Difference in distribution functions

Du, Jing; Koch, Forrest C.; Xia, Aihua; Jiang, Jiyang; Crawford, John D.; Lam, Ben C. P.

Published in:
NeuroImage

DOI:
10.1016/j.neuroimage.2021.118381

Publication date:
2021

NeuroImage 240 (2021) 118381

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Difference in distribution functions: A new diffusion weighted imaging metric for estimating white matter integrity

Jing Du[a,1,*], Forrest C. Koch[a,1], Aihua Xia[c], Jiyang Jiang[a], John D. Crawford[a], Ben C.P. Lam[a], Anbupalam Thalamuthu[a], Teresa Lee[a,b], Nicole Kochan[a], Chloe Fawns-Ritchie[d], Henry Brodaty[a,f], Qun Xu[e,g,*], Perminder S. Sachdev[a,b], Wei Wen[a,b]

[a] Centre for Healthy Brain Aging (CHeBA), School of Psychiatry, UNSW Sydney, New South Wales 2052, Australia
[b] Neuropsychiatric Institute (NPI), Euroa Centre, Prince of Wales Hospital, Randwick, New South Wales 2031, Australia
[c] School of Mathematics and Statistics, University of Melbourne, Melbourne, Victoria 3010, Australia
[d] Centre for Cognitive Ageing and Cognitive Epidemiology (CCACE), Department of Psychology, University of Edinburgh, Edinburgh, UK
[e] Department of Health Manage Centre, RenJi Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai 200127, China
[f] Dementia Centre for Research Collaboration, School of Psychiatry, UNSW Sydney, New South Wales 2052, Australia
[g] Department of Neurology, RenJi Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai 200127, China

ARTICLE INFO

Keywords:
Diffusion weighted imaging
White matter
Ageing
Cognition

ABSTRACT

Diffusion weighted imaging (DWI) is a widely recognized neuroimaging technique to evaluate the microstructure of brain white matter. The objective of this study is to establish an improved automated DWI marker for estimating white matter integrity and investigating ageing related cognitive decline. The concept of Wasserstein distance was introduced to help establish a new measure: difference in distribution functions (DDF), which captures the difference of reshaping one's mean diffusivity (MD) distribution to a reference MD distribution. This new DWI measure was developed using a population-based cohort (n=19,369) from the UK Biobank. Validation was conducted using the data drawn from two independent cohorts: the Sydney Memory and Ageing Study, a community-dwelling sample (n=402), and the Renji Cerebral Small Vessel Disease Cohort Study (RCCS), which consisted of cerebral small vessel disease (CSVD) patients (n=171) and cognitively normal controls (NC) (n=43). DDF was associated with age across all three samples and better explained the variance of changes than other established DWI measures, such as fractional anisotropy, mean diffusivity and peak width of skeletonized mean diffusivity (PSMD). Significant correlations between DDF and cognition were found in the UK Biobank cohort and the MAS cohort. Binary logistic analysis and receiver operator characteristic curve analysis of RCCS demonstrated that DDF had higher sensitivity in distinguishing CSVD patients from NC than the other DWI measures. To demonstrate the flexibility of DDF, we calculated regional DDF which also showed significant correlation with age and cognition. DDF can be used as a marker for monitoring the white matter microstructural changes and ageing related cognitive decline in the elderly.

1. Introduction

Diffusion weighted imaging (DWI) is a non-invasive imaging technique widely used to investigate the microstructure of cerebral white matter *in vivo* (Terence et al., 2008). White matter integrity is critical to normal brain structure and function, and its disruption may accelerate brain ageing as well as brain diseases. Various DWI measures have been developed for investigation of white matter, with notable examples of FA (fractional anisotropy) and MD (mean diffusivity). DWI measures has been reported associated with microstructural changes across the lifespan (Mwangi et al., 2013). FA and MD (Basser and Pierpaoli, 1996) are the two most commonly used DWI indexes to depict the directionality and magnitude of diffusion in the cerebral white matter. Changes of FA and MD have also been widely observed to be associated with cognitive decline in the healthy elderly (Abe et al., 2002; Grieve et al., 2007; Lebel et al., 2012) and in populations affected by small vessel disease or Alzheimer's disease (Croall et al., 2017; Li et al., 2018).

Efforts have been made to develop DWI-derived metrics that are sensitive to changes in specific cognitive domains or specific neuropathologies. Peak width of skeletonized mean diffusivity (PSMD) is a novel measure introduced in 2016 (Baykara et al., 2016). PSMD has been extensively used to investigate cerebral small vessel disease (CSVD), and is defined as the difference between the 95th and 5th percentiles

* Corresponding authors.
E-mail addresses: jing.du@student.unsw.edu.au (J. Du), xuqun@renji.com (Q. Xu).
[1] Co-first authors.

https://doi.org/10.1016/j.neuroimage.2021.118381.
Received 19 May 2021; Received in revised form 27 June 2021; Accepted 8 July 2021
Available online 9 July 2021.

of skeletonized MD. It has been found to be highly correlated with processing speed, a cognitive domain most detrimentally affected by CSVD (Lam et al., 2019). However, if two individuals have similar level of variances in their MD distributions, their PSMD would not be much different, even if the locations of the modes (the peaks) of the two distributions are very different. This is because PSMD captures only the variability of MD distribution without accounting for the shift of the distribution.

The aim of our study, therefore, was to develop a new DWI metric which would capture the comprehensive profile of MD distribution of the whole brain white matter by considering the accumulative shift of MD of each voxel. This new measure should meet two major criteria: able to capture the change of age, and the change in cognition. Furthermore, it should be able to differentiate diseased brains from healthy ones. Our new DWI metric is a special case under the general framework which was developed in this study. We name this general framework the difference in distribution function (DDF). Under this framework, in addition to MD, other DWI indexes such as FA, AD (axial diffusivity) and RD (radial diffusivity) were also computed and investigated in Supplementary.

Three independent cohorts were used to develop and validate the new framework and measures. As a discovery dataset, we used UK Biobank, a large-scale cohort study with comprehensive data on lifestyle, physical and mental health, brain MRI and cognition (Collins, 2012). We investigated the associations between DDF and age and cognition in this cohort. Reliability and predictive validity of DDF were examined using two independent cohorts: the Sydney Memory and Ageing Study (MAS) (Sachdev et al., 2010) and the Renji Cerebral Small Vessel Disease Cohort Study (RCCS) (Du et al., 2019).

