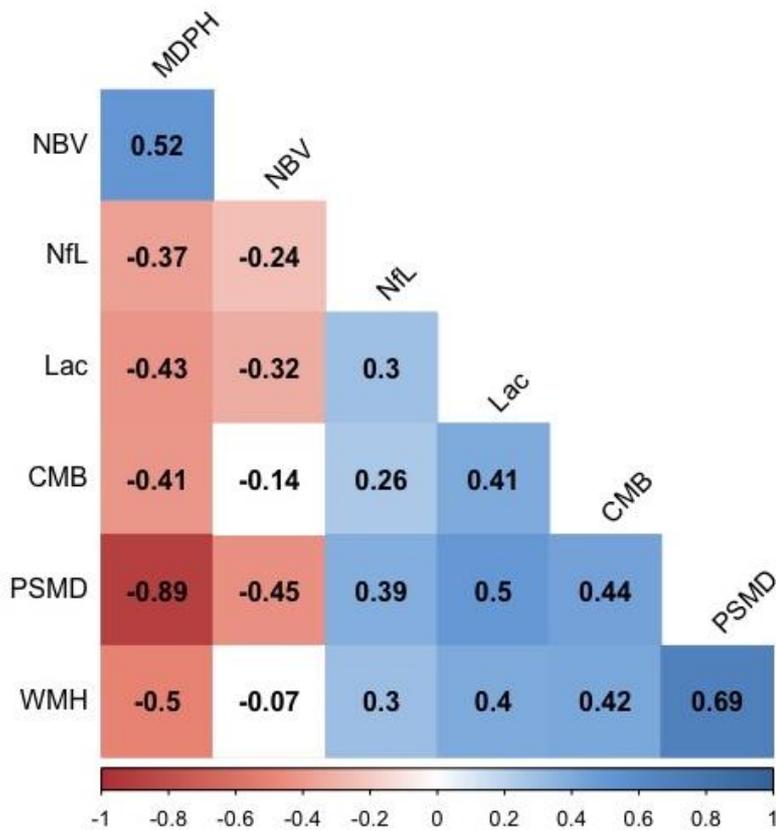


Figure 2. Higher levels of neurofilament light chain (NfL) was related to a higher SVD marker burden on imaging

Spearman's rank correlation shows the cross-sectional associations between MRI markers and neurofilament light chain (NfL). Blue and red quartiles show significant positive and negative correlations. White tiles show no significant association. NfL= neurofilament light chain, NBV= brain volume, Lac= lacune count, WMH= white matter hyperintensity, CMB= cerebral microbleed count, MDPH= mean diffusivity normalised peak height, PSMD= peak width of skeletonized mean diffusivity

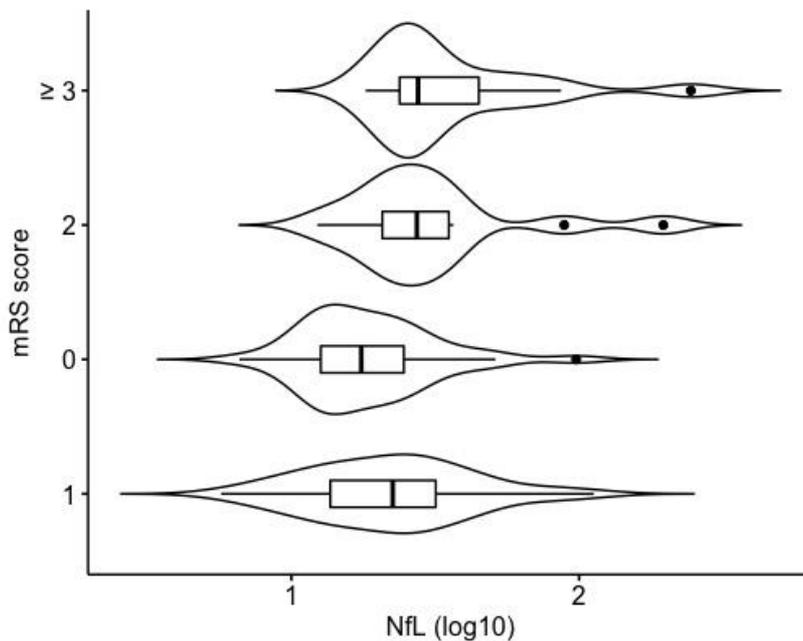


Figure 3. Higher levels of neurofilament light chain (NfL) was associated with greater disability

mRS score= modified Rankin score, NfL(log10)= neurofilament light chain (NfL) log10 transformed

107 patients had complete baseline DTI and NfL data of whom 19 converted to dementia during follow-up. Higher baseline NfL predicted the risk of converting to dementia while accounting for the clinical markers ($HR= 1.676$ (95% $CI= 1.183-2.373$), $p=0.004$) (Table 4). MRI parameters also predicted both cognitive decline and dementia (Table 4). The Area under the curve (AUC) for prediction of dementia was 0.775 for NfL, 0.791 for MDPH and 0.758 for PSMD. To determine whether NfL contributed additional information above that provided by DTI, we additionally included baseline MDPH in the model resulting an increase in AUC of 0.804. Baseline NfL predicted changes during follow-up in the number of lacunes and CMB, and also NBV independent of their initial MRI baseline values (Table 5). The model's AUC was 0.773 and 0.822 for lacunes and CMB respectively.

3.2.2 Prediction by change in NfL

There was no significant overall change in NfL over the 3 years follow-up period (Table 3). In view of this lack of change in NfL, there was no association with dementia conversion ($N= 15$) or change in any conventional MRI measure (Table 5, Table 6). In contrast there was a significant change in DTI over time (both assessed by MDPH and PSMD) (Table 3) that predicted dementia conversion (Table 6).

Table 3. Annualized cognitive and imaging change rate and over 5 years and 3 years respectively.

SVD markers				
Cognition (N= 90)	Estimated mean baseline value (CI)	Estimated mean annual change (CI)	Wald test	p Value
Global	-0.486 (-0.653, -0.319)	-0.026 (-0.044, -0.008)	-2.83	0.005
PS	-0.826 (-0.998, -0.655)	-0.059 (-0.088, -0.029)	-3.93	<0.001
Imaging characteristics (N= 90)	Estimated mean baseline value (CI)	Estimated mean annual change rate (CI)	Wald test	p Value
NfL (pg/mL) (log 10)	1.330 (1.279, 1.381)	0.0011 (-0.017, 0.019)	0.12	0.903
MDPH (mm ² /s)	0.015 (0.002, 0.016)	-0.0004 (-0.0005, -0.0003)	-10.34	< 0.001
PSMD (mm ² /s)	3.744e-04 (3.544e-04, 3.943e04)	1.316e-05 (9.467e-06, 1.686e-05)	6.981	< 0.001
WMH (% brain) (log 10)	0.438 (0.370, 0.506)	0.084 (0.075, 0.093)	17.95	< 0.001
NBV (ml)	1294.21 (1277.007, 1311.410)	-9.017 (-10.764, -7.270)	-10.12	< 0.001
	Median baseline value (range)	No of patients with incident findings		
Lacune	2 (0- 26)	24		
CMB	0 (0- 41)	32		

Yearly rates of change are defined as the mean estimates of the fixed effects from the linear mixed effect models with their 95% confidence interval (CI).

SVD- Small Vessel Disease, Global- Global Cognition, PS–Processing Speed, NfL- Serum Neurofilament Light Chain, MDPH- Mean Diffusivity Normalised Peak Height, PSMD- Peak Width of Skeletonised Mean Diffusivity, WMH- White Matter Hyperintensity Lesion Load, NBV- Normalised Brain Volume, Lacune- Lacune count, CMB- Cerebral Microbleeds count. Significant at P-value <0.05

Table 4. Longitudinal analysis. DTI and NfL baseline markers predicting cognitive decline and dementia conversion.

	Baseline Marker Prediction									
	Dementia Conversion (N= 19)				Decline in Global			Decline in PS		
	β (SE)	P-value	HR (95% CI)	AUC	β (SE)	P-value	Adj. R	β (SE)	P-value	Adj. R
NfL (pg/mL) (log 10)	0.516 (0.178)	0.004	1.676 (1.183, 2.373)	0.775	-0.335 (0.094)	0.001	0.353	-0.277 (0.114)	0.017	0.050
MDPH (mm ² /s)	-0.848 (0.265)	0.001	0.428 (0.255, 0.719)	0.791	0.148 (0.106)	0.169	0.265	0.139 (0.123)	0.262	-0.006
PSMD (mm ² /s)	0.507 (0.155)	0.001	1.661 (1.225, 2.252)	0.758	-0.215 (0.098)	0.032	0.290	-0.143 (0.116)	0.224	-0.003
NfL (pg/mL) (log10) indep. MDPH (mm ² /s)	0.369 (0.179)	0.040	1.446 (1.017, 2.055)	0.804						
NfL (pg/mL) (log10) indep. PSMD (mm ² /s)					-0.303 (0.095)	0.002	0.365			

Values show standardized regression coefficients: β and standard errors (SE) for predictor variables in regression models of dementia conversion and decline in cognition. NfL- Serum Neurofilament Light Chain, MDPH- Mean Diffusivity Normalised Peak Height, PSMD- Peak Width of Skeletonised Mean Diffusivity, Global- Global Cognition, PS-Processing Speed, HR- Hazard ratio, CI- Confidence Interval, Adj R²- Adjusted explained variance, AUC- Area under the Curve. Significant at P-value <0.05

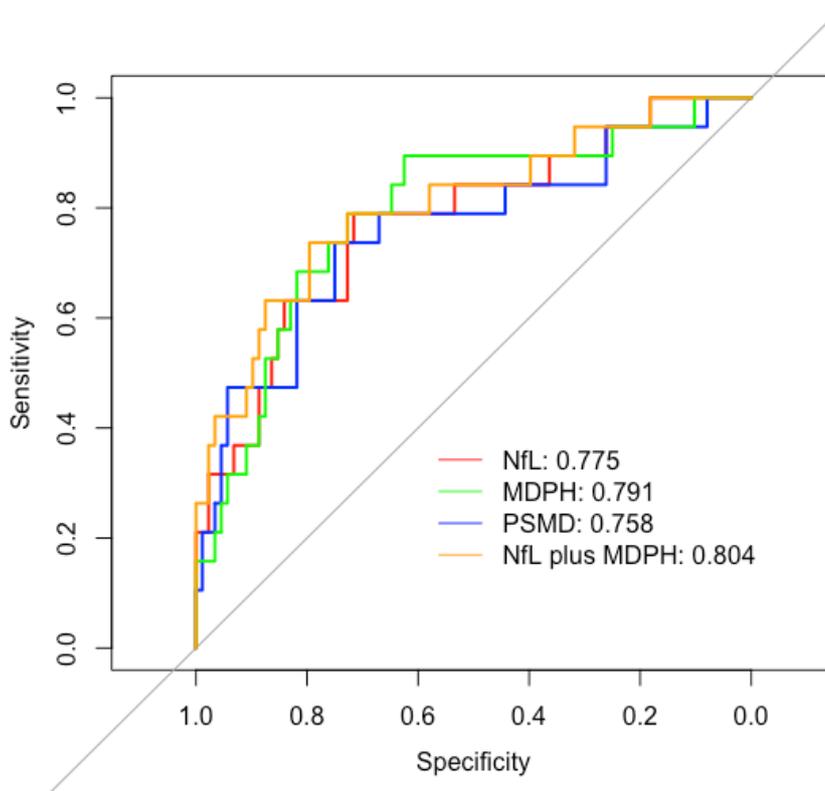


Figure 3. Baseline neurofilament light chain (NfL) and/ or DTI measures together with the clinical markers classified well dementia conversion and no-dementia conversion

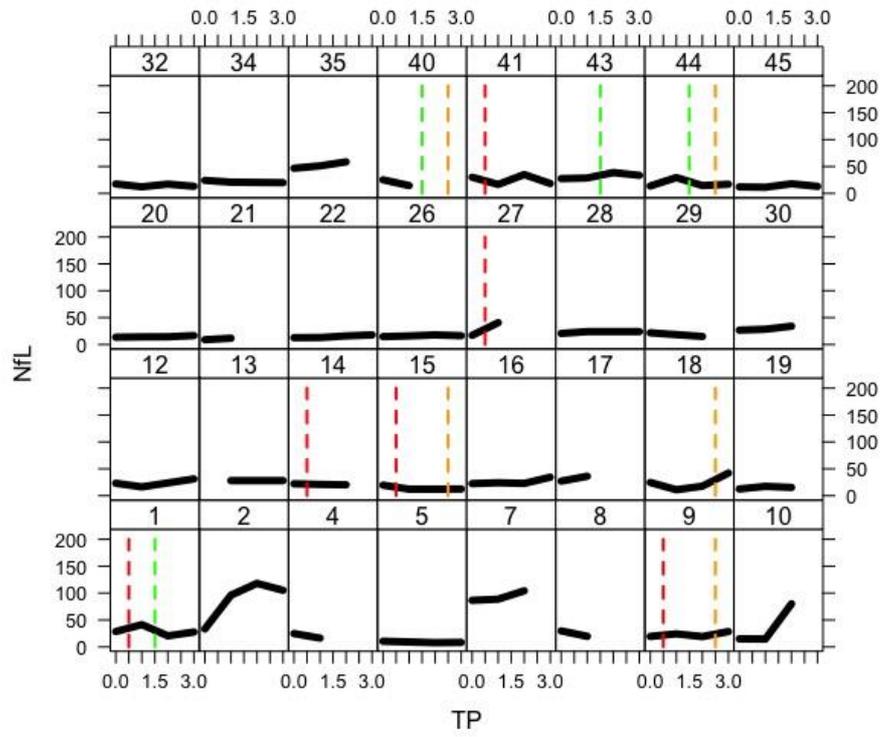
For each predictive model the area under the curve (AUC) shows the diagnostic discriminatory ability of a binary classifier system with varying thresholds. NfL= neurofilament light chain, MDPH= mean diffusivity normalised peak height, PSMD= peak width of skeletonized mean diffusivity

Table 5. Longitudinal analysis. Change in DTI but not in NfL predicts cognitive decline and dementia conversion.

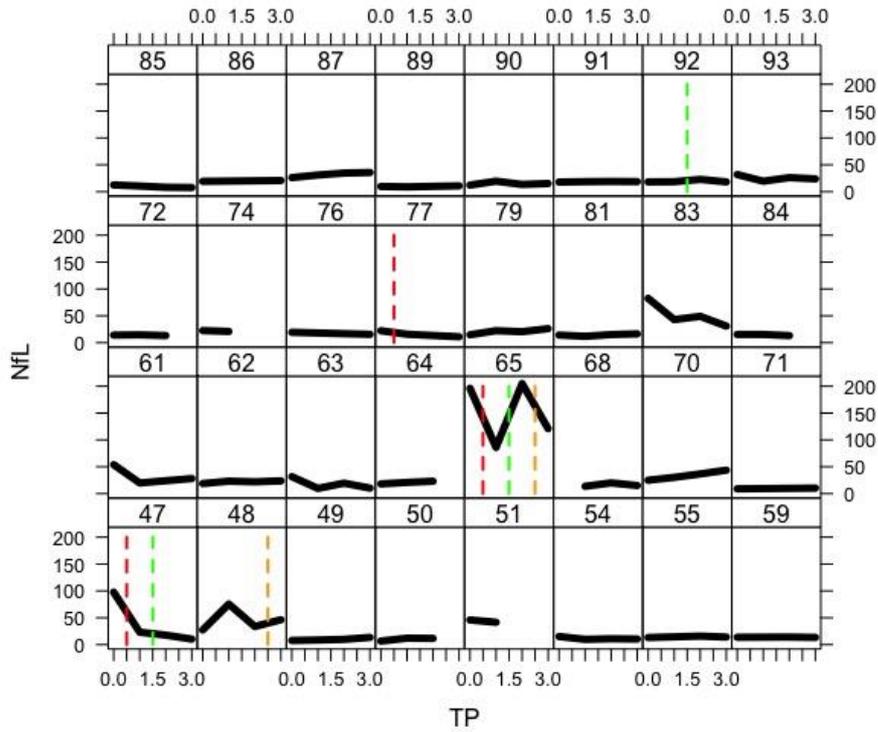
Cohort	Change Marker Prediction									
	Dementia Conversion (N= 15)				Decline in Global			Decline in PS		
	β (SE)	P-value	HR (95% CI)	AUC	β (SE)	P-value	Adj. R	β (SE)	P-value	Adj. R
NfL (pg/mL) (log 10)	0.178 (0.301)	0.554	1.195 (0.663, 2.155)	0.705	-0.051 (0.104)	0.625	0.070	0.035 (0.108)	0.751	-0.009
MDPH (mm ² /s)	-0.813 (0.297)	0.006	0.444 (0.248, 0.794)	0.788	0.260 (0.102)	0.012	0.134	0.049 (0.110)	0.655	-0.008
PSMD (mm ² /s)	0.594 (0.293)	0.043	1.812 (1.020, 3.219)	0.761	-0.362 (0.100)	0.001	0.199	-0.230 (0.108)	0.036	0.040

Values show standardized regression coefficients: β and standard errors (SE) for predictor variables in regression models of dementia conversion and decline in cognition. NfL- Serum Neurofilament Light Chain, MDPH- Mean Diffusivity Normalised Peak Height, PSMD- Peak Width of Skeletonised Mean Diffusivity, Global- Global Cognition, PS-Processing Speed, HR- Hazard ratio, CI- Confidence Interval, Adj R²- Adjusted explained variance, AUC- Area under the Curve. Significant at P-value <0.05

A)



B)



C)

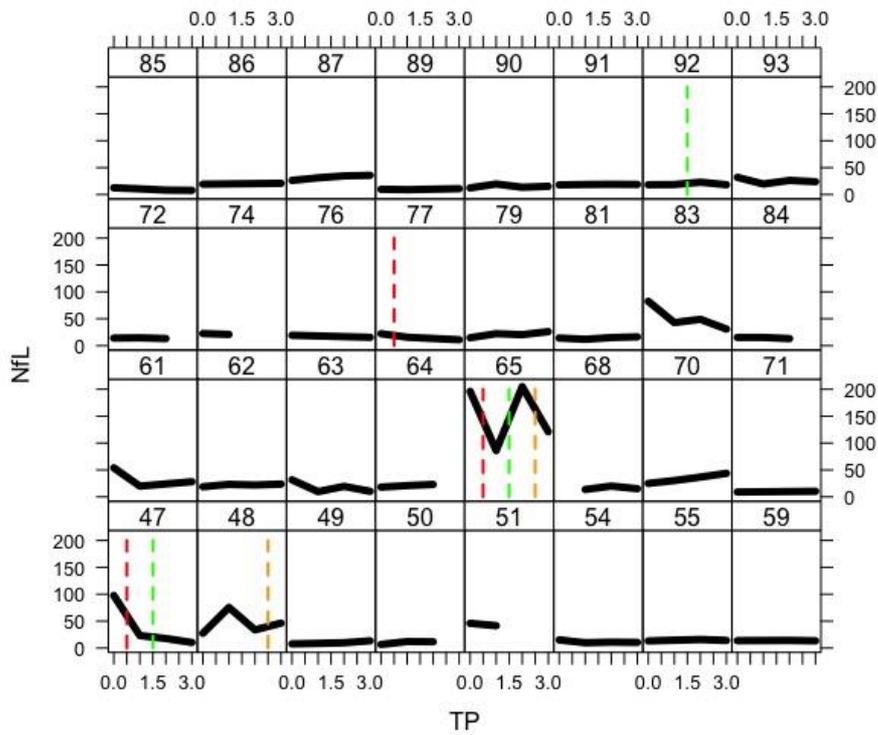
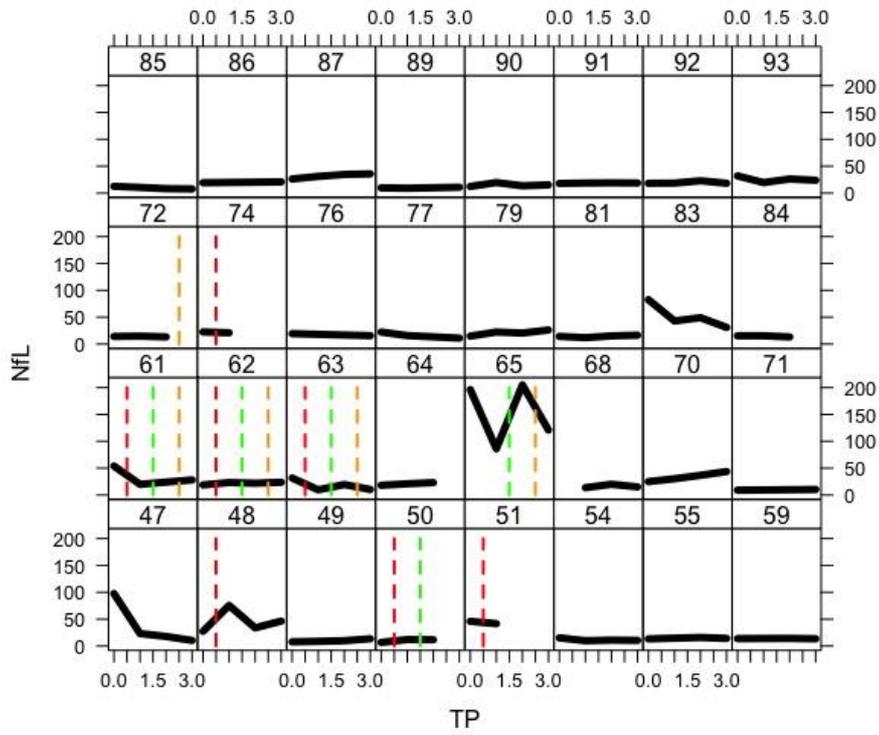


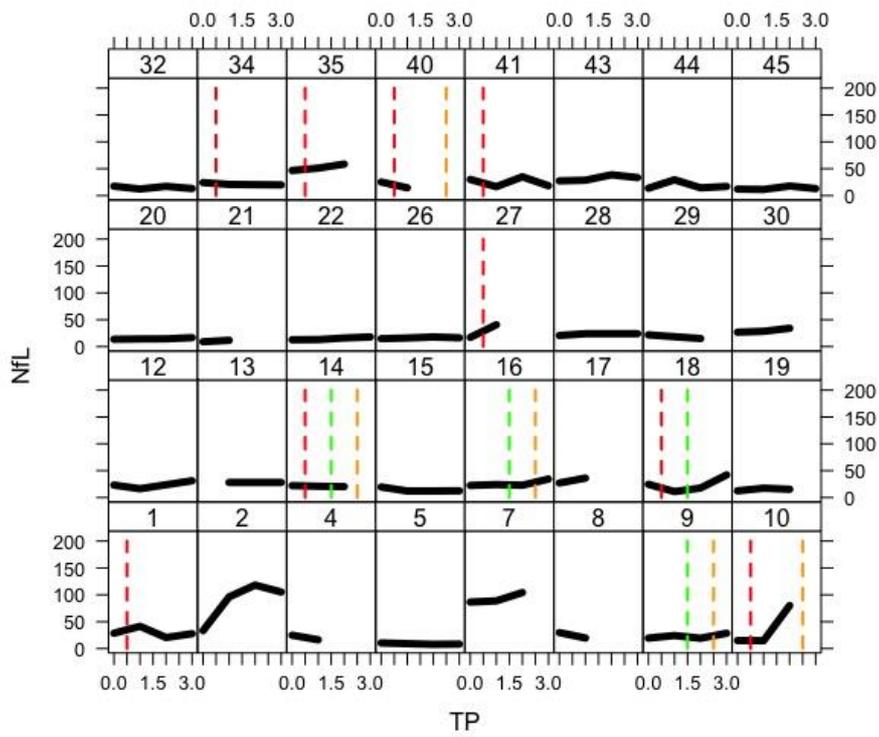
Fig 5A-C. Neurofilament light chain (NfL) measures and incidences of new lacune over time per patient

The black line refers to neurofilament light chain (NfL) levels over time. The red, green and orange dotted lines represents an incidence of a new lacune in a patient between baseline and time point 1, time point 1 and 2, time point 2 and 3 respectively. NfL= neurofilament light chain

A)



B)



C)

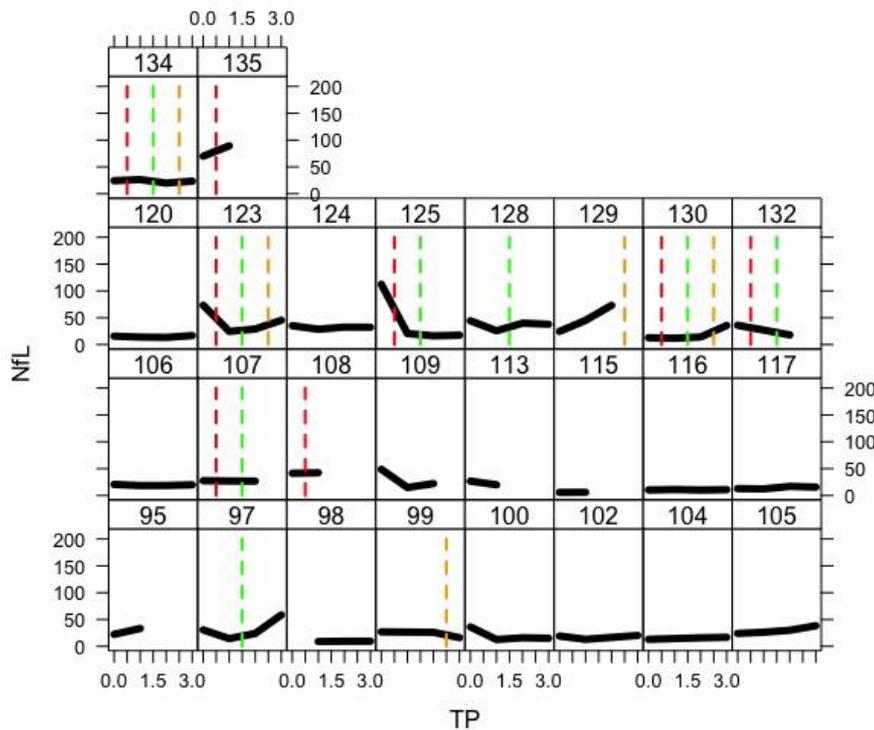


Fig 6A-C. Neurofilament light chain (NfL) measures and incidences of new cerebral microbleeds over time per patient

The black line refers to neurofilament light chain (NfL) levels over time. The red, green and orange dotted lines represents an incidence of a new cerebral microbleed in a patient between baseline and time point 1, time point 1 and 2, time point 2 and 3 respectively.

4. Discussion

The results showed that baseline NfL was associated with impaired cognitive function, disability and future dementia conversion. On the other hand, change in NfL over a 3 years period was not detectable during the study and was not associated with cognitive decline and dementia conversion. These findings indicate that while NfL may be a useful predictor of outcome, it is not likely to be a suitable surrogate marker for a clinical trial in SVD. The lack of a detectable change in NfL may be partly explained by previously reported dynamic changes in NfL post lacunar stroke where NfL was elevated but eventually returned to normal NfL levels after 15 months³⁰⁶. In contrast to NfL, diffusion tensor imaging (DTI) measures showed

significant change and predicted the clinical outcome measure supporting the role of DTI as a surrogate marker in SVD.

On the other hand, baseline NfL predicted dementia independently of age in SVD. This finding is not in line with previous evidence in SVD where baseline NfL predicted dementia only when not accounting for age³⁰⁹. The inconsistency may be explained by the different stage of SVD between cohorts. As shown in Chapter 3, SVD markers show different predictive powers over various SVD cohorts with imaging measures being stronger and clinical markers such as age being less important in more severe SVD compared to milder SVD. This explanation is supported by previous findings in CADASIL where baseline NfL predicted survival also independently of age³⁰⁸.

Although NfL may not have specific features required for a surrogate marker, NfL may still be an important marker for selecting an SVD group with an increased risk of progressing to dementia in a clinical trial. As the blood-based NfL marker can be easily acquired across large number of patients and as NfL was associated with dementia risk, the marker may be used in selecting a group more prone to dementia conversion in SVD. This may potentially increase the statistical power and may require a lower sample size for a clinical trial.

This chapter's cross-sectional findings further replicate previous evidence¹⁹³ that NfL was significantly associated with cognitive function and characteristic imaging markers in sporadic SVD. We further show that NfL correlates with disability not only in CADASIL but also in sporadic SVD.

This study has a number of strengths. Repeated sampling of both MRI and blood was performed and data on dementia conversion at 5 years was available for all subjects. MRI included both conventional markers and DTI, allowing the comparative performance of prediction by NfL levels with those of MRI markers to be determined. It also has limitations. NfL levels at more than one time point were not available on all subjects. It has been shown that dropouts had worse cognitive function than those who remained in the study, which might result in an underestimation of the magnitude of cognitive, MRI, and NfL changes over time²⁰¹.

To conclude, NfL may be a useful marker for selecting a SVD patient group with a higher risk for dementia conversion potentially reducing the required sample size needed for a clinical trial. NfL may therefore be useful as a prognostic marker. However,

we did not find any evidence that NfL could be used as surrogate clinical endpoint in a clinical trial as there was no significant change over the 3 years.

Chapter 7 Discussion

The work presented in this thesis focused on developing and testing imaging and circulating markers in patients with SVD to determine the most suitable surrogate markers, which can be employed as a clinical endpoint for a phase-2 clinical trial.

1. Summary of the findings

Chapter 2 presented the first multicenter randomized clinical trial in SVD involving DTI as a primary surrogate marker. While there were no treatment effects between the intensive and standard BP group, monthly changes in BP were negatively associated with WM lesion load growth, which was a secondary imaging endpoint. In contrast, there was no association between monthly changes in BP and the primary surrogate marker DTI, which measured changes in the WM microstructure in and outside of the WM lesion region. Follow-up analysis then showed that there were differences in DTI measurement changes depending on the scanner type (Siemens vs. Phillips). While it remained undetermined whether the significant differences in DTI changes between scanner types is a systematic scanner problem for multicenter trials, it raised the question of whether conventional MRI markers may be more robust than DTI markers for multicenter studies and therefore more suitable as a surrogate marker in this trial design. To be used as a clinical endpoint in a trial, any changes in surrogate marker must be associated with clinical outcomes such as dementia conversion.

Chapter 3, therefore, tested the clinical relevance of conventional MRI (brain volume, WMH, lacune count, and CMB count) and DTI measures across multiple cohorts characterized by different designs and SVD progressions. The results showed that while the DTI measure MD Median at baseline was associated with impaired cognitive function and predicted dementia conversion across all SVD cohorts, change in DTI overtime only predicted dementia conversion in severe SVD, characterized by lacunar stroke and confluent WMH, but not in mild SVD or MCI patients where the underlying pathology is not primarily vascular. Conventional MRI measures such as brain volume and lacune count further increased prediction only in severe SVD and monogenic SVD but not in mild SVD or MCI. The importance of clinical markers, such as age, sex, and education, or pre-morbid IQ, overall

decreased as the disease severity increased. These findings indicated that using DTI together with conventional MRI markers as surrogate markers is likely to be more successful in more severe SVD cohorts where the vascular factor is the primary determinant for dementia conversion.

Chapter 4 addressed the question, how the DTI data should be analyzed to have an optimal surrogate marker for a future phase 2 clinical trial in SVD. Over the recent 5 years, automatic and semi-automatic markers have been developed as surrogate markers for SVD. The chapter's goal was to compare these to two more conventional all WM histogram measures and to determine the optimal surrogate marker for a future phase 2 clinical trial based on several criteria: 1) markers' baseline association with impaired cognitive function, 2) the markers' baseline prediction for dementia conversion, 3) the markers' change over time associated with dementia conversion, 4) the markers' minimum sample size required for a phase II clinical trial. The previous chapters pointed out two further aspects to consider. Chapter 2 demonstrated scanner-/site-dependent DTI differences may be an issue that should be avoided as it may compromise the statistical power of the trial. Chapter 3 showed that the DTI marker's performance, as a surrogate marker may be dependent on the inclusion criteria. Surrogate markers are likely to be more successful in more severe SVD. There were several important findings in Chapter 4. First, as shown in Chapter 3, while most baseline measures predicted dementia conversion, significant changes in all DTI measures were only associated with dementia conversion in severe SVD but not in mild SVD. Second, while MD Median and DSEG significantly changed over 2 years in the multicenter trial PRESERVE, PSMD and Geff did not. There were, however, no significant site differences for any DTI changes. Visual inspection showed that there were more outliers in PSMD, Geff, and MD Median than in the combined all WM histogram measure PC1. Third, the sample size for a hypothetical clinical trial in SVD significantly varied both across the different markers and the different cohorts included. The automatic and semi-automatic markers PSMD and DSEG required the lowest sample size in sporadic SVD, whereas DSEG and MD Median had the lowest estimate for CADASIL. The global network measure Geff required high sample sizes across all cohorts. The sample size was overall higher in the MCI group compared to the vascular cohorts. Based on these findings, we concluded that recently developed automatic and semi-automatic markers PSMD and DSEG may be useful surrogate endpoints especially in patients with severe SVD. In contrast to PSMD, DSEG showed significant change in the multicenter study indicating that it may be less susceptible to scanner differences and may be better

employed as a surrogate marker for a phase-2 multicenter clinical trial. A summary indicating the markers' performance in each of the cohort studies is shown in Table 1.