2. Materials and methods

2.1. Participants

2.1.1. UK Biobank

UK Biobank (www.ukbiobank.ac.uk) (Sudlow et al., 2015) is a large-scale ongoing biomedical study which began in 2006 with a focus on investigating the contributions of genetic and environmental influences on ageing and diseases. The UK Biobank project received ethical approval from the Research Ethics Committee within the terms of an Ethics and Governance Framework and all participants provided written informed consent. We excluded seven participants with dementia, resulting in 19,369 participants being included in our study since the release date.

As shown in supplementary Table e-1, educational attainment of UK Biobank participants was measured in seven categories, which we binarized into either having attained a college/university degree or not (1 = college/university degree or above, 0 = otherwise). The age of a participant was the age at the MRI scan.

2.1.1.1. MRI acquisition. Brain imaging data from the UK Biobank cohort were obtained from three dedicated imaging centres (Manchester, Reading and Newcastle). The scanners had the same system, parameters and manufacturer (3T Siemens Skyra, software VD13) and a standard Siemens 32-channel receive head coil. We used T1-weighted, FLAIR (fluid-attenuated inversion recovery) and DWI scans. The parameters for each MRI modality can be found in previous publications (Alfaro-Almagro et al., 2018; Miller et al., 2016) of UK Biobank.

2.1.1.2. Cognition. Cognitive tests were administered on a touchscreen computer. The short-term test-retest reliability and validity of the cognitive tests (as correlations with other well-validated, standardised cognitive tests) have been reported to be moderate to high, with mean Pearson $r = 0.55$, ranging from 0.40 to 0.89, and mean Pearson $r = 0.53$, ranging from 0.22 to 0.83, respectively (Fawns-Ritchie and Deary, 2020).

Cognitive test data were acquired at the imaging visit (instance 2). Seven tests from the UK Biobank battery of cognitive tests were selected for the current study to represent three cognitive domains (Cox et al., 2019; Kendall et al., 2017). “Reaction Time” (average time to correctly identify matches in a “snap”-like card game task), “Trail Making Test A” (time taken to complete a numeric path), and “Symbol Digit Substitution” (number of correct symbol number matches within time limit) formed the *Processing Speed domain*. “Numeric Memory” (maximum number of digits remembered correctly) and “Pairs Matching” (number of incorrect visual matching) contributed to the *Memory domain*. “Trail Making Test B” (time taken to complete an alphanumeric path) and “Fluid Intelligence” (total number of questions that required logic and reasoning correctly answered) formed the *Executive Function domain*.

2.1.2. MAS

The first validation sample was drawn from Wave 2 of the community-dwelling population-based Sydney Memory and Ageing Study (MAS) (Sachdev et al., 2010), which is a longitudinal study examining the predictors of cognitive decline in a nondemented, community-dwelling sample. They were recruited randomly through the electoral roll from two electorates of Eastern Sydney, New South Wales, Australia, where registration on the electoral roll is compulsory. Participants with dementia were excluded from the study at baseline. MRI scans were available for Waves 1, 2 and 4 for about half of the participants in each wave. We used the data of Wave 2 in our study as there were no DWI data available in Wave 1 and Wave 4 had fewer scans than Wave 2. Participants with dementia at Wave 2 were excluded (final $n = 402$). The study was approved by the Human Research Ethics Committees of the University of New South Wales (UNSW Sydney) and the South Eastern Sydney and Illawarra Area Health Service (SESLHD). Written informed consent was obtained from participants and informants.

2.1.2.1. MRI acquisition. The brain MRI data of Wave 2 were acquired on Philips 3T Achieva Quasar Dual scanner (Philips Medical Systems, Best, The Netherlands) at Neuroscience Research Australia (NeuRA), Sydney, New South Wales, Australia. T1-weighted, FLAIR and DWI scans were used in this study. The standard protocol can be found in our previous publications (Jiang et al., 2015).

2.1.2.2. Cognition. A battery of interview-based cognitive tests covering five cognitive domains were administered by trained psychologists. These tests (Sachdev et al., 2010) include Attention/processing speed: Digit Symbol-Coding, Trail Making test A; Executive function: Controlled Oral Word Association Test, Trail Making Test B; Memory: Logical Memory Story A delayed recall, Rey Auditory Verbal Learning Test, Benton Visual Retention Test recognition; Language: Boston Naming Test – 30 items, Semantic Fluency (Animals); Visuo-spatial function: Block Design.

2.1.3. RCCS

The second validation cohort is Renji Cerebral Small Vessel Disease Corhort Study (RCCS), a cohort study of post-stroke cerebral small vessel disease (CSVD) patients with varying levels of cognitive performance recruited consecutively by the Renji Hospital, Shanghai Jiao Tong University School of Medicine. Written informed consent was obtained from participants and informants and the study was approved by the Research Ethics Committee of Renji Hospital School of Medicine, Shanghai Jiao Tong University (Du et al., 2019). This study included detailed cognitive tests across different domains and multi-modal MRI sequences, including structural MRI, DWI and functional MRI. Forty-three age and sex matched healthy community-dwelling participants without CSVD were recruited from the Tangqiao community, Pudong New District in Shanghai as normal controls (NC). Detailed recruitment and exclusion criteria for CSVD patients have been published elsewhere (Du et al., 2019). The following inclusion criteria were applied for the NC: (1) education ≥ 6 years; (2) no history of clinical stroke; (3) no history of severe diseases of important organs including liver, heart, lung; (4) visual rating score