Chapter 5 focused on the reproducibility of the imaging markers by comparing them across multiple designs and cohorts. Reproducibility on the same scanner as well as between different scanners is central when it comes to determining the optimal surrogate marker. The marker should yield highly consistent reliable results to avoid limiting the statistical power to detect a treatment effect. Between-scanner findings overall demonstrated only a moderate reproducibility of the markers. Skeletonized markers PSMD and MSMD and the network measure Geff were particularly impacted by discrepancies in MD distributions that were likely an issue with the scanner itself. Whereas MSMD and MD Median showed the highest reproducibility both between the 2 sessions and the 2 visits in the younger healthy control cohort, DSEG, MD Median and Geff demonstrated high reproducibility both in SVD patients and age-matched control participants. The findings indicate that MD Median may overall be more reproducible than other markers such as PSMD. Furthermore, despite its automation, it is recommended that every step of the computational pipeline is checked in detail as both issues coming from the scanner itself as well as from image processing may occur.

Problems with marker reproducibility underline 2 important aspects. First, running the imaging pipeline of automatic markers without checking the processing steps in detail may significantly impair its reproducibility. As a consequence, any DTI measures may still be labor-intensive and require trained personnel. Second, reproducibility is compromised in multicenter trials where different scanner types are employed. Both of these limitations do not exist for blood-based circulating markers. In Chapter 6 it was determined whether the circulating blood-based marker NfL may be suitable as a surrogate marker for a phase-2 clinical trial. The results confirmed previous evidence that the circulating blood-based marker NfL at baseline was associated with impaired cognitive function and disability in SVD. While NfL at baseline predicted later dementia conversion and decline in cognition both independently of the clinical markers as well as of the clinical markers and DTI measure, there was no significant change over the 3 years, which did not predict dementia conversion. It could, however, be shown that baseline NfL also predicted changes in lacune count, CMB count, and brain atrophy. These findings indicated that NfL may not be a suitable surrogate marker for a phase-2 clinical trial in SVD.

Table 1. Summary table shows which imaging markers were effective in each of the cohorts included in the thesis. The table is color-coded indicating a statistically significant effect/ lower minimum sample size than the other measures (*in green*) vs. statistically non-significant effect/ higher minimum sample size than the other measures (*in red*). The table further indicates whether the imaging markers were suitable as surrogate markers (+++). For the surrogate marker evaluation only imaging measures were included which have been tested in the cross-sectional and longitudinal analysis as well as where a minimum sample size estimated were computed (*in blue*).

Cohort	Cohort characteristic	Marker	Performance (description)	Shown suitable as surrogate marker (+++)
PRESERVE (N= 111)	Severe SVD	MDPH	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional analysis significant change over 2 years showed no ITT and PP treatment effect change in BP was not associated with the markers change 	
		NBV	<ul style="list-style-type: none"> significant change over 2 years showed no ITT and PP treatment effect change in BP was not associated with the markers change 	
		WMH lesion load	<ul style="list-style-type: none"> significant change over 2 years showed no ITT and PP treatment effect 	

			<ul style="list-style-type: none"> • change in BP was associated with the marker's change 	
		Lacune count	<ul style="list-style-type: none"> • significant change over 2 years • showed no ITT and PP treatment effect • change in BP was not associated with the markers change 	
		CMB count	<ul style="list-style-type: none"> • significant change over 2 years • showed no ITT and PP treatment effect • change in BP was not associated with the markers change 	
		MD median	<ul style="list-style-type: none"> • significant association with impaired cognitive function in cross-sectional analysis • significant change over 2 years • no difference in the marker between scanner sites 	
		FA median	<ul style="list-style-type: none"> • significant association with impaired cognitive function in cross-sectional analysis • significant change over 2 years 	
		FAPH	<ul style="list-style-type: none"> • significant association with impaired cognitive function in cross-sectional analysis • significant change over 2 years 	
		FA pkval	<ul style="list-style-type: none"> • significant association with impaired cognitive function in cross-sectional analysis • significant change over 2 years 	

		MD pkval	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional analysis no significant change over 2 years 	
		PC1	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional analysis marker required a higher sample size than other markers 	
		PSMD	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional analysis no significant change over 2 years 	
		DSEG	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional analysis significant change over 2 years 	
		Geff	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional analysis no significant change over 2 years 	
SCANS (N= 121)	Severe SVD	MD median	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional analysis the marker contributes more in explaining the model's variance in impaired cognitive function than other imaging markers 	+++

			<ul style="list-style-type: none"> • baseline marker predicted dementia conversion • significant change over 3 years • marker's change over 3 years was associated with dementia conversion over 5 years • marker required a lower sample size than other markers for a 3-years randomized clinical trial 	
		MDPH	<ul style="list-style-type: none"> • significant association with impaired cognitive function in cross-sectional analysis • baseline marker predicted dementia conversion • significant change over 3 years • marker's change over 3 years was associated with dementia conversion over 5 years 	+++
		FA median	<ul style="list-style-type: none"> • significant association with impaired cognitive function in cross-sectional analysis • baseline marker predicted dementia conversion • significant change over 3 years • marker's change over 3 years was not associated with dementia conversion over 5 years 	
		FAPH	<ul style="list-style-type: none"> • significant association with impaired 	

			<p>cognitive function in cross-sectional analysis</p> <ul style="list-style-type: none"> • baseline marker predicted dementia conversion • no significant change over 3 years • marker's change over 3 years was associated with dementia conversion over 5 years 	
		FA pkval	<ul style="list-style-type: none"> • significant association with impaired cognitive function in cross-sectional analysis • baseline marker predicted dementia conversion • significant change over 3 years • marker's change over 3 years was associated with dementia conversion over 5 years 	
		MD pkval	<ul style="list-style-type: none"> • significant association with impaired cognitive function in cross-sectional analysis • baseline marker predicted dementia conversion • significant change over 3 years • marker's change over 3 years was not associated with dementia conversion over 5 years 	
		NBV	<ul style="list-style-type: none"> • the marker contributes more in explaining the model's variance in 	

			<ul style="list-style-type: none"> impaired cognitive function than other imaging markers significant change over 3 years 	
		WMH	<ul style="list-style-type: none"> the marker contributes less in explaining the model's variance in impaired cognitive function than other imaging markers significant change over 3 years 	
		Lacune count	<ul style="list-style-type: none"> the marker contributes more in explaining the model's variance in impaired cognitive function than other imaging markers 	
		CMB count	<ul style="list-style-type: none"> the marker contributes less in explaining the model's variance in impaired cognitive function than other imaging markers 	
		PC1	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional analysis baseline marker predicted dementia conversion marker consisting of DTI changes over 3 years was associated with dementia conversion over 5 years marker required a higher sample size than other markers for a 3-years randomized clinical trial 	
		PSMD	<ul style="list-style-type: none"> significant association with impaired 	+++

			<p>cognitive function in cross-sectional analysis</p> <ul style="list-style-type: none"> • baseline marker predicted dementia conversion • significant change over 3 years • marker's change over 3 years was associated with dementia conversion over 5 years • marker required a lower sample size than other markers for a 3-years randomized clinical trial 	
		DSEG	<ul style="list-style-type: none"> • significant association with impaired cognitive function in cross-sectional analysis • baseline marker predicted dementia conversion • significant change over 3 years marker's change over 3 years was associated with dementia conversion over 5 years • marker required a lower sample size than other markers for a 3-years randomized clinical trial 	+++
		Geff	<ul style="list-style-type: none"> • significant association with impaired cognitive function in cross-sectional analysis • baseline marker predicted dementia conversion • significant change over 3 years marker's 	

			<ul style="list-style-type: none"> • change over 3 years was associated with dementia conversion over 5 years • marker required a higher sample size than other markers for a 3-years randomized clinical trial 	
RUN DMC (N= 503)		MD median	<ul style="list-style-type: none"> • significant association with impaired cognitive function in cross-sectional analysis • baseline marker predicted dementia conversion • significant change over 4 years • marker's change over 4 years was not associated with dementia conversion • marker required a higher sample size than other markers 	
		MDPH	<ul style="list-style-type: none"> • significant association with impaired cognitive function in cross-sectional analysis • baseline marker predicted dementia conversion • no significant change over 4 years • marker's change over 4 years was not associated with dementia conversion 	
		FA median	<ul style="list-style-type: none"> • significant association with impaired cognitive function in cross-sectional analysis • baseline marker did not predict dementia conversion 	

			<ul style="list-style-type: none"> • no significant change over 4 years • marker's change over 4 years was not associated with dementia conversion 	
		FAPH	<ul style="list-style-type: none"> • significant association with impaired cognitive function in cross-sectional analysis • baseline marker did not predict dementia conversion 	
		FA pkval	<ul style="list-style-type: none"> • no significant association with impaired cognitive function in cross-sectional analysis • baseline marker did not predict dementia conversion • significant change over 4 years 	
		MD pkval	<ul style="list-style-type: none"> • significant association with impaired cognitive function in cross-sectional analysis • baseline marker did not predict dementia conversion • significant change over 4 years 	
		TBV	<ul style="list-style-type: none"> • the marker contributes more in explaining the model's variance in impaired cognitive function than other imaging markers • significant change over 4 years 	
		WMH	<ul style="list-style-type: none"> • the marker contributes less in explaining the model's variance in impaired cognitive function than other 	

			<ul style="list-style-type: none"> imaging markers significant change over 4 years 	
		Lacune count	<ul style="list-style-type: none"> the marker contributes less in explaining the model's variance in impaired cognitive function than other imaging markers 	
		CMB count	<ul style="list-style-type: none"> the marker contributes less in explaining the model's variance in impaired cognitive function than other imaging markers 	
		PC1	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional analysis baseline marker predicted dementia conversion marker consisting of DTI changes over 3 years was not associated with dementia conversion over 4 years marker required a higher sample size than other markers for a 4-years randomized clinical trial 	
		PSMD	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional analysis baseline marker predicted dementia conversion significant change over 4 years change over 4 years was not associated 	

			<ul style="list-style-type: none"> with dementia conversion marker required a lower sample size than other markers for a 4-years randomized clinical trial 	
		DSEG	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional analysis baseline marker predicted dementia conversion significant change over 4 years change over 4 years was not associated with dementia conversion marker required a lower sample size than other markers for a 4-years randomized clinical trial 	
		Geff	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional analysis baseline marker predicted dementia conversion significant change over 4 years change over 4 years was not associated with dementia conversion marker required a higher sample size than other markers for a 4-years randomized clinical trial 	
HARMONISATION (N= 127)	MCI	MD median	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional 	

			<ul style="list-style-type: none"> analysis baseline marker predicted dementia conversion significant change over 2 years marker's change over 2 years was associated with dementia conversion marker required a high sample size than other markers 	
		MDPH	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional analysis baseline marker predicted dementia conversion significant change over 2 years marker's change over 2 years was not associated with dementia conversion 	
		FA median	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional analysis baseline marker did not predict dementia conversion significant change over 2 years marker's change over 2 years was not associated with dementia conversion 	
		FAPH	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional analysis baseline marker did not predict 	

			<ul style="list-style-type: none"> dementia conversion significant change over 2 years marker's change over 2 years was not associated with dementia conversion 	
		FA pkval	<ul style="list-style-type: none"> no significant association with impaired cognitive function in cross-sectional analysis baseline marker did not predict dementia conversion no significant change over 2 years marker's change over 2 years was not associated with dementia conversion 	
		MD pkval	<ul style="list-style-type: none"> no significant association with impaired cognitive function in cross-sectional analysis baseline marker did not predict dementia conversion significant change over 2 years marker's change over 2 years was not associated with dementia conversion 	
		ICV	<ul style="list-style-type: none"> the marker contributes less in explaining the model's variance in impaired cognitive function than other imaging markers no significant change over 2 years 	
		WMH	<ul style="list-style-type: none"> the marker contributes less in explaining the model's variance in impaired cognitive function than other 	

			<ul style="list-style-type: none"> imaging markers significant change over 2 years 	
		Lacune count	<ul style="list-style-type: none"> the marker contributes less in explaining the model's variance in impaired cognitive function than other imaging markers 	
		CMB count	<ul style="list-style-type: none"> the marker contributes less in explaining the model's variance in impaired cognitive function than other imaging markers 	
		PC1	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional analysis baseline marker predicted dementia conversion marker consisting of DTI changes over 3 years was not associated with dementia conversion over 4 years 	
		PSMD	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional analysis baseline marker predicted dementia conversion significant change over time marker's change over 2 years was not associated with dementia conversion 	
		DSEG	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional 	

			<ul style="list-style-type: none"> analysis baseline marker predicted dementia conversion significant change over time marker's change over 2 years was not associated with dementia conversion 	
		Geff	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional analysis baseline marker did not predict dementia conversion significant change over time marker's change over 2 years was not associated with dementia conversion 	
ASPS-Fam		MD median	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional analysis 	
		MDPH	<ul style="list-style-type: none"> no association with impaired cognitive function in cross-sectional analysis 	
		FA median	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional analysis 	
		FAPH	<ul style="list-style-type: none"> no association with impaired cognitive function in cross-sectional analysis 	
		FA pkval	<ul style="list-style-type: none"> no association with impaired cognitive function in cross-sectional analysis 	
		MD pkval	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional 	

			analysis	
		NBV	<ul style="list-style-type: none"> the marker contributes less in explaining the model's variance in impaired cognitive function than other imaging markers 	
		WMH	<ul style="list-style-type: none"> the marker contributes more in explaining the model's variance in impaired cognitive function than other imaging markers 	
		Lacune count	<ul style="list-style-type: none"> the marker contributes less in explaining the model's variance in impaired cognitive function than other imaging markers 	
		PC1	<ul style="list-style-type: none"> no significant association with impaired cognitive function in cross-sectional analysis 	
		PSMD	<ul style="list-style-type: none"> no significant association with impaired cognitive function in cross-sectional analysis 	
		DSEG	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional analysis 	
		Geff	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional analysis 	
CADASIL		MD median	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional analysis 	+++

			<ul style="list-style-type: none"> significant change over 1.5 years marker required a lower sample size than other markers the marker contributes more in explaining the model's variance in impaired cognitive function than other imaging markers 	
		MDPH	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional analysis 	
		FA median	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional analysis 	
		FAPH	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional analysis 	
		FA pkval	<ul style="list-style-type: none"> no significant association with impaired cognitive function in cross-sectional analysis 	
		MD pkval	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional analysis 	
		NBV	<ul style="list-style-type: none"> the marker contributes more in explaining the model's variance in impaired cognitive function than other imaging markers 	
		WMH	<ul style="list-style-type: none"> the marker contributes less in explaining the model's variance in 	

			impaired cognitive function than other imaging markers	
		Lacune count	<ul style="list-style-type: none"> the marker contributes more in explaining the model's variance in impaired cognitive function than other imaging markers 	
		CMB count	<ul style="list-style-type: none"> the marker contributes less in explaining the model's variance in impaired cognitive function than other imaging markers 	
		PC1	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional analysis marker required a higher sample size than other markers 	
		PSMD	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional analysis significant change over 1.5 years marker required a lower sample size than other markers 	
		DSEG	<ul style="list-style-type: none"> significant association with impaired cognitive function in cross-sectional analysis significant change over 1.5 years marker required a lower sample size than other markers 	
		Geff	<ul style="list-style-type: none"> significant association with impaired 	

			<p>cognitive function in cross-sectional analysis</p> <ul style="list-style-type: none"> • no significant change over 1.5 years • marker required a higher sample size than other markers 	
--	--	--	---	--

MD Median= mean diffusivity median of the all WM histogram, MDPH= mean diffusivity normalised peak height of the histogram, MD pkval= mean diffusivity peak value of the histogram, FA Median= fractional anisotropy median of the all WM histogram, MDPH= fractional anisotropy normalised peak height of the histogram, MD pkval= fractional anisotropy peak value of the histogram, PC1= scores of the first principal component, PSMD= peak width of skeletonized mean diffusivity, DSEG= diffusion tensor image segmentation, Geff= global efficiency network measure, NBV= normalised brain value, ICV= intracranial volume, TBV= total brain volume, WMH= white matter hyperintensity, CMB= cerebral microbleeds, RCT= randomized controlled trial, ITT= intention-to-treat analysis, PP= per-protocol analysis

2. Five recommendations when planning future phase-2 clinical trials

There are five recommendations as a result of this work. They may be useful when planning a future phase-2 clinical trial in SVD.