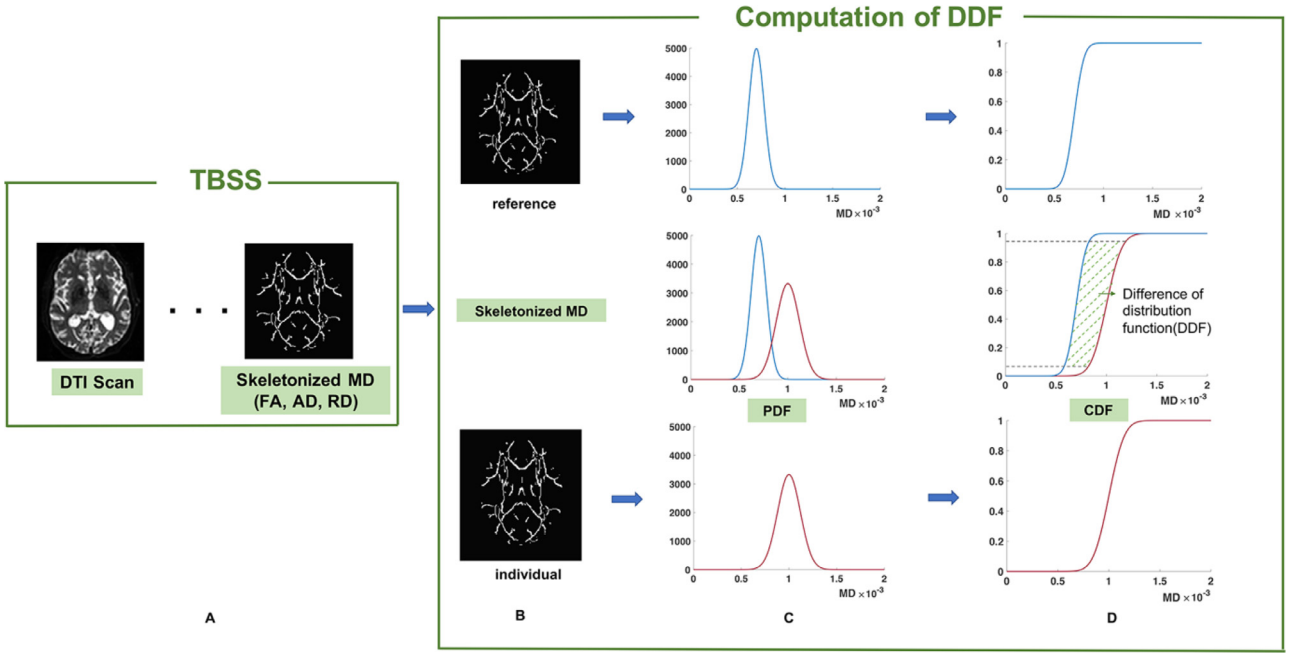

Fig. 1. Flowchart of DDF computation. A: Generating skeletonized MD map (TBSS). B-D: The computation of DDF: Column C shows the PDF for skeletonized MD distribution. Column D shows the CDF for skeletonized MD distribution. The upper row is the reference, the lower row is an individual subject, and row in the middle shows the comparison between individual and reference. The value of DDF is the area with green dash lines. The grey dash lines are the 5th to 95th percentiles of the differences of the inversed distribution functions. Abbreviations: DDF, Difference in distribution functions; TBSS, Tract-Based Spatial Statistics; MD, mean diffusivity; PDF, probability density function; CDF, cumulative distribution function.

of WMH graded using the Fazekas score (Fazekas et al., 1987) in T2 FLAIR sequence ≤ 1 and without other obvious structural abnormalities on MRI scans; (5) no severe depressive symptoms (17-item Hamilton Depression Rating Scale score (Hamilton, 1960) ≥ 24); (6) cognitive tests scores were within normal range; (7) no vascular risks factors such as hypercholesterolemia, diabetes mellitus, hypertension and smoking.

2.1.3.1. MRI acquisition. Standardized T1-weighted 3D fast spoiled gradient recalled (SPGR) sequence images, T2-weighted FLAIR, and DWI were obtained using a 3T MRI scanner (Signa HDxt; GE HealthCare, Milwaukee, WI, USA). Detailed parameters can be found in previous publication (Du et al., 2019).

2.1.3.2. Cognitive testing. Standardized detailed multi-domain cognitive tests were administered by a trained neuropsychologist. Processing speed was assessed by Trail Making Test A, Stroop colour-word test A (Stroop, 1935) and Symbol digital substitution. Executive function was assessed by Trail Making Test B and Stroop colour-word test C. Memory function was evaluated by auditory verbal learning test (short term and long term recall) (AVLT) (Qihao Guo et al., 2009). Language function was assessed by semantic verbal fluency test (VFT, 1-minute animal naming test) (Benton, 1989) and Boston naming test (BNT) (Guo et al., 2006). Spatial function was assessed by Rey-Osterrieth complex figure test (Rey-O copy) (Shin et al., 2006).

2.2. Development of DDF – a general framework using W-index

Definition: DDF computation is illustrated in Fig. 1. Inspired by the concept of Wasserstein distance (WD) (Villani, 2003) we began by considering two cumulative distribution functions (DF) with F_R for the DF of some reference distribution of a random variable R and F_S for the target distribution of a random variable S. F_R can be the distribution of a group of healthy controls or an entire population, and F_S is the DF of a DWI metric of a participant. The framework we propose is then in the form of:

$$d(F_R, F_S) = \int_0^1 \phi(x,\ F_R^{-1}(x) - F_S^{-1}(x))dx,$$

where ϕ is a weighting function applied to the differences of the inverse DFs. The inverse DF gives the quantile, for example, $F_R^{-1}(0.5) - F_S^{-1}(0.5)$ is the difference between the medians of the two distributions. Hence $d(F_R,\ F_S)$ measures the accumulated weighted differences of the corresponding quantiles. Furthermore, the use of inverse DFs allows the integral to be defined cleanly in terms of quantiles as opposed to specific values or units of measurement. As $d(F_R, F_S)$ captures the differences between F_R and F_S using the mass transportation idea in the Wasserstein distance, we call it *W-index* of S in terms of the reference R. Ideally, if we know the benchmark distribution F_R, the W-index would be able to tell the health status of the brain S. It is, however, in reality difficult to find the perfect benchmark distribution F_R because it would be impossible to define a perfectly healthy brain.

It is worth noting the case when ϕ is the identity function, that is $\phi(x, y) = y$ for $l \le x \le u$ and $\phi(x, y) = 0$ for $x < l$ or $x > u$, we have

$$\begin{aligned} d(F_R, F_S) &= \int_l^u \left(F_R^{-1}(x) - F_S^{-1}(x)\right)dx = \int_l^u F_R^{-1}(x)dx - \int_l^u F_S^{-1}(x)dx \\ &= \mathbb{E}\left(R \times \mathbf{1}_{\left[F_R^{-1}(l) < R < F_R^{-1}(u)\right]}\right) - \mathbb{E}\left(S \times \mathbf{1}_{\left[F_S^{-1}(l) < S < F_S^{-1}(u)\right]}\right), \end{aligned}$$

where $\mathbf{1}_{[a<R<b]}$ is an *indicator function* which takes value 1 if random variable R is between a and b; otherwise, it is 0. In other words, $\mathbb{E}(R \times \mathbf{1}_{[a<R<b]})$ is called the *trimmed mean* of R for its values between a and b. Hence, when $\phi(x, y) = y$ for $l \le x \le u$ and $\phi(x, y) = 0$ for $x < l$ or $x > u$, $d(F_R,\ F_S)$ is simply the difference between the truncated means of the two distributions. In this case, when comparing DWI measures of subjects evaluated against the same reference, the truncated mean of the reference distribution can be treated as constant and thus no longer required in the computation.

We point out that PSMD is a special case of our framework. When there is no reference, i.e., $F_R^{-1}(x) = 0$ for all x, and $\phi(x, y) = \frac{1}{f_S(y)}$ for

$0.05 \le x \le 0.95$ and $\phi(x, y) = 0$ for $x < 0.05$ or $x > 0.95$, where f_S is the density of F_S, then

$$PSMD = \int_{0.05}^{0.95} \phi\left(x, F_S^{-1}(x)\right) dx = \int_{0.05}^{0.95} dF_S^{-1}(x) = F_S^{-1}(0.95) - F_S^{-1}(0.05).$$

Implementation: A software package, (https://github.com/ForrestCKoch/DCDF) available through the Python Package Index (https://pypi.org/project/dcdf/) (PyPI) was developed to perform numerical integration using Reimann sums over user specified functions (ϕ). These functions can be defined using Python 3 syntax and any functions available through the Numpy library (https://numpy.org/). The reference distribution is first estimated by calculating the running average of cumulative bin frequencies over a user supplied reference list of images. DDF framework based DWI metrics of the subjects were then computed in parallel against the generated reference. DDF references were computed specifically for each different cohort. One hundred participants were randomly selected from UK Biobank for the calculation of reference distribution F_R, we conducted the visual inspection of their images to ensure that there were no severe brain lesions. For MAS cohort, all MAS Wave 2 participants were included as the reference distribution. In RCCS, we used all the normal controls as the reference.

Preprocessing: DWI data were first skeletonized using Tract-Based Spatial Statistics (TBSS), part of the FSL (Smith et al., 2006). AD, FA, MD, and RD images were projected onto the resulting skeleton using FA derived projection parameters. These images were further masked using a standard skeleton threshold at an FA of 0.3 and a mask provided with the PSMD tool to exclude regions adjacent to the ventricles (Baykara et al., 2016). Mean MD (AD or FA or RD) values were calculated by averaging each individual's skeletonized MD data (AD or FA or RD).

Data exploration: We conducted an initial search of functional forms to find measures which correlated well with both age and cognition in the MAS cohort. For each of the AD, FA, MD, and RD map, we calculated measures for Eqs. (1)–(3).

$$\int_0^1 \phi\left(\, x,\ F_R^{-1}(x) - F_S^{-1}(x)\right) dx, \tag{1}$$

with

$$\phi(x, y) = y \text{ for } l \le x \le u \text{ and } \phi(x, y) = 0 \text{ for } x < l \text{ or } x > u, \tag{2}$$

or

$$\phi(x, y) = e^{\theta y} \text{ for } l \le x \le u \text{ and } \phi(x, y) = 0 \text{ for } x < l \text{ or } x > u. \tag{3}$$

If we use Eq. (2), it is the simplest form without taking its absolute value. In most of the cases, the difference between the DWI measure distributions of a reference and a subject, i.e. $D(x) = F_R^{-1}(x) - F_S^{-1}(x)$, is very small. Therefore, for Eq. (3), we considered $\theta = 10^i, i \in \{0,\ 1,\ 2, 3\}$. Here we used $(l,\ u) = (0.05, 0.95)$, as we considered the DWI values between (0, 0.05) and (0.95, 1.0) as outliers or noises. Top performing measures were then selected for UK Biobank, MAS cohort and RCCS cohort. For $(l, u) = (0.05, 0.95)$, we applied Eqs. (1) and (3) with $\theta = 10$ $(i = 1)$, 100 $(i = 2)$, and 1000 $(i = 3)$. In this study we focused on MD using Eqs. (1) and (3) with $\theta = 1000$ $(i = 3)$, $(l, u) = (0.05, 0.95)$, which had best performance correlating with age and cognition and was denoted as $DDF^{MD}_{1000,\ 0.05,0.95}$.

ROI analysis: In addition to TBSS analysis for the whole brain, region of interest (ROI) analysis was conducted to evaluate the regional DDF for local anatomical information. As an example, we used JHU white matter tractography atlas (probability threshold at 0.25) (Wakana et al., 2007) and computed DDF for the eleven (we combined right and left of the same structure) white matter fibre tracts. An illustration of JHU atlas was shown in Supplementary Fig. e-1, detailed information for the tracts can be found in Supplementary Table e-2.

2.3. Computation of white matter hyperintensity (WMH) volume

WMH volumes were extracted by an automated pipeline UBO Detector (Jiang et al., 2018) for all three cohorts. For MAS and RCCS participants, WMH volumes were extracted by an automated pipeline UBO Detector. Briefly, individual FLAIR images were warped to a standard space, and segmented into clusters using FSL FAST. A k-nearest neighbour (k-NN)-based algorithm was then applied to the clusters for the classification of WMH and non-WMH. Periventricular WMH was defined as WMH voxels within 12 mm from a periventricular mask in standard space, and the rest WMH voxels were regarded as deep WMH.

UBO Detector generated satisfying WMH segmentations for older participants (e.g. MAS) and individuals with cerebrovascular pathology (e.g. RCCS). However, we found UBO Detector did not perform satisfactorily in scans with very subtle WMHs, such as the younger participants in UK Biobank, because subtle WMH burden tended to disappear after the interpolation in spatial transformation. Therefore, we modified the UBO Detector to extract WMH in the individual FLAIR space. All atlases and masks in the standard space were transformed to individual space by applying the reversed flow map. Intracranial volume (ICV) was included as a covariate when analysing WMH volume in UK Biobank.

2.4. Computation of cognitive domain and global cognition scores

All the raw test scores were first transformed into z-scores using the mean and standard deviation values of a healthy reference subsample in each study following previous practice (Lipnicki et al., 2013). Transformation was used to reduce skewness if the absolute value was larger than 1 and outliers using ±3 standard deviation (SD) were winsorized. The z-scores of corresponding tests were averaged to form domain scores and were again standardized using the means and SDs of the reference sample. Global cognition scores were computed in a similar way by averaging the domain scores and subsequently transforming to z-scores. The use of this standardization procedure assists comparison of cognitive performance across different distributions and samples and interpretation of effects. Because UK Biobank does not have tests on the language and visuospatial domains, we computed two sets of global cognition scores in MAS and RCCS, one with and the other without the language and visuospatial domain scores.

To identify the healthy reference subsample, we screened UK Biobank participants and excluded those participants who had severe disorders which may have a significant influence on their cognition. Details of the disease codes (field ID 20002) used can be found in Supplementary Table e-3. Similarly, we used a subsample of healthy individuals at baseline in MAS who were free from medical conditions that may affect cognition for the standardization of test scores. For RCCS, we used the healthy controls as our reference for standardization.