2.1 Think carefully about single vs. multicenter trial designs

Employing imaging measures as a surrogate endpoint in a phase-2 clinical trial still requires more than 100 patients with sporadic SVD per treatment arm for a treatment effect size of 20%. A multicenter trial design may be the answer in achieving the targeted sample size but may also create further challenges which are not only due to a higher administrative workload caused by coordinating the respective sites. When it comes to imaging, a multicenter trial design involving different scanner types may influence the statistical power of the study. This impact may be stronger and differently pronounced for certain imaging measures than for others. Chapter 2 indicated that the change in DTI measure may depend on scanner vendors (Siemens vs. Phillips). Chapter 5 further showed that the reproducibility of all imaging measures tested is not high between scanner types. The aim is though not to discourage setting up a multicenter imaging trial in SVD (on the contrary, they are very much needed to test new clinical therapies in SVD) but to be aware of these challenges and to know how to minimize their influence. One way of mitigating it would be to run a pilot imaging study prior to the start of the multicenter clinical trial. In this pilot study, a small number of patients would undergo MRI with the various MRI machines prior to the trial's onset. This would allow determining how reproducible the imaging measures are across the different scanner types. As a consequence, it may be decided not to employ certain machines for the study or to increase the overall sample size to meet the statistical power. Generally, it is further recommended that only scanner sites with a sufficient number of patients per site are included. Having sites characterized by strongly differing sample sizes may complicate the statistical analysis later in the course of the study.

2.2 Think thoroughly about the specific SVD population to be included

SVD is a heterogeneous disease. This work demonstrated that also changes in the imaging markers are associated with later dementia conversion depending on the disease severity. Chapter 3 demonstrated that changes in DTI measures were only associated with dementia conversion in the severe SVD but not in the mild SVD cohort. Similarly, change in conventional MRI markers only added predictive value to

later dementia conversion in the severe SVD but not in the mild SVD cohort. Brain volume and lacune count in combination with the clinical and other imaging measures were important predictors only in severe SVD and monogenic SVD but not in mild SVD, MCI, or individuals with normal neurological functioning. Chapter 4 further showed that differences in the inclusion criteria may also require a larger minimum sample size for a clinical trial. More studies are though needed which test these differences in other cohorts as there were significant methodological differences between the cohorts tested. For now, it is important to point out that these differences in the imaging markers' performance as surrogate markers may exist depending on the SVD severity. It is therefore recommended to think thoroughly about the specific SVD population to be included in the clinical trial and to review the gathered evidence about how certain imaging markers perform as surrogate markers in different SVD populations. The Fazekas scale for classifying WM lesions ⁷³, the presence of a lacunar stroke or the blood-based marker NfL may help to narrow down the SVD subpopulation of interest.

2.3 Think carefully about the cognitive measures to be employed

Patients with SVD often show cognitive impairment particularly in processing speed and executive function. Longitudinally cognitive decline is however far from being homogeneous as some SVD patients show sharp cognitive declines over a few years while others demonstrate little change in cognitive function. In line with previous evidence, there was no cognitive decline over the 2 years in the PRESERVE clinical trial. It can therefore be concluded that currently available measures of cognition should not be employed as a primary measure in a 2-3 years phase-2 clinical trial in SVD unless a significantly larger sample size is considered. On the other hand, cognitive measures in a clinical trial over time give valuable clinical information about the individual patient. In large sample sizes, it allows explorative analysis to determine why some patients show significant deterioration or convert to dementia while others do not. Recent evidence has shown that some cognitive measures are more sensitive in measuring impairment in cognition than others ³¹³. It is therefore advised to read carefully the literature before deciding which cognitive measure to employ. Future studies should also examine whether paper-based cognitive tests are really necessary or if some tests show high validity and reliability also in a digital format. This may significantly reduce the amount of time to evaluate the tests afterward particularly in a clinical trial with a large sample size and may also minimize any human-made error rate.

2.4 Think thoroughly about possible reasons for a drop in sample size

As presented in Chapter 2, the clinical trial suffered from a significant drop in sample size over the 2 years. From the 111 patients enrolled at the start of the trial, only 81 patients had useable DTI measures at both time points. This was not only due to human-related factors such as withdrawal of consent or death but also due to a rigorous detailed evaluation of the DTI image quality. It is therefore important to be aware of both factors when planning the sample size in an imaging clinical trial. One possible way of being aware of image quality problems already in the course of the clinical trial is to have concurrent blinded image quality control. This may help to identify systematic image quality problems earlier and may allow for early technical interventions. Apart from a concurrent image quality evaluation, it is also essential to think about how to avoid losing too much sample size due to modifiable human causes such as withdrawal of consent or missing the appointment. This is a particularly important issue to think about in a population characterized by older age, a higher prevalence of multiple diseases, cognitive impairment, and disability. It is advised to review in detail previous studies or clinical trials with a similar aged population to come up with a comprehensive strategy to attenuate a significant drop in sample size throughout the clinical trial. One way of achieving this would be to reduce the number of in-person visits to the clinical site, which may require a long travel time and may be stressful, and to have more remote assessments in the patient's home.

2.5 Check the computational steps of the imaging marker despite its automation

In recent years great progress has been made in developing automatic or semi-automatic imaging measures in SVD cohorts. Chapter 4 demonstrated that DTI measures such as PSMD or DSEG may be suitable as surrogate markers particularly in severe small vessel disease marked by lacunar stroke. Using these measures in phase-2 clinical trials could help to reduce the overall amount of time for the MRI analysis in a clinical trial. Chapter 5, however, also demonstrated that despite its automatic marker computation, detailed quality control concerning the segmentation and registration may still be very much needed. Given these pieces of evidence in the thesis, it is therefore advised to employ the automatic measure as a surrogate endpoint in a clinical trial while making sure that every computational step of the image measure's pipeline is still checked by an MRI expert.

3. How would a future phase-2 clinical trial in SVD look like?

3.1 Patient selection

Since NfL as a blood marker can be more easily acquired and be centrally computed, it would help to screen for SVD patients in a larger group of the elderly population who are at a higher risk for dementia conversion. As NfL is a non-specific SVD marker, NfL together with imaging such as an MRI score system¹⁶⁶ could additionally be used in a second stage to further narrow down the patient group of the trial. On the other hand, although selecting patients with more severe SVD compared to milder disease stage may increase the statistical power of the study, it may likely decrease the responsiveness of the treatment intervention due to the already caused significant brain damage.

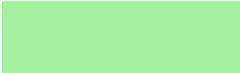
3.2 Selecting the surrogate marker

The automatic and semi-automatic markers PSMD and DSEG showed similar associations with the clinical outcome as MD Median and required the lowest sample size estimates in the single-center sporadic SVD cohorts making them eligible surrogate markers⁶⁶. Conventional MRI measures may be considered as a secondary endpoint in severe sporadic or monogenic SVD as they account for a significant variation in impaired cognitive function and showed an added predictive value for dementia conversion. A color-coded summary evaluating the performance of the various DTI measures as surrogate markers in severe sporadic and monogenic SVD can be found in Table 2. For multicenter trials, DSEG may be preferred over PSMD as it showed to be more robust in the multicenter trial and had overall a higher reproducibility between scanners. Within-center reproducibility was high for MD Median making it a good choice as a surrogate marker for a single-center clinical trial. The fact that MD Median showed significant change over the 2-years multicenter PRESERVE trial may indicate that it is also robust for a multicenter trial. Selecting the right imaging marker as a surrogate endpoint also depends on the type of treatment administered in the clinical trial. The specific treatment may be associated with an expected treatment effect size, which in turn may impact the duration of the study as well as the overall minimum sample size required. A summary of the aspects to consider when planning a phase-2 clinical trial in severe SVD can be found in Figure 1.

Table 2. Overall evaluation of the imaging measures for a phase-2 clinical trial in severe and monogenic SVD

Marker	Cross-sectional association with cognition	Baseline marker predicts later dementia conversion*	Significant change over time	Change in the marker associated with dementia conversion*	Sample size requirement	Multicenter
MD median						
PC1						
PSMD						
DSEG						
Geff						

PC1= scores of the first principal component, PSMD= peak width of skeletonized mean diffusivity, DSEG= diffusion tensor image segmentation, Geff= global efficiency network measure, MD median= mean diffusivity median of the all WM histogram

 = consistently good performance across SCANS, PRESERVE and CADASIL

 = less consistently good performance across SCANS, PRESERVE and CADASIL

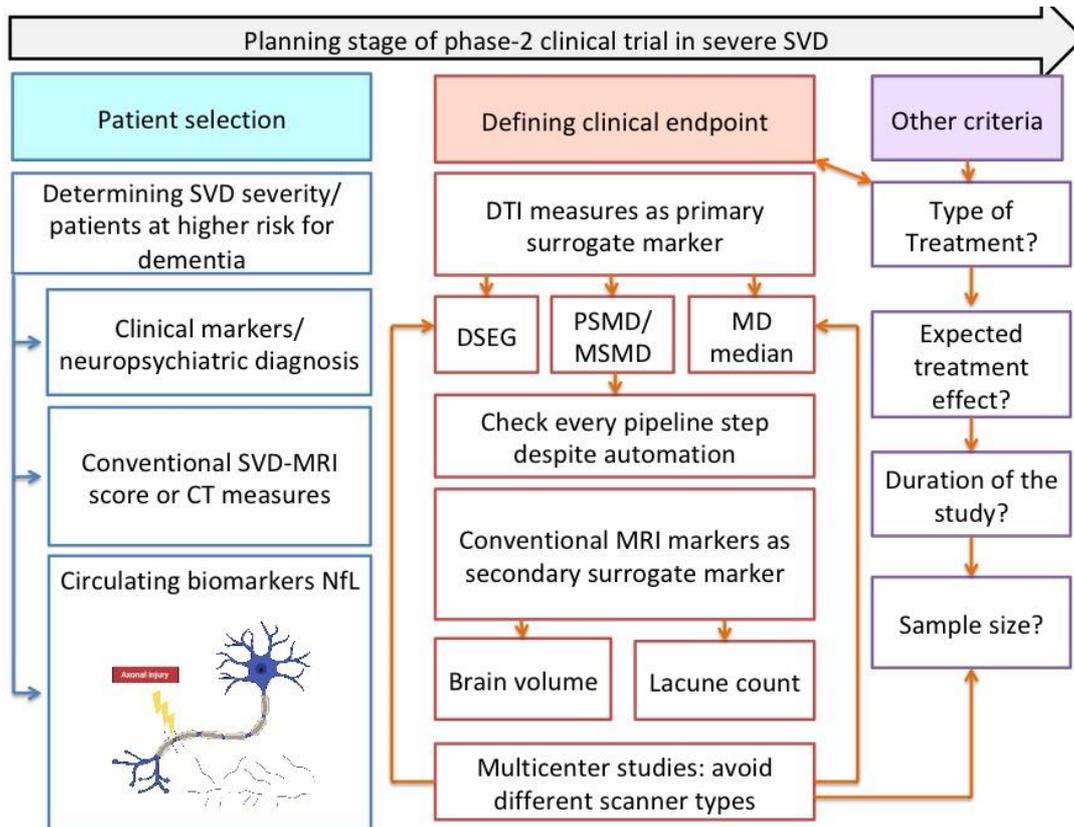


Figure 1. Aspects to consider when planning a phase-2 clinical trial in severe SVD. Selecting patients with severe SVD and with a higher risk for dementia based on a combination of clinical markers, imaging scores and circulating biomarkers such as NfL may increase the statistical power for a clinical trial. DTI markers may be suitable as surrogate endpoints but may be impacted by the study’s design (single vs. multicenter) and by the type of treatment to be administered. Conventional MRI markers such as change in brain volume and lacune count incidence may increase the accuracy in predicting clinical outcome measures such as dementia conversion beyond DTI. It is advised to check in detail every computation step in PSMD despite its automatic property.

4. Conclusion

This work evaluated MRI and serum markers as clinical endpoints for future phase-2 clinical trials in SVD. Over the recent years much progress has been made in developing new surrogate markers. This work compared their performance and outlined how a possible future phase-2 clinical trial may look like. The results showed that aspects of single- vs. multicenter designs, and selecting the right SVD target population are central cornerstones and may impact the marker's performance in the clinical trial. My work has highlighted a number of considerations that need to be considered carefully when planning a phase-2 clinical trial, and I hope that by using them the success of such trials can be improved.

Reference

1. Pantoni L. Cerebral small vessel disease: from pathogenesis and clinical characteristics to therapeutic challenges. *Lancet Neurol.* 2010. doi:10.1016/S1474-4422(10)70104-6
2. Caplan LR. Lacunar infarction and small vessel disease: Pathology and pathophysiology. *J Stroke.* 2015. doi:10.5853/jos.2015.17.1.2
3. Wardlaw JM, Smith C, Dichgans M. Mechanisms of sporadic cerebral small vessel disease: Insights from neuroimaging. *Lancet Neurol.* 2013. doi:10.1016/S1474-4422(13)70060-7
4. Bailey EL, Smith C, Sudlow CLM, Wardlaw JM. Pathology of lacunar ischemic stroke in humans - A systematic review. *Brain Pathol.* 2012. doi:10.1111/j.1750-3639.2012.00575.x
5. Wardlaw JM, Smith C, Dichgans M. Small vessel disease: mechanisms and clinical implications. *Lancet Neurol.* 2019;18(7):684-696. doi:10.1016/S1474-4422(19)30079-1
6. Shi Y, Thrippleton MJ, Makin SD, et al. Cerebral blood flow in small vessel disease: A systematic review and meta-analysis. *J Cereb Blood Flow Metab.* 2016;36(10):1653-1667. doi:10.1177/0271678X16662891
7. Promjunyakul N, Lahna D, Kaye JA, et al. Characterizing the white matter hyperintensity penumbra with cerebral blood flow measures. *NeuroImage Clin.* 2015. doi:10.1016/j.nicl.2015.04.012
8. Bernbaum M, Menon BK, Fick G, et al. Reduced blood flow in normal white matter predicts development of leukoaraiosis. *J Cereb Blood Flow Metab.* 2015. doi:10.1038/jcbfm.2015.92
9. Ten Dam VH, Van Den Heuvel DMJ, De Craen AJM, et al. Decline in total cerebral blood flow is linked with increase in periventricular but not deep white matter hyperintensities. *Radiology.* 2007. doi:10.1148/radiol.2431052111
10. Van Der Veen PH, Muller M, Vincken KL, et al. Longitudinal Relationship between Cerebral Small-Vessel Disease and Cerebral Blood Flow. *Stroke.* 2015;46(5):1233-1238. doi:10.1161/STROKEAHA.114.008030
11. Daneman R, Prat A. The blood–brain barrier. *Cold Spring Harb Perspect Biol.* 2015. doi:10.1101/cshperspect.a020412
12. Zhang W, Zhu L, An C, et al. The blood brain barrier in cerebral ischemic injury – Disruption and repair. *Brain Hemorrhages.* 2020;1(1):34-53. doi:10.1016/j.hest.2019.12.004
13. Alafuzoff I, Adolfsson R, Bucht G, Winblad B. Albumin and immunoglobulin in plasma and cerebrospinal fluid, and blood-cerebrospinal fluid barrier function in patients with dementia of alzheimer type and multi-infarct dementia. *J Neurol Sci.* 1983. doi:10.1016/0022-510X(83)90157-0
14. Wada H. Blood-Brain Barrier Permeability of the Demented Elderly as Studied by Cerebrospinal Fluid-Serum Albumin Ratio. *Intern Med.* 1998. doi:10.2169/internalmedicine.37.509
15. Dichgans M, Wick M, Gasser T. Cerebrospinal fluid findings in CADASIL. *Neurology.* 1999. doi:10.1212/WNL.53.1.233
16. Tomimoto H, Akiguchi I, Suenaga T, et al. Alterations of the blood-brain barrier and glial cells in white-matter lesions in cerebrovascular and Alzheimer's disease patients. *Stroke.* 1996. doi:10.1161/01.STR.27.11.2069
17. Topakian R, Barrick TR, Howe FA, Markus HS. Blood-brain barrier permeability is increased in normal-appearing white matter in patients with lacunar stroke and leukoaraiosis. *J Neurol Neurosurg Psychiatry.* 2010. doi:10.1136/jnnp.2009.172072
18. Zhang CE, Wong SM, Van De Haar HJ, et al. Blood-brain barrier leakage is more widespread in patients with cerebral small vessel disease. *Neurology.*