2.5. Statistical analyses

Regression analysis: Linear regression analysis was employed for exploring the relationship between DWI measures and age after controlling for sex, education and imaging centre for UK Biobank, with age as the independent variable and DWI measure as the dependent variable. Linear regression analyses between DWI measure and the global cognition and the main cognitive domain scores were then performed after controlling for age, sex, education and imaging centre for UK Biobank. Hierarchical multiple linear regression was conducted to assess the unique contribution of an individual imaging measure to global cognition. Multicollinearity tests were performed to make sure that the independent variables were not highly correlated by considering variance inflation factors (VIF) above 10 indicating multicollinearity. Demographic characteristics including age, sex, education attainment and imaging centres were initially entered in step 1. Total WMH volume was added in step 2 to examine whether WMH volume provided additional contribution to the prediction of cognition. FA and MD were subsequently entered as

Table 1
Sample characteristics.

	UK Biobank (n = 19369)	MAS (n = 402)	RCCS	
			CSVD (n = 171)	NC (n = 43)
Demographics				
Age, years, mean ± SD	63.11±7.45	77.82±4.58	65.38±7.73	65.74±4.81
range (min, max)	(45.16, 80.71)	(70.29, 90.40)	(50.00, 85.00)	(56.00, 75.00)
Male, number (%)	9123(47.10)	189(46.90)	130(76.00)	31(72.09)
Education, years, mean ± SD	-	11.84±3.59	10.58±3.00	11.67±2.74
Qualification, college number (%)	9009(46.50)	-	-	-
Neuroimaging measures				
Total WMH volume mm^3, mean ± SD	2166.92±2422.31	15387.24±14336.46	23379.76±34064.06	6724.57±7727.95
FA, mean ± SD	0.516±0.018	0.448±0.023	0.560±0.029	0.590±0.026
MD, $\times 10^{-3}$, mean ± SD	0.747±0.024	0.780±0.033	0.788±0.036	0.749±0.028
PSMD, $\times 10^{-3}$, mean ± SD	0.226±0.032	0.408±0.067	0.547±0.116	0.457±0.086

After excluding outliers of neuroimaging measures, in UK Biobank, the number of participants for analysis were: FA n = 19261, MD n = 19229, PSMD n = 19155, Total WMH volume n = 17473; in MAS, FA n = 396, MD n = 398, PSMD n = 395, Total WMH volume n = 401; in CSVD patients of RCCS, FA n = 170, MD n = 171, PSMD n = 169, Total WMH volume n = 170; in NC of RCCS, 43 participants were included for all neuroimaging measures.
Abbreviations: CSVD = cerebral small vessel disease; NC = normal control; FA = mean fractional anisotropy of the skeleton; MD = mean of mean diffusivity of the skeleton; PSMD = peak width of skeletonized mean diffusivity; SD = standard deviation; WMH = white matter hyperintensity.

Table 2
The associations between DWI measures and age, cognition in UK Biobank.

	Age (n=19369)		Processing speed (n = 7936)		Executive (n = 7682)		Memory (n = 8519)		Global (n = 7626 three domains)	
	standardized coefficient	P	standardized coefficient	p	standardized coefficient	p	standardized coefficient	p	standardized coefficient	p
FA	-0.351	<0.001*	0.091	<0.001*	0.076	<0.001*	0.028	0.016	0.082	<0.001*
MD	0.320	<0.001*	-0.059	<0.001*	-0.040	0.001*	-0.010	0.368	-0.047	<0.001*
PSMD	0.445	<0.001*	-0.112	<0.001*	-0.068	<0.001*	-0.025	0.043	-0.086	<0.001*
$DDF^{MD}_{1000,\ 0.05,0.95}$	-0.545	<0.001*	0.117	<0.001*	0.069	<0.001*	0.031	0.017	0.086	<0.001*

Abbreviations: FA = mean fractional anisotropy of the skeleton; MD = mean of mean diffusivity of the skeleton; PSMD = peak width of skeletonized mean diffusivity; $DDF^{MD}_{1000,\ 0.05,0.95}$ = MD computed using Eqs. 1 and 3 with $\theta = 1000$ and $(l,\ u) = (0.05, 0.95)$;
Two tailed p-value of $p < 0.05$ was considered statistically significant. Reported p values refer to the raw p values before Bonferroni correction. *Statistically significant after Bonferroni correction (adjusted p-value: 0.05/4 = 0.0125).

the conventional DWI measures in step 3. In step 4, PSMD was entered. Finally, $DDF^{MD}_{1000,\ 0.05,0.95}$ was entered to in step 5 to examine its contributions to cognitive outcomes in addition to the demographic characteristics and other neuroimaging measures already considered.

Binary logistic regression analysis was used to assess the diagnostic power of DWI measures for classifying CSVD patients from cognitive normal controls in RCCS. Receiver operator characteristic (ROC) curve analysis assessed the sensitivity and specificity of different DWI measures for predicting CSVD status from normal controls. Area under the curve (AUC) and corresponding 95% confidence intervals were calculated to compare the diagnostic ability of different DWI measures. Univariate General linear model (GLM) was applied to identify group difference of regional DDF values between CSDV and NC after correcting for age, sex and education. Null hypothesis for AUC value was set at 0.5.

All statistical analyses were performed in SPSS 26.0 (IBM corporation, USA). Outliers of DWI measures using the ± 3 SD range were excluded. All DWI measures were transformed to Z scores. A two-tailed p-value of 0.05 was considered statistically significant for all analyses. Bonferroni correction was used for adjusting multiple comparisons when conducting the linear regression for different DWI measures and different ROIs.

3. Results

Demographics and neuroimaging characteristics of three study samples are summarized in Table 1. This research examined 19,369 UK Biobank participants without dementia, 402 MAS participants without dementia, and 214 RCCS participants. Both UK Biobank and MAS were recruited from community-dwelling populations. RCCS contained two subsamples: a group of CSVD patients, and a group of cognitively normal controls. MAS had a higher mean age and age range than UK Biobank and RCCS.

3.1. UK Biobank as exploratory sample

We first conducted regression analysis on DWI measures and age after controlling for sex, education and imaging centre, with age as independent variable and DWI measures as dependent variables. The regression coefficient between $DDF^{MD}_{1000,\ 0.05,0.95}$ and age was the highest compared to those of other DWI measures, after controlling for sex, education and imaging centre (standardized β = -0.545, p<0.001) (Table 2). Fig. 2 depicts the associations between DWI measures (FA, MD, PSMD and DDF) and age in UK Biobank fitted with linear regression model. Table 2 also shows the results of associations between DWI measures and different cognitive domains and global cognition in UK Biobank.

The results of hierarchical regression analyses in UK Biobank were listed in Table 3. VIFs of these independent variables were less than 10. After controlling for demographics and all the other neuroimaging metrics, including WMH volume, skeletonized mean FA and MD and PSMD, $DDF^{MD}_{1000,\ 0.05,0.95}$ still explained a small additional variance of global cognition significantly (ΔR^2=0.001, p = 0.007).

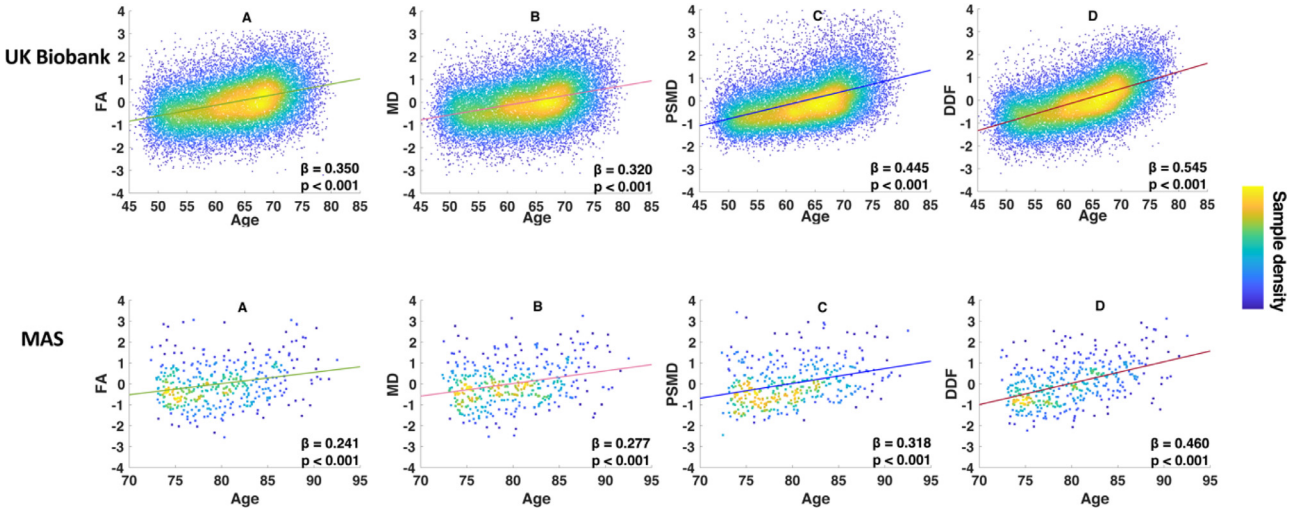

Fig. 2. Associations between age and DWI measures in UK Biobank (upper panel) and MAS (lower panel). Each point in scatter plots A, B, C and D represents an individual. Linear regression model is shown fitting the DWI measure as a function of age. The inverse relationships were transformed to be positive for easy comparison for FA and DDF. The x axis is age in years, and the y axis is the unitless values of residuals of DWI measures after correcting for sex, education and imaging centre. Standardized regression coefficients and p-values are shown at the lower right corner of each subplot, colour bar indicates the density of the points; FA = mean fractional anisotropy of the skeleton; MD = mean of mean diffusivity of the skeleton; PSMD = peak width of skeletonized mean diffusivity; DDF here represents $DDF^{MD}_{1000,\ 0.05,0.95}$= MD computed using Eqs. 1 and 3 with θ=1000 and (l,u)=(0.05,0.95).

Table 3
Hierarchical regression analyses of association of neuroimaging measures with global cognition in UK Biobank.

Measures	R^2	ΔR^2	p
Model 1 Demographics (age, sex, education and imaging centre)	0.179	-	<0.001
Model 2 Demographics + Total WMH volume	0.181	0.002	<0.001
Model 3 Demographics + Total WMH volume + (FA+MD)	0.186	0.005	<0.001
Model 4 Demographics + Total WMH volume + (FA+MD)+PSMD	0.187	0.001	0.002
Model 5 Demographics + Total WMH volume + (FA+MD)+PSMD+ $DDF^{MD}_{1000,\ 0.05,0.95}$	0.188	0.001	0.007

Those participants who had global cognition (7626 out of total 19369 participants in our sample) were included in the hierarchical regression analysis.
Abbreviations: FA = mean fractional anisotropy of the skeleton; MD = mean of mean diffusivity of the skeleton; PSMD = peak width of skeletonized mean diffusivity; $DDF^{MD}_{1000,\ 0.05,0.95}$ = MD computed using Eqs. 1 and 3 with θ =1000 and $(l,\ u)$ = (0.05, 0. 95); WMH = white matter hyperintensity; R^2 = Proportion of variance explained in sample; ΔR^2 = Change in the proportion of variance explained. Reported p values are significance for ΔR^2. Two tailed p-value of p < 0.05 was considered statistically significant.

3.2. Validation in MAS

For validation, we investigated whether similar relationships between the new DWI measure $DDF^{MD}_{1000,\ 0.05,0.95}$ and age and cognition could be observed in the MAS cohort. Consistent with the results in UK Biobank, $DDF^{MD}_{1000,\ 0.05,0.95}$ had the highest regression coefficient with age (standardized $\beta = -0.460$, $p < 0.001$) (Table 4 and Fig. 2) after controlling for sex and education. Results of associations between DWI measures and cognition in MAS are shown in Table 4. Consistent with that of UK Biobank, in comparison with other DWI measures, $DDF^{MD}_{1000,\ 0.05,0.95}$ demonstrated the highest regression coefficient with global cognition.

Subsequent hierarchical regression analysis in MAS (Table 5) showed that $DDF^{MD}_{1000,\ 0.05,0.95}$ ($\Delta R^2 = 0.016$, $p = 0.005$) explained significant additional variance of cognition after controlling for demographics, WMH, FA, MD and PSMD. No multicollinearity existed with all VIFs being smaller than 10.

3.3. Validation in RCCS

To test whether $DDF^{MD}_{1000,\ 0.05,0.95}$ could differentiate CSVD patients from cognitively normal controls, we further extended our validation to RCCS. The analysis was applied to two groups of participants: the CSVD patients (n = 171), and cognitively normal participants (n= 43).

Binary logistic regression: Binary logistic regression analysis was used to identify associated DWI measures which may significantly discriminate CSVD patients. DWI measures including FA, MD, PSMD, $DDF^{MD}_{1000,\ 0.05,0.95}$ were entered into the model as independent variables, and the binary outcome of disease status (1 indicates CSVD, 0 indicates NC) was set as the dependent variable. Age, sex and education were also entered into the model as covariables. FA and $DDF^{MD}_{1000,\ 0.05,0.95}$ were inversely transformed to make their odds ratios (ORs) comparable with the other DWI measures. AD and RD were excluded from the analysis for their high VIFs. The results (Table 6) indicated that $DDF^{MD}_{1000,\ 0.05,0.95}$ was the only significant DWI measure (OR=4.633, p = 0.007, 95% CI = 1.527–14.057) that discriminated CSVD patients form NC.