2017. doi:10.1212/WNL.0000000000003556
19. Bronge L, Wahlund LO. White matter lesions in dementia: An MRI study on blood-brain barrier dysfunction. *Dement Geriatr Cogn Disord*. 2000;11(5):263-267. doi:10.1159/000017248
 20. Hanyu H, Asano T, Tanaka Y, Iwamoto T, Takasaki M, Abe K. Increased blood-brain barrier permeability in white matter lesions of Binswanger's disease evaluated by contrast-enhanced MRI. *Dement Geriatr Cogn Disord*. 2002. doi:10.1159/000058326
 21. Smith EE, Eichler F. Cerebral amyloid angiopathy and lobar intracerebral hemorrhage. *Arch Neurol*. 2006. doi:10.1001/archneur.63.1.148
 22. Choi JC. Genetics of cerebral small vessel disease. *J Stroke*. 2015;17(1):7-16. doi:10.5853/jos.2015.17.1.7
 23. Biffi A, Shulman JM, Jagiella JM, et al. Genetic variation at CR1 increases risk of cerebral amyloid angiopathy. *Neurology*. 2012. doi:10.1212/WNL.0b013e3182452b40
 24. Biffi A, Greenberg SM. Cerebral amyloid angiopathy: A systematic review. *J Clin Neurol*. 2011. doi:10.3988/jcn.2011.7.1.1
 25. Charidimou A, Boulouis G, Gurol ME, et al. Emerging concepts in sporadic cerebral amyloid angiopathy. *Brain*. 2017. doi:10.1093/brain/awx047
 26. Chabriat H, Joutel A, Dichgans M, Tournier-Lasserre E, Bousser MG. CADASIL. *Lancet Neurol*. 2009. doi:10.1016/S1474-4422(09)70127-9
 27. Joutel A, Corpechot C, Ducros A, et al. Notch3 mutations in CADASIL, a hereditary adult-onset condition causing stroke and dementia. *Nature*. 1996. doi:10.1038/383707a0
 28. Joutel A, Vahedi K, Corpechot C, et al. Strong clustering and stereotyped nature of Notch3 mutations in CADASIL patients. *Lancet*. 1997. doi:10.1016/S0140-6736(97)08083-5
 29. Locatelli M, Padovani A, Pezzini A. Pathophysiological Mechanisms and Potential Therapeutic Targets in Cerebral Autosomal Dominant Arteriopathy With Subcortical Infarcts and Leukoencephalopathy (CADASIL). *Front Pharmacol*. 2020;11(March):1-13. doi:10.3389/fphar.2020.00321
 30. Rutten JW, Dauwerse HG, Gravesteijn G, et al. Archetypal NOTCH3 mutations frequent in public exome: implications for CADASIL. *Ann Clin Transl Neurol*. 2016;3(11):844-853. doi:10.1002/acn3.344
 31. Chabriat H, Bousser MG. Neuropsychiatric manifestations in CADASIL. *Dialogues Clin Neurosci*. 2007.
 32. Drazyk AM, Tan RYY, Tay J, Traylor M, Das T, Markus HS. Encephalopathy in a Large Cohort of British Cerebral Autosomal Dominant Arteriopathy With Subcortical Infarcts and Leukoencephalopathy Patients. *Stroke*. 2019;50(2):283-290. doi:10.1161/STROKEAHA.118.023661
 33. Opherk C, Peters N, Herzog J, Luedtke R, Dichgans M. Long-term prognosis and causes of death in CADASIL: A retrospective study in 411 patients. *Brain*. 2004;127(11):2533-2539. doi:10.1093/brain/awh282
 34. Buffon F, Porcher R, Hernandez K, et al. Cognitive profile in CADASIL. *J Neurol Neurosurg Psychiatry*. 2006. doi:10.1136/jnnp.2005.068726
 35. Guey S, Mawet J, Hervé D, et al. Prevalence and characteristics of migraine in CADASIL. *Cephalalgia*. 2016. doi:10.1177/0333102415620909
 36. Reyes S, Viswanathan A, Godin O, et al. Apathy: A major symptom in Cadasil. *Neurology*. 2009. doi:10.1212/01.wnl.0000344166.03470.f8
 37. Adib-Samii P, Brice G, Martin RJ, Markus HS. Clinical spectrum of CADASIL and the effect of cardiovascular risk factors on phenotype: Study in 200 consecutively recruited individuals. *Stroke*. 2010;41(4):630-634. doi:10.1161/STROKEAHA.109.568402
 38. Bersano A, Bedini G, Oskam J, et al. CADASIL: Treatment and Management Options. *Curr Treat Options Neurol*. 2017;19(9). doi:10.1007/s11940-017-

- 0468-z
39. Charlton RA, Morris RG, Nitkunan A, Markus HS. The cognitive profiles of CADASIL and sporadic small vessel disease. *Neurology*. 2006;66(10):1523-1526. doi:10.1212/01.wnl.0000216270.02610.7e
 40. Dichgans M. Cognition in CADASIL. In: *Stroke*. ; 2009. doi:10.1161/STROKEAHA.108.534412
 41. Jahn H. Memory loss in alzheimer's disease. *Dialogues Clin Neurosci*. 2013. doi:10.31887/dcms.2013.15.4/hjahn
 42. 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. *PLoS One*. 2015;10(8):1-15. doi:10.1371/journal.pone.0135523
 43. Chui HC, Zarow C, Mack WJ, et al. Cognitive impact of subcortical vascular and Alzheimer's disease pathology. *Ann Neurol*. 2006. doi:10.1002/ana.21009
 44. Zeestraten EA, Lawrence AJ, Lambert C, et al. Change in multimodal MRI markers predicts dementia risk in cerebral small vessel disease. *Neurology*. 2017. doi:10.1212/WNL.0000000000004594
 45. Vermeer SE, Prins ND, Den Heijer T, Hofman A, Koudstaal PJ, Breteler MMB. Silent brain infarcts and the risk of dementia and cognitive decline. *N Engl J Med*. 2003. doi:10.1056/NEJMoa022066
 46. Staekenborg SS, Su T, Van Straaten ECW, et al. Behavioural and psychological symptoms in vascular dementia; differences between small- and large-vessel disease. *J Neurol Neurosurg Psychiatry*. 2010;81(5):547-551. doi:10.1136/jnnp.2009.187500
 47. American Psychiatric Association. DSM-5 Diagnostic Classification. In: *Diagnostic and Statistical Manual of Mental Disorders*. ; 2013. doi:10.1176/appi.books.9780890425596.x00diagnosticclassification
 48. Robert P, Lanctôt KL, Agüera-Ortiz L, et al. Is it time to revise the diagnostic criteria for apathy in brain disorders? The 2018 international consensus group. *Eur Psychiatry*. 2018;54:71-76. doi:10.1016/j.eurpsy.2018.07.008
 49. Withall A, Brodaty H, Altendorf A, Sachdev PS. A longitudinal study examining the independence of apathy and depression after stroke : the Sydney Stroke Study. 2020;(2011):264-273. doi:10.1017/S1041610209991116
 50. De Groot JC, De Leeuw FE, Oudkerk M, Hofman A, Jolles J, Breteler MMB. Cerebral white matter lesions and depressive symptoms in elderly adults. *Arch Gen Psychiatry*. 2000. doi:10.1001/archpsyc.57.11.1071
 51. Tay J, Tuladhar AM, Hollocks MJ, et al. Apathy is associated with large-scale white matter network disruption in small vessel disease. *Neurology*. 2019;92(11):E1157-E1167. doi:10.1212/WNL.0000000000007095
 52. Hollocks MJ, Lawrence AJ, Brookes RL, et al. Differential relationships between apathy and depression with white matter microstructural changes and functional outcomes. *Brain*. 2015;138(12):3803-3815. doi:10.1093/brain/awv304
 53. Dalen JW Van, Wanrooij LL Van, Charante EPM Van. Apathy is associated with incident dementia in community-dwelling older people. 2018;0. doi:10.1212/WNL.0000000000004767
 54. Tay J, Morris RG, Tuladhar AM, Husain M, De Leeuw FE, Markus HS. Apathy, but not depression, predicts all-cause dementia in cerebral small vessel disease. *J Neurol Neurosurg Psychiatry*. 2020. doi:10.1136/jnnp-2020-323092
 55. Valenti R, Poggesi A, Pescini F, Inzitari D, Pantoni L. Psychiatric disturbances in CADASIL: A brief review. *Acta Neurol Scand*. 2008. doi:10.1111/j.1600-0404.2008.01015.x
 56. De Laat KF, Van Norden AGW, Gons RAR, et al. Gait in elderly with cerebral small vessel disease. *Stroke*. 2010. doi:10.1161/STROKEAHA.110.583229
 57. Pinter D, Ritchie SJ, Doubal F, et al. Impact of small vessel disease in the

- brain on gait and balance. *Sci Rep*. 2017. doi:10.1038/srep41637
58. Baezner H, Blahak C, Poggesi A, et al. Association of gait and balance disorders with age-related white matter changes: The LADIS Study. *Neurology*. 2008;70(12):935-942. doi:10.1212/01.wnl.0000305959.46197.e6
 59. Finsterwalder S, Wuehr M, Gesierich B, et al. Minor gait impairment despite white matter damage in pure small vessel disease. *Ann Clin Transl Neurol*. 2019. doi:10.1002/acn3.50891
 60. Norrving B. Long-term prognosis after lacunar infarction. *Lancet Neurol*. 2003;2(4):238-245. doi:10.1016/S1474-4422(03)00352-1
 61. Benavente O, White CL, Roldan AM. Small vessel strokes. *Curr Cardiol Rep*. 2005. doi:10.1007/s11886-005-0006-6
 62. Petty GW, Brown RD, Whisnant JP, Sicks JRD, O'Fallon WM, Wiebers DO. Ischemic stroke subtypes: A population-based study of functional outcome, survival, and recurrence. *Stroke*. 2000. doi:10.1161/01.STR.31.5.1062
 63. Bamford J, Sandercock P, Jones L, Warlow C. The natural history of lacunar infarction: The oxfordshire community stroke project. *Stroke*. 1987. doi:10.1161/01.STR.18.3.545
 64. 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. *PLoS One*. 2015;10(8):1-15. doi:10.1371/journal.pone.0135523
 65. Lawrence AJ, Zeestraten EA, Morris RG, et al. Change in multimodal MRI markers predicts dementia risk in cerebral small vessel disease. *Neurology*. 2017;89(18):1869-1876. doi:10.1212/wnl.0000000000004594
 66. Benjamin P, Zeestraten E, Lambert C, 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):228-240. doi:10.1038/jcbfm.2015.113
 67. Eggink E, Moll van Charante EP, van Gool WA, Richard E. A Population Perspective on Prevention of Dementia. *J Clin Med*. 2019. doi:10.3390/jcm8060834
 68. Wardlaw JM, Smith EE, Biessels GJ, et al. Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. *Lancet Neurol*. 2013. doi:10.1016/S1474-4422(13)70124-8
 69. Symms M, Jäger HR, Schmierer K, Yousry TA. A review of structural magnetic resonance neuroimaging. *J Neurol Neurosurg Psychiatry*. 2004. doi:10.1136/jnnp.2003.032714
 70. Le Bihan D. How MRI Makes the Brain Visible. In: *Make Life Visible*. ; 2020. doi:10.1007/978-981-13-7908-6_20
 71. Hachinski VC, Potter P, Merskey H. Leuko-Araiosis: An Ancient Term for a New Problem. *Can J Neurol Sci / J Can des Sci Neurol*. 1986;13(S4):533-534. doi:10.1017/S0317167100037264
 72. Gouw AA, Seewann A, Van Der Flier WM, et al. Heterogeneity of small vessel disease: A systematic review of MRI and histopathology correlations. *J Neurol Neurosurg Psychiatry*. 2011. doi:10.1136/jnnp.2009.204685
 73. Fazekas F, Chawluk JB, Alavi A, Hurtig HI, Zimmerman RA. MR signal abnormalities at 1.5 T in Alzheimer's dementia and normal aging. *Am J Roentgenol*. 1987. doi:10.2214/ajr.149.2.351
 74. Wardlaw JM, Valdés Hernández MC, Muñoz-Maniega S. What are white matter hyperintensities made of? Relevance to vascular cognitive impairment. *J Am Heart Assoc*. 2015. doi:10.1161/JAHA.114.001140
 75. Pantoni L. Leukoaraiosis: From an ancient term to an actual marker of poor prognosis. *Stroke*. 2008;39(5):1401-1403. doi:10.1161/STROKEAHA.107.505602
 76. Sachdev PS, Wen W, Christensen H, Jorm AF. White matter hyperintensities

- are related to physical disability and poor motor function. *J Neurol Neurosurg Psychiatry*. 2005. doi:10.1136/jnnp.2004.042945
77. Longstreth WT, Manolio TA, Arnold A, et al. Clinical correlates of white matter findings on cranial magnetic resonance imaging of 3301 elderly people: The cardiovascular health study. *Stroke*. 1996. doi:10.1161/01.STR.27.8.1274
 78. Nylander R, Kilander L, Ahlström H, Lind L, Larsson EM. Small vessel disease on neuroimaging in a 75-Year-Old Cohort (PIVUS): Comparison with cognitive and executive tests. *Front Aging Neurosci*. 2018;10(JUL):1-8. doi:10.3389/fnagi.2018.00217
 79. Bolandzadeh N, Davis JC, Tam R, Handy TC, Liu-Ambrose T. The association between cognitive function and white matter lesion location in older adults: A systematic review. *BMC Neurol*. 2012;12(1):1. doi:10.1186/1471-2377-12-126
 80. Griffanti L, Jenkinson M, Suri S, et al. Classification and characterization of periventricular and deep white matter hyperintensities on MRI: A study in older adults. *Neuroimage*. 2018;170(March 2017):174-181. doi:10.1016/j.neuroimage.2017.03.024
 81. Kim KW, MacFall JR, Payne ME. Classification of White Matter Lesions on Magnetic Resonance Imaging in Elderly Persons. *Biol Psychiatry*. 2008. doi:10.1016/j.biopsych.2008.03.024
 82. Schmidt R, Schmidt H, Haybaeck J, et al. Heterogeneity in age-related white matter changes. *Acta Neuropathol*. 2011. doi:10.1007/s00401-011-0851-x
 83. Wright CB, Festa JR, Paik MC, et al. White matter hyperintensities and subclinical infarction: Associations with psychomotor speed and cognitive flexibility. *Stroke*. 2008. doi:10.1161/STROKEAHA.107.484147
 84. 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. doi:10.1136/jnnp.2003.014910
 85. 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. doi:10.1161/STROKEAHA.107.507475
 86. Lawrence AJ, Patel B, Morris RG, 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. doi:10.1371/journal.pone.0061014
 87. Cook IA, Leuchter AF, Morgan ML, et al. Longitudinal Progression of Subclinical Structural Brain Disease in Normal Aging. *Am J Geriatr Psychiatry*. 2004. doi:10.1097/00019442-200403000-00010
 88. Van Den Heuvel DMJ, Ten Dam VH, De Craen AJM, et al. Increase in periventricular white matter hyperintensities parallels decline in mental processing speed in a non-demented elderly population. *J Neurol Neurosurg Psychiatry*. 2006. doi:10.1136/jnnp.2005.070193
 89. Van Dijk EJ, Prins ND, Vrooman HA, Hofman A, Koudstaal PJ, Breteler MMB. Progression of cerebral small vessel disease in relation to risk factors and cognitive consequences: Rotterdam scan study. *Stroke*. 2008;39(10):2712-2719. doi:10.1161/STROKEAHA.107.513176
 90. Maillard P, Carmichael O, Fletcher E, Reed B, Mungas D, DeCarli C. Coevolution of white matter hyperintensities and cognition in the elderly. *Neurology*. 2012;79(5):442-448. doi:10.1212/WNL.0b013e3182617136
 91. Van Leijsen EMC, Van Uden IWM, Ghafoorian M, et al. Nonlinear temporal dynamics of cerebral small vessel disease. *Neurology*. 2017. doi:10.1212/WNL.0000000000004490
 92. Gouw AA, Van Der Flier WM, Fazekas F, et al. Progression of white matter hyperintensities and incidence of new lacunes over a 3-year period: The