ROC curves analysis: Summary of the ROC curve analysis is shown in Fig. 3. The highest AUC was found for $DDF^{MD}_{1000,\ 0.05,0.95}$ (AUC = 0.826, S.E. = 0.035, 95% CI = 0.758–0.895, $p < 0.001$). The second highest AUC was for MD (AUC = 0.798, S.E. = 0.035, 95% CI = 0.729–0.867, $p < 0.001$). PSMD was found to have an AUC = 0.750 (AUC = 0.750, S.E. = 0.043, 95% CI = 0.666–0.833, $p < 0.001$). FA shows the smallest diagnostic sensitivity (AUC = 0.778, S.E. = 0.038, CI = 0.704–0.851, $p < 0.001$).

Linear regression analysis was carried out between age and DWI measures within CSVD patients in RCCS. In line with results from the

Table 4
The associations between DWI measures and cognition in MAS.

	Age (n = 402)		Processing speed (n = 387)		Executive (n = 371)		Memory (n = 402)		Language (n = 397)		Spatial (n = 394)		Global (n = 368 five domains)		Global [a] (n = 369 three domains without spatial and language)	
	standardized coefficient	p	standardized coefficient	P	standardized coefficient	p	standardized coefficient	p	standardized coefficient	p	standardized coefficient	p	standardized coefficient	p	standardized coefficient	p
FA	-0.241	<0.001*	0.167	0.001*	0.169	0.001*	0.081	0.082	0.130	0.009*	0.108	0.030	0.181	<0.001*	0.181	<0.001*
MD	0.277	<0.001*	-0.106	0.035	-0.070	0.170	-0.011	0.805	-0.099	0.046	-0.118	0.016	-0.120	0.013	-0.091	0.061
PSMD	0.318	<0.001*	-0.207	<0.001*	-0.201	<0.001*	-0.022	0.647	-0.047	0.367	-0.146	0.004*	-0.161	0.001*	-0.179	<0.001*
$DDF^{MD}_{1000,\ 0.05,0.95}$	-0.460	<0.001*	0.239	<0.001*	0.175	0.002*	0.093	0.067	0.154	0.005*	0.270	<0.001*	0.253	<0.001*	0.217	<0.001*

[a]We computed the global cognition in MAS without spatial and language domains for comparison with UK Biobank which does not have tests on the language and spatial domains.
Abbreviations: FA = mean fractional anisotropy of the skeleton; MD = mean of mean diffusivity of the skeleton; PSMD = peak width of skeletonized mean diffusivity; $DDF^{MD}_{1000,\ 0.05,0.95}$ = MD computed using Eqs. 1 and 3 with θ =1000 and $(l,\ u)$ = (0.05, 0. 95);
Two tailed p-value of p < 0.05 was considered statistically significant. Reported p-values refer to the raw p-values before Bonferroni correction. *Statistically significant after Bonferroni correction (adjusted p-value: 0.05/4 =0.0125).

Fig. 3. Receiver Operator Characteristic analysis predicting CSVD status. The black line is the reference line indicating area under curves (AUC) where the null hypothesis (p < 0.05) is not rejected. Variable with AUC under 0.5 would have no predictive value. FA = mean fractional anisotropy of the skeleton; MD = mean of mean diffusivity of the skeleton; PSMD = peak width of skeletonized mean diffusivity; $DDF^{MD}_{1000,\ 0.05,0.95}$ = MD computed using Eqs. 1 and 3 with θ=1000 and $(l,\ u)$ = (0.05, 0. 95).

Table 5
Hierarchical regression analyses of association of neuroimaging measures with global cognition in MAS.

Measures	R^2	ΔR^2	p
Model 1 Demographics (age, sex and education)	0.231	-	<0.001
Model 2 Demographics + Total WMH volume	0.231	0.000	0.717
Model 3 Demographics + Total WMH volume + (FA+MD)	0.263	0.032	0.001
Model 4 Demographics + Total WMH volume + (FA+MD)+PSMD	0.280	0.017	0.004
Model 5 Demographics + Total WMH volume + (FA+MD)+PSMD+ $DDF^{MD}_{1000,\ 0.05,0.95}$	0.296	0.016	0.005

Those participants who had global cognition with five cognitive domains (368 out of total 402 participants in our sample) were included in the hierarchical regression analysis.
Demographics in MAS included age, sex and education.
Abbreviations: FA = mean fractional anisotropy of the skeleton; MD = mean of mean diffusivity of the skeleton; PSMD = peak width of skeletonized mean diffusivity; $DDF^{MD}_{1000,\ 0.05,0.95}$ = MD computed using Eqs. 1 and 3 with θ =1000 and $(l,\ u)$ = (0.05, 0.95); WMH = white matter hyperintensity; R^2 = Proportion of variance explained in sample; ΔR^2 = Change in the proportion of variance explained.
Reported p-values are significance for ΔR^2. Two tailed p-value of p < 0.05 was considered statistically significant.

Table 6
Binary logistic regression results for group comparison of DWI measures among two groups in RCCS.

Variable	B	S.E.	p-value	OR	95% CI for OR Lower	Upper
Sex	0.204	0.542	0.707	1.226	0.424	3.546
Age	-0.150	0.044	0.001	0.861	0.790	0.938
Education	-0.101	0.075	0.177	0.904	0.781	1.047
FA	0.251	0.664	0.706	1.285	0.350	4.725
MD	0.319	0.644	0.621	1.376	0.389	4.865
PSMD	0.765	0.413	0.064	2.148	0.955	4.829
$DDF^{MD}_{1000,\ 0.05,0.95}$	1.533	0.566	0.007	4.633	1.527	14.057

Abbreviations: FA = mean fractional anisotropy of the skeleton; MD = mean of mean diffusivity of the skeleton; PSMD = peak width of skeletonized mean diffusivity; $DDF^{MD}_{1000,\ 0.05,0.95}$ = MD computed using Eqs. 1 and 3 with θ =1000 and $(l,\ u)$ = (0.05, 0.95); B = regression coefficient; S.E. = standard error; OR = odds ratio; CI = confidence interval. Two tailed p-value of p < 0.05 was considered statistically significant.

Table 7
The associations between regional DDF in eleven tracts and age, cognition in UK Biobank.