- leukoaraiosis and disability study. *Stroke*. 2008. doi:10.1161/STROKEAHA.107.498535
93. Schmidt R, Petrovic K, Ropele S, Enzinger C, Fazekas F. Progression of leukoaraiosis and cognition. *Stroke*. 2007. doi:10.1161/STROKEAHA.107.489112
 94. Verhaaren BFJ, Vernooij MW, De Boer R, et al. High blood pressure and cerebral white matter lesion progression in the general population. *Hypertension*. 2013. doi:10.1161/HYPERTENSIONAHA.111.00430
 95. Garde E, Lykke Mortensen E, Rostrup E, Paulson OB. Decline in intelligence is associated with progression in white matter hyperintensity volume. *J Neurol Neurosurg Psychiatry*. 2005. doi:10.1136/jnnp.2004.055905
 96. Longstreth WT, Arnold AM, Beauchamp NJ, et al. Incidence, manifestations, and predictors of worsening white matter on serial cranial magnetic resonance imaging in the elderly: The cardiovascular health study. *Stroke*. 2005. doi:10.1161/01.STR.0000149625.99732.69
 97. Van Straaten ECW, Harvey D, Scheltens P, et al. Periventricular white matter hyperintensities increase the likelihood of progression from amnesic mild cognitive impairment to dementia. *J Neurol*. 2008. doi:10.1007/s00415-008-0874-y
 98. van Leijssen EMC, Bergkamp MI, van Uden IWM, et al. Cognitive consequences of regression of cerebral small vessel disease. *Eur Stroke J*. 2019;4(1):85-89. doi:10.1177/2396987318820790
 99. Patel B, Markus HS. Magnetic resonance imaging in cerebral small vessel disease and its use as a surrogate disease marker. *Int J Stroke*. 2011. doi:10.1111/j.1747-4949.2010.00552.x
 100. Schmidt R, Scheltens P, Erkinjuntti T, et al. White matter lesion progression: A surrogate endpoint for trials in cerebral small-vessel disease. *Neurology*. 2004;63(1):139-144. doi:10.1212/01.WNL.0000132635.75819.E5
 101. Nasrallah IM, Pajewski NM, Auchus AP, et al. Association of intensive vs standard blood pressure control with cerebral white matter lesions. *JAMA - J Am Med Assoc*. 2019;322(6):524-534. doi:10.1001/jama.2019.10551
 102. Shi Y, Wardlaw JM. Update on cerebral small vessel disease: A dynamic whole-brain disease. *Stroke Vasc Neurol*. 2016;1(3):83-92. doi:10.1136/svn-2016-000035
 103. Okazaki S, Hornberger E, Griebe M, Gass A, Hennerici MG, Szabo K. MRI characteristics of the evolution of supratentorial recent small subcortical infarcts. *Front Neurol*. 2015;6(MAY):1-6. doi:10.3389/fneur.2015.00118
 104. Loos CMJ, Staals J, Wardlaw JM, Van Oostenbrugge RJ. Cavitation of deep lacunar infarcts in patients with first-ever lacunar stroke: A 2-year follow-up study with MR. *Stroke*. 2012;43(8):2245-2247. doi:10.1161/STROKEAHA.112.660076
 105. Wardlaw JM. What is a lacune? *Stroke*. 2008;39(11):2921-2922. doi:10.1161/STROKEAHA.108.523795
 106. Bamford J, Sandercock P, Dennis M, Warlow C, Burn J. Classification and natural history of clinically identifiable subtypes of cerebral infarction. *Lancet*. 1991. doi:10.1016/0140-6736(91)93206-O
 107. Fisher CM. Lacunar strokes and infarcts: A review. *Neurology*. 1982. doi:10.1212/wnl.32.8.871
 108. Yamamoto Y, Akiguchi I, Oiwa K, Hayashi M, Kasai T, Ozasa K. Twenty-four-hour blood pressure and MRI as predictive factors for different outcomes in patients with lacunar infarct. *Stroke*. 2002. doi:10.1161/str.33.1.297
 109. Miyao S, Takano A, Teramoto J, Takahashi A. Leukoaraiosis in relation to prognosis for patients with lacunar infarction. *Stroke*. 1992. doi:10.1161/01.STR.23.10.1434
 110. Choi P, Ren M, Phan TG, et al. Silent infarcts and cerebral microbleeds modify

- the associations of white matter lesions with gait and postural stability: Population-based study. *Stroke*. 2012;43(6):1505-1510. doi:10.1161/STROKEAHA.111.647271
111. Schneider JA, Arvanitakis Z, Bang W, Bennett DA. Mixed brain pathologies account for most dementia cases in community-dwelling older persons. *Neurology*. 2007. doi:10.1212/01.wnl.0000271090.28148.24
 112. Vermeer SE, Longstreth WT, Koudstaal PJ. Silent brain infarcts: a systematic review. *Lancet Neurol*. 2007. doi:10.1016/S1474-4422(07)70170-9
 113. Baykara E, Gesierich B, Adam R, 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). doi:10.1002/ana.24758
 114. Patel B, Lawrence AJ, Chung AW, et al. Cerebral microbleeds and cognition in patients with symptomatic small vessel disease. *Stroke*. 2013;44(2):356-361. doi:10.1161/STROKEAHA.112.670216
 115. Gregoire SM, Chaudhary UJ, Brown MM, et al. The Microbleed Anatomical Rating Scale (MARS). *Neurology*. 2009. doi:10.1212/WNL.0b013e3181c34a7d
 116. Cordonnier C, Potter GM, Jackson CA, et al. Improving interrater agreement about brain microbleeds: Development of the Brain Observer MicroBleed Scale (BOMBS). *Stroke*. 2009. doi:10.1161/STROKEAHA.108.526996
 117. Greenberg SM, Vernooij MW, Cordonnier C, et al. Cerebral microbleeds: a guide to detection and interpretation. *Lancet Neurol*. 2009. doi:10.1016/S1474-4422(09)70013-4
 118. Jia Z, Mohammed W, Qiu Y, Hong X, Shi H. Hypertension increases the risk of cerebral microbleed in the territory of posterior cerebral artery: A study of the association of microbleeds categorized on a basis of vascular territories and cardiovascular risk factors. *J Stroke Cerebrovasc Dis*. 2014. doi:10.1016/j.jstrokecerebrovasdis.2012.12.016
 119. Elmståhl S, Ellström K, Siennicki-Lantz A, Abul-Kasim K. Association between cerebral microbleeds and hypertension in the Swedish general population "Good Aging in Skåne" study. *J Clin Hypertens*. 2019. doi:10.1111/jch.13606
 120. Staals J, Van Oostenbrugge RJ, Knottnerus ILH, Rouhl RPW, Henskens LHG, Lodder J. Brain microbleeds relate to higher ambulatory blood pressure levels in first-ever lacunar stroke patients. *Stroke*. 2009. doi:10.1161/STROKEAHA.109.558049
 121. Ding J, Sigurdsson S, Garcia M, et al. Risk factors associated with incident cerebral microbleeds according to location in older people: The Age, Gene/Environment Susceptibility (AGES)-Reykjavik study. *JAMA Neurol*. 2015. doi:10.1001/jamaneurol.2015.0174
 122. Gao Z, Wang W, Wang Z, et al. Cerebral microbleeds are associated with deep white matter hyperintensities, but only in hypertensive patients. *PLoS One*. 2014. doi:10.1371/journal.pone.0091637
 123. Poels MMF, Vernooij MW, Ikram MA, et al. Prevalence and risk factors of cerebral microbleeds: An update of the rotterdam scan study. In: *Stroke*. ; 2010. doi:10.1161/STROKEAHA.110.595181
 124. Jeerakathil T, Wolf PA, Beiser A, et al. Cerebral microbleeds: Prevalence and associations with cardiovascular risk factors in the Framingham Study. *Stroke*. 2004. doi:10.1161/01.STR.0000131809.35202.1b
 125. Cordonnier C, Al-Shahi Salman R, Wardlaw J. Spontaneous brain microbleeds: Systematic review, subgroup analyses and standards for study design and reporting. *Brain*. 2007. doi:10.1093/brain/awl387
 126. Poels MMF, Ikram MA, Van Der Lugt A, et al. Incidence of cerebral microbleeds in the general population: The Rotterdam Scan Study. *Stroke*. 2011. doi:10.1161/STROKEAHA.110.607184
 127. Nishikawa T, Ueba T, Kajiwara M, Miyamatsu N, Yamashita K. Cerebral microbleeds in patients with intracerebral hemorrhage are associated with

- previous cerebrovascular diseases and white matter hyperintensity, but not with regular use of antiplatelet agents. *Neurol Med Chir (Tokyo)*. 2009. doi:10.2176/nmc.49.333
128. Fan YH, Zhang L, Lam WWM, Mok VCT, Wong KS. Cerebral microbleeds as a risk factor for subsequent intracerebral hemorrhages among patients with acute ischemic stroke. *Stroke*. 2003. doi:10.1161/01.STR.0000090841.90286.81
 129. Bokura H, Saika R, Yamaguchi T, et al. Microbleeds are associated with subsequent hemorrhagic and ischemic stroke in healthy elderly individuals. *Stroke*. 2011. doi:10.1161/STROKEAHA.110.601922
 130. Boulanger JM, Coutts SB, Eliasziw M, et al. Cerebral microhemorrhages predict new disabling or fatal strokes in patients with acute ischemic stroke or transient ischemic attack. *Stroke*. 2006. doi:10.1161/01.STR.0000204237.66466.5f
 131. Thijs V, Lemmens R, Schoofs C, et al. Microbleeds and the risk of recurrent stroke. *Stroke*. 2010. doi:10.1161/STROKEAHA.110.588020
 132. Van Norden AGW, Van Den Berg HAC, De Laat KF, Gons RAR, Van Dijk EJ, De Leeuw FE. Frontal and temporal microbleeds are related to cognitive function: The Radboud University Nijmegen diffusion tensor and magnetic resonance cohort (RUN DMC) study. *Stroke*. 2011. doi:10.1161/STROKEAHA.111.629634
 133. Qiu C, Cotch MF, Sigurdsson S, et al. Cerebral microbleeds, retinopathy, and dementia: The AGES-Reykjavik Study. *Neurology*. 2010. doi:10.1212/WNL.0b013e3182020349
 134. Werring DJ, Frazer DW, Coward LJ, et al. Cognitive dysfunction in patients with cerebral microbleeds on T2*-weighted gradient-echo MRI. *Brain*. 2004. doi:10.1093/brain/awh253
 135. Knudsen KA, Rosand J, Karluk D, Greenberg SM. Clinical diagnosis of cerebral amyloid angiopathy: Validation of the boston criteria. *Neurology*. 2001. doi:10.1212/WNL.56.4.537
 136. Linn J, Halpin A, Demaerel P, et al. Prevalence of superficial siderosis in patients with cerebral amyloid angiopathy. *Neurology*. 2010. doi:10.1212/WNL.0b013e3181dad605
 137. Fazekas F, Kleinert R, Roob G, et al. Histopathologic analysis of foci of signal loss on gradient-echo T2*- weighted MR images in patients with spontaneous intracerebral hemorrhage: Evidence of microangiopathy-related microbleeds. *Am J Neuroradiol*. 1999.
 138. Van Norden AGW, Van Den Berg HAC, De Laat KF, Gons RAR, Van Dijk EJ, De Leeuw FE. Frontal and temporal microbleeds are related to cognitive function: The Radboud University Nijmegen diffusion tensor and magnetic resonance cohort (RUN DMC) study. *Stroke*. 2011;42(12):3382-3386. doi:10.1161/STROKEAHA.111.629634
 139. Liem MK, Lesnik Oberstein SAJ, Haan J, et al. MRI correlates of cognitive decline in CADASIL: A 7-year follow-up study. *Neurology*. 2009. doi:10.1212/01.wnl.0000339038.65508.96
 140. Meier IB, Gu Y, Guzaman VA, et al. Lobar microbleeds are associated with a decline in executive functioning in older adults. *Cerebrovasc Dis*. 2014. doi:10.1159/000368998
 141. Gregoire SM, Smith K, Jäger HR, et al. Cerebral microbleeds and long-term cognitive outcome: Longitudinal cohort study of stroke clinic patients. *Cerebrovasc Dis*. 2012;33(5):430-435. doi:10.1159/000336237
 142. Mathalon DH, Sullivan E V., Rawles JM, Pfefferbaum A. Correction for head size in brain-imaging measurements. *Psychiatry Res Neuroimaging*. 1993. doi:10.1016/0925-4927(93)90016-B
 143. Appelman APA, Exalto LG, Van Der Graaf Y, Biessels GJ, Mali WPTM,

- Geerlings MI. White matter lesions and brain atrophy: More than shared risk factors? A systematic review. *Cerebrovasc Dis*. 2009;28(3):227-242. doi:10.1159/000226774
144. Aribisala BS, Valdés Hernández MC, Royle NA, et al. Brain atrophy associations with white matter lesions in the ageing brain: The Lothian Birth Cohort 1936. *Eur Radiol*. 2013;23(4):1084-1092. doi:10.1007/s00330-012-2677-x
 145. Jagust WJ, Zheng L, Harvey DJ, et al. Neuropathological basis of magnetic resonance images in aging and dementia. *Ann Neurol*. 2008. doi:10.1002/ana.21296
 146. Nitkunan A, Lanfranconi S, Charlton RA, Barrick TR, Markus HS. Brain atrophy and cerebral small vessel disease a prospective follow-up study. *Stroke*. 2011;42(1):133-138. doi:10.1161/STROKEAHA.110.594267
 147. Jokinen H, Lipsanen J, Schmidt R, et al. Brain atrophy accelerates cognitive decline in cerebral small vessel disease The LADIS study. *Neurology*. 2012;78(22):1785-1792. doi:10.1212/WNL.0b013e3182583070
 148. Lambert C, Benjamin P, Zeestraten E, Lawrence AJ, Barrick TR, Markus HS. Longitudinal patterns of leukoaraiosis and brain atrophy in symptomatic small vessel disease. *Brain*. 2016. doi:10.1093/brain/aww009
 149. Duering M, Csanadi E, Gesierich B, et al. Incident lacunes preferentially localize to the edge of white matter hyperintensities: insights into the pathophysiology of cerebral small vessel disease. *Brain*. 2013;136(9):2717-2726. doi:10.1093/brain/awt184
 150. Braffman BH, Zimmerman RA, Trojanowski JQ, Gonatas NK, Hickey WF, Schlaepfer WW. Brain MR: Pathologic correlation with gross and histopathology. 1. Lacunar infarction and Virchow-Robin spaces. *Am J Roentgenol*. 1988;151(3):551-558. doi:10.2214/ajr.151.3.551
 151. Brown R, Benveniste H, Black SE, et al. Understanding the role of the perivascular space in cerebral small vessel disease. *Cardiovasc Res*. 2018;114(11):1462-1473. doi:10.1093/cvr/cvy113
 152. Rudie JD, Rauschecker AM, Nabavizadeh SA, Mohan S. Neuroimaging of Dilated Perivascular Spaces: From Benign and Pathologic Causes to Mimics. *J Neuroimaging*. 2018;28(2):139-149. doi:10.1111/jon.12493
 153. Groeschel S, Hanefeld F. Virchow-Robin spaces on magnetic resonance images: normative data and their association with diseases. *Neuropediatrics*. 2006. doi:10.1055/s-2006-953596
 154. MacLulich AMJ, Wardlaw JM, Ferguson KJ, Starr JM, Seckl JR, Deary IJ. Enlarged perivascular spaces are associated with cognitive function in healthy elderly men. *J Neurol Neurosurg Psychiatry*. 2004. doi:10.1136/jnnp.2003.030858
 155. Passiak BS, Liu D, Kresge HA, et al. Perivascular spaces contribute to cognition beyond other small vessel disease markers. *Neurology*. 2019. doi:10.1212/WNL.00000000000007124
 156. Doubal FN, MacLulich AMJ, Ferguson KJ, Dennis MS, Wardlaw JM. Enlarged Perivascular Spaces on MRI Are a Feature of Cerebral Small Vessel Disease. *Stroke*. 2010. doi:10.1161/STROKEAHA.109.564914
 157. Zhu YC, Tzourio C, Soumaré A, Mazoyer B, Dufouil C, Chabriat H. Severity of dilated virchow-robin spaces is associated with age, blood pressure, and MRI markers of small vessel disease: A population-based study. *Stroke*. 2010. doi:10.1161/STROKEAHA.110.591586
 158. Potter GM, Doubal FN, Jackson CA, et al. Enlarged perivascular spaces and cerebral small vessel disease. *Int J Stroke*. 2015. doi:10.1111/ijvs.12054
 159. Benjamin P, Trippier S, Lawrence AJ, et al. Lacunar infarcts, but not perivascular spaces, are predictors of cognitive decline in cerebral small-vessel disease. *Stroke*. 2018;49(3). doi:10.1161/STROKEAHA.117.017526

160. Benjamin P, Trippier S, Lawrence AJ, et al. Lacunar infarcts, but not perivascular spaces, are predictors of cognitive decline in cerebral small-vessel disease. *Stroke*. 2018;49(3):586-593. doi:10.1161/STROKEAHA.117.017526
161. Staals J, Makin SDJ, Doubal FN, Dennis MS, Wardlaw JM. Stroke subtype, vascular risk factors, and total MRI brain small-vessel disease burden. *Neurology*. 2014. doi:10.1212/WNL.0000000000000837
162. Klarenbeek P, Van Oostenbrugge RJ, Rouhl RPW, Knottnerus ILH, Staals J. Ambulatory blood pressure in patients with lacunar stroke: Association with total MRI burden of cerebral small vessel disease. *Stroke*. 2013. doi:10.1161/STROKEAHA.113.002545
163. Huijts M, Duits A, Van Oostenbrugge RJ, Kroon AA, De Leeuw PW, Staals J. Accumulation of MRI markers of cerebral small vessel disease is associated with decreased cognitive function. A study in first-ever lacunar stroke and hypertensive patients. *Front Aging Neurosci*. 2013. doi:10.3389/fnagi.2013.00072
164. Staals J, Booth T, Morris Z, et al. Total MRI load of cerebral small vessel disease and cognitive ability in older people. *Neurobiol Aging*. 2015. doi:10.1016/j.neurobiolaging.2015.06.024
165. Uiterwijk R, van Oostenbrugge RJ, Huijts M, De Leeuw PW, Kroon AA, Staals J. Total cerebral small vessel disease MRI score is associated with cognitive decline in executive function in patients with hypertension. *Front Aging Neurosci*. 2016. doi:10.3389/fnagi.2016.00301
166. Amin AI Olama A, Wason JMS, Tuladhar AM, et al. Simple MRI score aids prediction of dementia in cerebral small vessel disease. *Neurology*. 2020;94(12):e1294-e1302. doi:10.1212/WNL.00000000000009141
167. O'Sullivan M, Summers PE, Jones DK, Jarosz JM, Williams SCR, Markus HS. Normal-appearing white matter in ischemic leukoaraiosis: A diffusion tensor MRI study. *Neurology*. 2001. doi:10.1212/WNL.57.12.2307
168. Le Bihan D, Mangin JF, Poupon C, et al. Diffusion tensor imaging: Concepts and applications. *J Magn Reson Imaging*. 2001. doi:10.1002/jmri.1076
169. Soares JM, Marques P, Alves V, Sousa N. A hitchhiker's guide to diffusion tensor imaging. *Front Neurosci*. 2013. doi:10.3389/fnins.2013.00031
170. O'Sullivan M, Summers PE, Jones DK, Jarosz JM, Williams SCR, Markus HS. Normal-appearing white matter in ischemic leukoaraiosis: A diffusion tensor MRI study. *Neurology*. 2001;57(12):2307-2310. doi:10.1212/WNL.57.12.2307
171. Song SK, Sun SW, Ju WK, Lin SJ, Cross AH, Neufeld AH. Diffusion tensor imaging detects and differentiates axon and myelin degeneration in mouse optic nerve after retinal ischemia. *Neuroimage*. 2003. doi:10.1016/j.neuroimage.2003.07.005
172. Van Norden AGW, De Laat KF, Van Dijk EJ, et al. Diffusion tensor imaging and cognition in cerebral small vessel disease. The RUN DMC study. *Biochim Biophys Acta - Mol Basis Dis*. 2012. doi:10.1016/j.bbadis.2011.04.008
173. D'Souza MM, Gorthi SP, Vadwala K, et al. Diffusion tensor tractography in cerebral small vessel disease: correlation with cognitive function. *Neuroradiol J*. 2018;31(1). doi:10.1177/1971400916682753
174. Jokinen H, Schmidt R, Ropele S, et al. Diffusion changes predict cognitive and functional outcome: The LADIS study. *Ann Neurol*. 2013;73(5):576-583. doi:10.1002/ana.23802
175. Holtmannspötter M, Peters N, Opherk C, et al. Diffusion magnetic resonance histograms as a surrogate marker and predictor of disease progression in CADASIL a two-year follow-up study. *Stroke*. 2005;36(12):2559-2565. doi:10.1161/01.STR.0000189696.70989.a4
176. Croall ID, Lohner V, Moynihan B, et al. Using DTI to assess white matter microstructure in cerebral small vessel disease (SVD) in multicentre studies.

- Clin Sci*. 2017;131(12):1361-1373. doi:10.1042/CS20170146
177. Cuadrado-Godia E, Dwivedi P, Sharma S, et al. Cerebral small vessel disease: A review focusing on pathophysiology, biomarkers, and machine learning strategies. *J Stroke*. 2018. doi:10.5853/jos.2017.02922
 178. Low A, Mak E, Rowe JB, Markus HS, O'Brien JT. Inflammation and cerebral small vessel disease: A systematic review. *Ageing Res Rev*. 2019;53(February):100916. doi:10.1016/j.arr.2019.100916
 179. Ravaglia G, Forti P, Maioli F, et al. Blood inflammatory markers and risk of dementia: The Conselice Study of Brain Aging. *Neurobiol Aging*. 2007. doi:10.1016/j.neurobiolaging.2006.08.012
 180. Miralbell J, Soriano JJ, Spulber G, et al. Structural brain changes and cognition in relation to markers of vascular dysfunction. *Neurobiol Aging*. 2012. doi:10.1016/j.neurobiolaging.2011.09.020
 181. Van Dijk EJ, Prins ND, Vermeer SE, et al. C-reactive protein and cerebral small-vessel disease: The Rotterdam scan study. *Circulation*. 2005. doi:10.1161/CIRCULATIONAHA.104.506337
 182. Aribisala BS, Wiseman S, Morris Z, et al. Circulating inflammatory markers are associated with magnetic resonance imaging-visible perivascular spaces but not directly with white matter hyperintensities. *Stroke*. 2014. doi:10.1161/STROKEAHA.113.004059
 183. Satizabal CL, Zhu YC, Mazoyer B, Dufouil C, Tzourio C. Circulating IL-6 and CRP are associated with MRI findings in the elderly: The 3C-Dijon Study. *Neurology*. 2012. doi:10.1212/WNL.0b013e318248e50f
 184. Shoamanesh A, Preis SR, Beiser AS, et al. Inflammatory biomarkers, cerebral microbleeds, and small vessel disease: Framingham heart study. *Neurology*. 2015. doi:10.1212/WNL.0000000000001279
 185. Wada M, Nagasawa H, Kurita K, et al. Cerebral small vessel disease and C-reactive protein: Results of a cross-sectional study in community-based Japanese elderly. *J Neurol Sci*. 2008. doi:10.1016/j.jns.2007.06.053
 186. Barzilay JI, Lovato JF, Murray AM, et al. Albuminuria and cognitive decline in people with diabetes and normal renal function. *Clin J Am Soc Nephrol*. 2013;8(11):1907-1914. doi:10.2215/CJN.11321112
 187. Takae K, Hata J, Ohara T, et al. Albuminuria increases the risks for both Alzheimer disease and vascular dementia in community-dwelling Japanese elderly: The hisayama study. *J Am Heart Assoc*. 2018;7(2). doi:10.1161/JAHA.117.006693
 188. Yamasaki K, Hata J, Furuta Y, et al. Association of albuminuria with white matter hyperintensities volume on brain magnetic resonance imaging in elderly Japanese — the Hisayama study —. *Circ J*. 2020;84(6):935-942. doi:10.1253/circj.CJ-19-1069
 189. Khalil M, Teunissen CE, Otto M, et al. Neurofilaments as biomarkers in neurological disorders. *Nat Rev Neurol*. 2018;14(10):577-589. doi:10.1038/s41582-018-0058-z
 190. Mattsson N, Andreasson U, Zetterberg H, et al. Association of plasma neurofilament light with neurodegeneration in patients with Alzheimer disease. *JAMA Neurol*. 2017. doi:10.1001/jamaneurol.2016.6117
 191. Rohrer JD, Woollacott IOC, Dick KM, et al. Serum neurofilament light chain protein is a measure of disease intensity in frontotemporal dementia. *Neurology*. 2016. doi:10.1212/WNL.0000000000003154
 192. Khalil M, Pirpamer L, Hofer E, et al. Serum neurofilament light levels in normal aging and their association with morphologic brain changes. *Nat Commun*. 2020. doi:10.1038/s41467-020-14612-6
 193. Duering M, Konieczny MJ, Tiedt S, et al. Serum neurofilament light chain levels are related to small vessel disease burden. *J Stroke*. 2018. doi:10.5853/jos.2017.02565

194. Khan U, Porteous L, Hassan A, Markus HS. Risk factor profile of cerebral small vessel disease and its subtypes. *J Neurol Neurosurg Psychiatry*. 2007;78(7):702-706. doi:10.1136/jnnp.2006.103549
195. Wright JT, Williamson JD, Whelton PK, et al. A randomized trial of intensive versus standard blood-pressure control. *N Engl J Med*. 2015;373(22):2103-2116. doi:10.1056/NEJMoa1511939
196. MacMahon S, Neal B, Tzourio C, et al. Randomised trial of a perindopril-based blood-pressure-lowering regimen among 6105 individuals with previous stroke or transient ischaemic attack. *Lancet*. 2001. doi:10.1016/S0140-6736(01)06178-5
197. Arima H, Chalmers J. PROGRESS: Prevention of recurrent stroke. *J Clin Hypertens*. 2011;13(9):693-702. doi:10.1111/j.1751-7176.2011.00530.x
198. Denker MG, Cohen DL. What is an appropriate blood pressure goal for the elderly: Review of recent studies and practical recommendations. *Clin Interv Aging*. 2013;8:1505-1516. doi:10.2147/CIA.S33087
199. Lesly A. Pearce, Leslie A. McClure, David C. Anderson, Claudia Jacova, Mukul Sharma, Robert G. Hart and ORB. Effects of long-term blood pressure lowering and dual antiplatelet therapy on cognition in patients with recent lacunar stroke: Secondary Prevention of Small Subcortical Strokes (SPS3) trial The SPS3 Investigators*. *Lancet Neurol*. 2014;13(12):1177-1185. doi:doi:10.1016/S1474-4422(14)70224-8.
200. Williamson JD, Pajewski NM, Auchus AP, et al. Effect of Intensive vs Standard Blood Pressure Control on Probable Dementia: A Randomized Clinical Trial. In: *JAMA - Journal of the American Medical Association*. ; 2019. doi:10.1001/jama.2018.21442
201. Zeestraten EA, Benjamin P, Lambert C, et al. Application of diffusion tensor imaging parameters to detect change in longitudinal studies in cerebral small vessel disease. *PLoS One*. 2016;11(1):1-16. doi:10.1371/journal.pone.0147836
202. Croall ID, Lohner V, Moynihan B, et al. Using DTI to assess white matter microstructure in cerebral small vessel disease (SVD) in multicentre studies. 2017;0:1361-1373. doi:10.1042/CS20170146
203. Croall ID, Tozer DJ, Moynihan B, et al. Effect of Standard vs Intensive blood pressure control on cerebral blood flow in small vessel disease the preserve randomized clinical trial. *JAMA Neurol*. 2018;75(6):720-727. doi:10.1001/jamaneurol.2017.5153
204. Jenkinson M, Smith S. A global optimisation method for robust affine registration of brain images. *Med Image Anal*. 2001. doi:10.1016/S1361-8415(01)00036-6
205. Dumont R, Willis JO, Veizel K, Zibulsky J. Wechsler Adult Intelligence Scale-Fourth Edition. In: *Encyclopedia of Special Education*. ; 2014. doi:10.1002/9781118660584.es2520
206. Boake C. From the Binet-Simon to the Wechsler-Bellevue: Tracing the history of intelligence testing. *J Clin Exp Neuropsychol*. 2002;24(3):383-405. doi:10.1076/jcen.24.3.383.981
207. Jaeger J. Digit symbol substitution test. *J Clin Psychopharmacol*. 2018;38(5):513-519. doi:10.1097/JCP.0000000000000941
208. Reitan RM. The relation of the Trail Making Test to organic brain damage. *J Consult Psychol*. 1955. doi:10.1037/h0044509
209. Delis DC, Kaplan E, Kramer JH. Delis-Kaplan executive function scale. . *San Antonio, TX Psychol Corp*. 2001.
210. Rey A. L'examen psychologique dans les cas d'encéphalopathie traumatique. *Arch Psychol (Geneve)*. 1941.
211. Bates D, Mächler M, Bolker BM, Walker SC. Fitting linear mixed-effects models using lme4. *J Stat Softw*. 2015;67(1). doi:10.18637/jss.v067.i01

212. Gunda B, Porcher R, Duering M, et al. ADC histograms from routine DWI for longitudinal studies in cerebral small vessel disease: A field study in CADASIL. *PLoS One*. 2014. doi:10.1371/journal.pone.0097173
213. Cooley SA, Heaps JM, Bolzenius JD, et al. Longitudinal change in performance on the montreal cognitive assessment in older adults. *Clin Neuropsychol*. 2015. doi:10.1080/13854046.2015.1087596
214. Zeestraten EA, Benjamin P, Lambert C, et al. Application of diffusion tensor imaging parameters to detect change in longitudinal studies in cerebral small vessel disease. *PLoS One*. 2016;11(1):1-16. doi:10.1371/journal.pone.0147836
215. Tombaugh TN, McIntyre NJ. The Mini-Mental State Examination: A Comprehensive Review. *J Am Geriatr Soc*. 1992. doi:10.1111/j.1532-5415.1992.tb01992.x
216. Barberger-Gateau P, Commenges D, Gagnon M, Letenneur L, Sauvel C, Dartigues J -F. Instrumental Activities of Daily Living as a Screening Tool for Cognitive Impairment and Dementia in Elderly Community Dwellers. *J Am Geriatr Soc*. 1992. doi:10.1111/j.1532-5415.1992.tb01802.x
217. van Norden AGW, de Laat KF, Gons RAR, et al. Causes and consequences of cerebral small vessel disease. The RUN DMC study: A prospective cohort study. Study rationale and protocol. *BMC Neurol*. 2011;11(1):29. doi:10.1186/1471-2377-11-29
218. van Uden IWM, Tuladhar AM, van der Holst HM, et al. Diffusion tensor imaging of the hippocampus predicts the risk of dementia; the RUN DMC study. *Hum Brain Mapp*. 2016. doi:10.1002/hbm.23029
219. McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's Dement*. 2011. doi:10.1016/j.jalz.2011.03.005
220. Román GC, Tatemichi TK, Erkinjuntti T, et al. Vascular dementia: Diagnostic criteria for research studies: Report of the ninds-airen international workshop*. *Neurology*. 1993. doi:10.1212/wnl.43.2.250
221. Hilal S, Chai YL, Van Veluw S, et al. Association between subclinical cardiac biomarkers and clinically manifest cardiac diseases with cortical cerebral microinfarcts. *JAMA Neurol*. 2017. doi:10.1001/jamaneurol.2016.5335
222. Van Veluw SJ, Hilal S, Kuijf HJ, et al. Cortical microinfarcts on 3T MRI: Clinical correlates in memory-clinic patients. *Alzheimer's Dement*. 2015;11(12):1500-1509. doi:10.1016/j.jalz.2014.12.010
223. Seiler S, Pirpamer L, Hofer E, et al. Magnetization transfer ratio relates to cognitive impairment in normal elderly. *Front Aging Neurosci*. 2014. doi:10.3389/fnagi.2014.00263
224. Davies G, Lam M, Harris SE, et al. Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. *Nat Commun*. 2018. doi:10.1038/s41467-018-04362-x
225. Tombaugh TN. Trail Making Test A and B: Normative data stratified by age and education. *Arch Clin Neuropsychol*. 2004. doi:10.1016/S0887-6177(03)00039-8
226. Nagahama Y, Okina T, Suzuki N, et al. Factor structure of a modified version of the Wisconsin Card Sorting Test: An analysis of executive deficit in Alzheimer's disease and mild cognitive impairment. *Dement Geriatr Cogn Disord*. 2003. doi:10.1159/000070683
227. Wechsler D. Wechsler adult intelligence scale - Third Edition (WAIS-III). *San Antonio*. 1997.
228. Brand N. Information processing in depression and anxiety. *Psychol Med*.

1987. doi:10.1017/S0033291700013040
229. Van Der Elst W, Van Boxtel M, Van Breukelen G, Jolles J. The Letter Digit Substitution Test: Normative data for 1,858 healthy participants aged 24-81 from the Maastricht Aging Study (MAAS): Influence of age, education, and sex. *J Clin Exp Neuropsychol*. 2006. doi:10.1080/13803390591004428
 230. Osterrieth PA. Le test de copie d'une figure complexe; contribution à l'étude de la perception et de la mémoire. *Arch Psychol (Geneve)*. 1944.
 231. Jensen AR, Rohwer WD. The stroop color-word test: A review. *Acta Psychol (Amst)*. 1966. doi:10.1016/0001-6918(66)90004-7
 232. Stroop JR. Studies of interference in serial verbal reactions. *J Exp Psychol*. 1935. doi:10.1037/h0054651
 233. Van Der Elst W, Van Boxtel MPJ, Van Breukelen GJP, Jolles J. Normative data for the Animal, Profession and Letter M Naming verbal fluency tests for Dutch speaking participants and the effects of age, education, and sex. *J Int Neuropsychol Soc*. 2006. doi:10.1017/S1355617706060115
 234. R.K. M, N. C. Verbal Series Attention Test: Clinical utility in the assessment of dementia. *Clin Neuropsychol*. 1996.
 235. Dubois B, Slachevsky A, Litvan I, Pillon B. The FAB: A frontal assessment battery at bedside. *Neurology*. 2000. doi:10.1212/WNL.55.11.1621
 236. Worcester DA. The Porteus Maze Test and Intelligence. *Psychol Bull*. 1951. doi:10.1037/h0050707
 237. Mack WJ, Freed DM, Williams BW, Henderson VW. Boston Naming Test: Shortened versions for use in Alzheimer's disease. *Journals Gerontol*. 1992. doi:10.1093/geronj/47.3.P154
 238. Isaacs B, Kennie AT. The set test as an aid to the detection of dementia in old people. *Br J Psychiatry*. 1973. doi:10.1192/bjp.123.4.467
 239. Smith a. Symbol Digit Modalities Test (SDMT). *Neuropsychol Assess*. 2004.
 240. Diller L, Ben Yishay Y, Gerstman LJ. Studies in Cognition and Rehabilitation in Hemiplegia. *NEW YORK UNIVMEDCENT*. 1974.
 241. Wechsler D. Wechsler memory scale - Third edition administration and scoring manual. *San Antonio, TX Psychol Corp*. 1997.
 242. Sunderland T, Hill JL, Mellow AM, et al. Clock Drawing in Alzheimer's Disease: A Novel Measure of Dementia Severity. *J Am Geriatr Soc*. 1989. doi:10.1111/j.1532-5415.1989.tb02233.x
 243. Wechsler D. Wechsler Adult Intelligence Scale (WAIS-3R). *Psychol Corp*. 1997.
 244. Sahadevan S, Lim PPJ, Tan NJL, Chan SP. Diagnostic performance of two mental status tests in the older Chinese: Influence of education and age on cut-off values. *Int J Geriatr Psychiatry*. 2000. doi:10.1002/(SICI)1099-1166(200003)15:3<234::AID-GPS99>3.0.CO;2-G
 245. Lambert C, Zeestraten E, Williams O, et al. Identifying preclinical vascular dementia in symptomatic small vessel disease using MRI. *NeuroImage Clin*. 2018;19(January):925-938. doi:10.1016/j.nicl.2018.06.023
 246. Ghafoorian M, Karssemeijer N, Van Uden IWM, et al. Automated detection of white matter hyperintensities of all sizes in cerebral small vessel disease. *Med Phys*. 2016. doi:10.1118/1.4966029
 247. Vrooman HA, Cocosco CA, van der Lijn F, et al. Multi-spectral brain tissue segmentation using automatically trained k-Nearest-Neighbor classification. *Neuroimage*. 2007. doi:10.1016/j.neuroimage.2007.05.018
 248. de Boer R, Vrooman HA, Ikram MA, et al. Accuracy and reproducibility study of automatic MRI brain tissue segmentation methods. *Neuroimage*. 2010. doi:10.1016/j.neuroimage.2010.03.012
 249. Gyanwali B, Shaik MA, Tan BY, Venketasubramanian N, Chen C, Hilal S. Risk Factors for and Clinical Relevance of Incident and Progression of Cerebral Small Vessel Disease Markers in an Asian Memory Clinic Population. *J*

- Alzheimer's Dis.* 2019. doi:10.3233/JAD-180911
250. Zwiers MP. Patching cardiac and head motion artefacts in diffusion-weighted images. *Neuroimage*. 2010. doi:10.1016/j.neuroimage.2010.06.014
 251. Grömping U. Relative importance for linear regression in R: The package relaimpo. *J Stat Softw*. 2006. doi:10.18637/jss.v017.i01
 252. Field A. *Discovering Statistics Using IBM SPSS Statistics.*; 2013.
 253. Robin X, Turck N, Hainard A, et al. pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 2011. doi:10.1186/1471-2105-12-77
 254. Stern Y. Cognitive reserve in ageing and Alzheimer's disease. *Lancet Neurol*. 2012. doi:10.1016/S1474-4422(12)70191-6
 255. Opherk C, Peters N, Herzog J, Luedtke R, Dichgans M. Long-term prognosis and causes of death in CADASIL: A retrospective study in 411 patients. *Brain*. 2004. doi:10.1093/brain/awh282
 256. Lawrence AJ, Chung AW, Morris RG, Markus HS, Barrick TR. Structural network efficiency is associated with cognitive impairment in small-vessel disease. *Neurology*. 2014;83(4):304-311. doi:10.1212/WNL.0000000000000612
 257. Tuladhar AM, Van Uden IWM, Rutten-Jacobs LCA, et al. Structural network efficiency predicts conversion to dementia. *Neurology*. 2016;86(12):1112-1119. doi:10.1212/WNL.0000000000002502
 258. Lawrence AJ, Zeestraten EA, Benjamin P, et al. Longitudinal decline in structural networks predicts dementia in cerebral small vessel disease. *Neurology*. 2018. doi:10.1212/WNL.0000000000005551
 259. Wen Q, Mustafi SM, Li J, et al. White matter alterations in early-stage Alzheimer's disease: A tract-specific study. *Alzheimer's Dement Diagnosis, Assess Dis Monit*. 2019. doi:10.1016/j.dadm.2019.06.003
 260. Gold BT, Johnson NF, Powell DK, Smith CD. White matter integrity and vulnerability to Alzheimer's disease: Preliminary findings and future directions. *Biochim Biophys Acta - Mol Basis Dis*. 2012. doi:10.1016/j.bbadis.2011.07.009
 261. Williams OA, Zeestraten EA, Benjamin P, 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. doi:10.1016/j.nicl.2017.08.016
 262. Smith SM, Jenkinson M, Johansen-Berg H, et al. Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. *Neuroimage*. 2006. doi:10.1016/j.neuroimage.2006.02.024
 263. Low A, Mak E, Stefaniak JD, et al. Peak Width of Skeletonized Mean Diffusivity as a Marker of Diffuse Cerebrovascular Damage. *Front Neurosci*. 2020. doi:10.3389/fnins.2020.00238
 264. Beaudet G, Tsuchida A, Petit L, et al. Age-Related Changes of Peak Width Skeletonized Mean Diffusivity (PSMD) Across the Adult Lifespan: A Multi-Cohort Study. *Front Psychiatry*. 2020. doi:10.3389/fpsy.2020.00342
 265. Lam BYK, Leung KT, Yiu B, et al. Peak width of skeletonized mean diffusivity and its association with age-related cognitive alterations and vascular risk factors. *Alzheimer's Dement Diagnosis, Assess Dis Monit*. 2019. doi:10.1016/j.dadm.2019.09.003
 266. McCreary CR, Beaudin AE, Subotic A, et al. Cross-sectional and longitudinal differences in peak skeletonized white matter mean diffusivity in cerebral amyloid angiopathy. *NeuroImage Clin*. 2020. doi:10.1016/j.nicl.2020.102280
 267. Vinciguerra C, Giorgio A, Zhang J, 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. doi:10.1016/j.msard.2018.11.011
 268. Williams OA, Zeestraten EA, Benjamin P, et al. Predicting Dementia in

- Cerebral Small Vessel Disease Using an Automatic Diffusion Tensor Image Segmentation Technique. *Stroke*. 2019;50(10):2775-2782. doi:10.1161/STROKEAHA.119.025843
269. Tuladhar AM, Lawrence A, Norris DG, Barrick TR, Markus HS, de Leeuw FE. Disruption of rich club organisation in cerebral small vessel disease. *Hum Brain Mapp*. 2017. doi:10.1002/hbm.23479
 270. Reijmer YD, Fotiadis P, Martinez-Ramirez S, et al. Structural network alterations and neurological dysfunction in cerebral amyloid angiopathy. *Brain*. 2015. doi:10.1093/brain/awu316
 271. Tuladhar AM, van Dijk E, Zwiers MP, et al. Structural network connectivity and cognition in cerebral small vessel disease. *Hum Brain Mapp*. 2016. doi:10.1002/hbm.23032
 272. Song SK, Sun SW, Ramsbottom MJ, Chang C, Russell J, Cross AH. Dysmyelination revealed through MRI as increased radial (but unchanged axial) diffusion of water. *Neuroimage*. 2002. doi:10.1006/nimg.2002.1267
 273. Winklewski PJ, Sabisz A, Naumczyk P, Jodzio K, Szurowska E, Szarmach A. Understanding the physiopathology behind axial and radial diffusivity changes- what do we know? *Front Neurol*. 2018. doi:10.3389/fneur.2018.00092
 274. Smith SM. Fast robust automated brain extraction. *Hum Brain Mapp*. 2002. doi:10.1002/hbm.10062
 275. Avants B, Tustison N, Song G. Advanced Normalization Tools (ANTs). *Insight J*. 2009.
 276. Pierpaoli C, Basser PJ. Toward a quantitative assessment of diffusion anisotropy. *Magn Reson Med*. 1996. doi:10.1002/mrm.1910360612
 277. Bullmore E, Sporns O. Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci*. 2009. doi:10.1038/nrn2575
 278. Desikan RS, Ségonne F, Fischl B, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*. 2006. doi:10.1016/j.neuroimage.2006.01.021
 279. Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM. Review FSL. *Neuroimage*. 2012.
 280. Tournier JD, Calamante F, Connelly A. MRtrix: Diffusion tractography in crossing fiber regions. *Int J Imaging Syst Technol*. 2012. doi:10.1002/ima.22005
 281. Latora V, Marchiori M. Efficient behavior of small-world networks. *Phys Rev Lett*. 2001. doi:10.1103/PhysRevLett.87.198701
 282. Newman MEJ. The structure and function of complex networks. *SIAM Rev*. 2003;45(2):167-256. doi:10.1137/S003614450342480
 283. Donohue MC, Gamst AC, Edland SD, Donohue MMC. Package 'longpower.' *Biometrics*. 2013.
 284. Chen D-G (Din), Peace KE. *Clinical Trial Data Analysis Using R.*; 2010. doi:10.1201/b10478
 285. Champely S. pwr: Basic functions for power analysis. *R Packag version 12-1*. 2017.
 286. De Guio F, Jouvent E, Biessels GJ, et al. Reproducibility and variability of quantitative magnetic resonance imaging markers in cerebral small vessel disease. *J Cereb Blood Flow Metab*. 2016. doi:10.1177/0271678X16647396
 287. De Guio F, Duering M, Fazekas F, 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(2):231-245. doi:10.1177/0271678X19888967
 288. Bernard C, Helmer C, Dilharreguy B, et al. Time course of brain volume changes in the preclinical phase of Alzheimer's disease. *Alzheimer's Dement*. 2014. doi:10.1016/j.jalz.2013.08.279

289. Winterer G, Androsova G, Bender O, et al. Personalized risk prediction of postoperative cognitive impairment – rationale for the EU-funded BioCog project. *Eur Psychiatry*. 2018. doi:10.1016/j.eurpsy.2017.10.004
290. Veenith T V., Carter E, Grossac J, et al. Inter Subject Variability and Reproducibility of Diffusion Tensor Imaging within and between Different Imaging Sessions. *PLoS One*. 2013. doi:10.1371/journal.pone.0065941
291. Lawrence AJ, Tozer DJ, Stamatakis EA, Markus HS. A comparison of functional and tractography based networks in cerebral small vessel disease. *NeuroImage Clin*. 2018. doi:10.1016/j.nicl.2018.02.013
292. McGraw KO, Wong SP. Forming Inferences about Some Intraclass Correlation Coefficients. *Psychol Methods*. 1996. doi:10.1037/1082-989X.1.1.30
293. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med*. 2016. doi:10.1016/j.jcm.2016.02.012
294. Ajilore O, Lamar M, Kumar A. Association of brain network efficiency with aging, depression, and cognition. *Am J Geriatr Psychiatry*. 2014. doi:10.1016/j.jagp.2013.10.004
295. Fischer FU, Wolf D, Scheurich A, Fellgiebel A. Association of structural global brain network properties with intelligence in normal aging. *PLoS One*. 2014;9(1). doi:10.1371/journal.pone.0086258
296. Koenis MMG, Brouwer RM, van den Heuvel MP, et al. Development of the brain's structural network efficiency in early adolescence: A longitudinal DTI twin study. *Hum Brain Mapp*. 2015;36(12):4938-4953. doi:10.1002/hbm.22988
297. Marengo S, Rawlings R, Rohde GK, et al. Regional distribution of measurement error in diffusion tensor imaging. *Psychiatry Res - Neuroimaging*. 2006;147(1):69-78. doi:10.1016/j.psychres.2006.01.008
298. Albi A, Pasternak O, Minati L, et al. Free water elimination improves test-retest reproducibility of diffusion tensor imaging indices in the brain: A longitudinal multisite study of healthy elderly subjects. *Hum Brain Mapp*. 2017;38(1):12-26. doi:10.1002/hbm.23350
299. Liu X, Yang Y, Sun J, et al. Reproducibility of diffusion tensor imaging in normal subjects: An evaluation of different gradient sampling schemes and registration algorithm. *Neuroradiology*. 2014;56(6):497-510. doi:10.1007/s00234-014-1342-2
300. Pfefferbaum A, Adalsteinsson E, Sullivan E V. Replicability of diffusion tensor imaging measurements of fractional anisotropy and trace in brain. *J Magn Reson Imaging*. 2003;18(4):427-433. doi:10.1002/jmri.10377
301. Luque Laguna PA, Combes AJE, Streffer J, et al. Reproducibility, reliability and variability of FA and MD in the older healthy population: A test-retest multiparametric analysis. *NeuroImage Clin*. 2020;26(July 2019). doi:10.1016/j.nicl.2020.102168
302. Croall ID, Lohner V, Moynihan B, et al. Using DTI to assess white matter microstructure in cerebral small vessel disease (SVD) in multicentre studies. *Clin Sci*. 2017. doi:10.1042/CS20170146
303. Galluzzi S, Marizzoni M, Babiloni C, et al. Clinical and biomarker profiling of prodromal Alzheimer's disease in workpackage 5 of the Innovative Medicines Initiative PharmaCog project: A "European ADNI study." *J Intern Med*. 2016;279(6):576-591. doi:10.1111/joim.12482
304. Zivadinov R, Bergsland N, Hagemeyer J, et al. Effect of switching from glatiramer acetate 20 mg/daily to glatiramer acetate 40 mg three times a week on gray and white matter pathology in subjects with relapsing multiple sclerosis: A longitudinal DTI study. *J Neurol Sci*. 2018;387(October 2017):152-156. doi:10.1016/j.jns.2018.02.023
305. Gaetani L, Blennow K, Calabresi P, Di Filippo M, Parnetti L, Zetterberg H. Neurofilament light chain as a biomarker in neurological disorders. *J Neurol*

- Neurosurg Psychiatry*. 2019;870-881. doi:10.1136/jnnp-2018-320106
306. Gattringer T, Pinter D, Enzinger C, et al. Serum neurofilament light is sensitive to active cerebral small vessel disease. *Neurology*. 2017. doi:10.1212/WNL.0000000000004645
 307. Tiedt S, Duering M, Barro C, et al. Serum neurofilament light a biomarker of neuroaxonal injury after ischemic stroke. *Neurology*. 2018;91(14):E1338-E1347. doi:10.1212/WNL.0000000000006282
 308. Gravesteijn G, Rutten JW, Verberk IMW, et al. Serum Neurofilament light correlates with CADASIL disease severity and survival. *Ann Clin Transl Neurol*. 2019;6(1):46-56. doi:10.1002/acn3.678
 309. Peters N, Leijssen E Van, Tuladhar AM, Barro C, Konieczny MJ. Serum Neurofilament Light Chain Is Associated with Incident Lacunes in Progressive Cerebral Small Vessel Disease. 2020;22(3):369-376.
 310. Disanto G, Barro C, Benkert P, et al. Serum Neurofilament light: A biomarker of neuronal damage in multiple sclerosis. *Ann Neurol*. 2017. doi:10.1002/ana.24954
 311. Rankin J. Cerebral Vascular Accidents in Patients over the Age of 60: II. Prognosis. *Scott Med J*. 1957. doi:10.1177/003693305700200504
 312. Van Swieten JC, Koudstaal PJ, Visser MC, Schouten H, Van Gijn J. Interobserver agreement for the assessment of handicap in stroke patients. *Stroke*. 1988. doi:10.1161/01.STR.19.5.604
 313. Brookes RL, Hollocks MJ, Khan U, Morris RG, Markus HS. The Brief Memory and Executive Test (BMET) for detecting vascular cognitive impairment in small vessel disease: A validation study. *BMC Med*. 2015. doi:10.1186/s12916-015-0290-y