	Age (n = 19369)		Processing speed (n = 7936)		Executive (n = 7682)		Memory (n = 8519)		Global (n = 7626 three domains)	
	standardized coefficient	p	standardized coefficient	p	standardized coefficient	p	standardized coefficient	p	standardized coefficient	p
ATR	-0.512	<0.001*	0.091	<0.001*	0.037	0.004*	-0.008	0.532	0.048	<0.001*
CGC	-0.218	<0.001*	0.045	<0.001*	0.039	<0.001*	0.006	0.594	0.040	<0.001*
CGH	-0.232	<0.001*	0.043	<0.001*	0.023	0.035	0.025	0.024	0.036	0.001*
CST	-0.299	<0.001*	0.066	<0.001*	0.031	0.007	0.004	0.722	0.039	<0.001*
Fmajor	-0.206	<0.001*	0.034	0.001*	0.006	0.565	-0.001	0.893	0.014	0.193
Fminor	-0.467	<0.001*	0.097	<0.001*	0.051	<0.001*	0.025	0.053	0.070	<0.001*
IFO	-0.417	<0.001*	0.081	<0.001*	0.041	0.001*	0.008	0.485	0.053	<0.001*
ILF	-0.277	<0.001*	0.054	<0.001*	0.033	0.003*	0.013	0.233	0.041	<0.001*
SLF	-0.495	<0.001*	0.099	<0.001*	0.054	<0.001*	0.001	0.950	0.063	<0.001*
SLFt	-0.150	<0.001*	0.038	<0.001*	0.015	0.179	0.004	0.694	0.024	0.025
UNC	-0.381	<0.001*	0.077	<0.001*	0.056	<0.001*	0.020	0.090	0.062	<0.001*

Abbreviations: ATR, Anterior thalamic radiation; CGC, Cingulum (cingulate gyrus); CGH, Cingulum (hippocampus); CST, Corticospinal tract; Fmajor, Forceps major; Fminor, Forceps minor; IFO, Inferior fronto-occipital fasciculus; ILF, Inferior longitudinal fasciculus; SLF, Superior longitudinal fasciculus; SLFt, Superior longitudinal fasciculus (temporal part); UNC, Uncinate fasciculus;
Two tailed p-value of $p < 0.05$ was considered statistically significant. Reported p-values refer to the raw p-values before Bonferroni correction. *Statistically significant after Bonferroni correction (adjusted p-value: 0.05/11 = 0.0045).

UK Biobank and MAS cohorts, $DDF^{MD}_{1000,\ 0.05,0.95}$ had the highest regression coefficient with age (standardized β = -0.508, p <0.001). The coefficient for PSMD was 0.368 with p <0.001 after Bonferroni correction. Coefficients for FA (standardized β = -0.186, p = 0.017), MD (standardized β = 0.154, p = 0.047) did not survive after Bonferroni correction.

3.4. ROI analysis

Significant correlations were found between age and $DDF^{MD}_{1000,\ 0.05,0.95}$ in all tracts in UK Biobank (see Table 7). For cognition, after correcting for age, sex, education and scanner, $DDF^{MD}_{1000,\ 0.05,0.95}$ in all 11 tracts were found significantly associated with processing speed. Executive function was associated with $DDF^{MD}_{1000,\ 0.05,0.95}$ in ATR (Anterior thalamic radiation), CGC (Cingulum - cingulate gyrus), Fminor (Forceps minor), IFO (Inferior fronto-occipital fasciculus), ILF (Inferior longitudinal fasciculus), SLF (Superior longitudinal fasciculus) and UNC (Uncinate fasciculus). We did not find any significant relationship between memory and DDF in UK Biobank after Bonferroni correction.

Relationships between regional $DDF^{MD}_{1000,\ 0.05,0.95}$ and age, cognition were replicated in MAS cohort (see Table 8). $DDF^{MD}_{1000,\ 0.05,0.95}$ showed significant correlation with age in most of the tracts except for SLFt (Superior longitudinal fasciculus - temporal part), and showed significant correlation with processing speed in ATR, CST (Corticospinal tract), Fmajor (Forceps major), Fminor, IFO, SLF and UNC. Only CST was found correlated with executive function and only CGH (Cingulum - hippocampus) was found correlated with memory. For RCCS cohort, we also conducted group comparison analysis for regional $DDF^{MD}_{1000,\ 0.05,0.95}$ between CSVD patients and normal controls. The results indicated that DDF values in all white matter tracts for CSVD patients were significantly lower than normal controls (see Supplementary Fig. e-2).

4. Discussion

The major objective of this study was to develop an improved automated neuroimaging marker of white matter integrity using DWI scans. To extract more information from the distributions of DWI derived metrics, we proposed a novel general framework based on the concept of Wasserstein distance to describe the cost of shaping one distribution into the other (Villani, 2009). There were several major points to be noted from our study. Firstly, $DDF^{MD}_{1000,\ 0.05,0.95}$ which was computed using the new framework based on MD better explained the variance of the chronological age and cognition of the participant, when compared to FA, MD, AD, RD and PSMD. Secondly, $DDF^{MD}_{1000,\ 0.05,0.95}$ had higher diagnostic accuracy for classifying CSVD from controls than traditional DWI measures we examined. Thirdly, our general framework can be applied to compute DWI-derived metrics other than MD, such as FA, AD, and RD, the results of which are shown in supplementary.

The strong relationship between $DDF^{MD}_{1000,\ 0.05,0.95}$ and the chronological age found in the UK Biobank was replicated in an independent community-dwelling cohort of older adults in MAS. Ageing is accompanied with gradual functional deterioration at the cellular and organismal levels (Wyss-Coray, 2016) and is the predominant risk factor for various chronic age-related brain diseases, such as Alzheimer's disease and cerebral small vessel disease (Partridge et al., 2018; Wardlaw et al., 2013). Ageing related brain function disruption drives progressive impairment of cognitive, emotional and circadian behaviours (Satoh et al., 2017), creating a pressing need to further investigate the role of ageing in brain function and understand the mechanisms of age-associated brain pathophysiology. It is generally accepted that our brains do not experience biological ageing at the same rate (Cole et al., 2019). Ageing brings about changes to the brain size, vasculature, and cognition. Most existing studies of brain age and brain health focus on grey matter (Cole et al., 2018; Franke et al., 2010) using structural MRI. As vascular burden has been proposed to relate to white matter degeneration in cognitively normal older individuals (Lakatta and Levy, 2003; Thom et al., 2006), the DDF measures derived from various DWI maps (FA, MD, AD and RD) showed a high correlation with chronological age, and could be considered as a neuroimaging tool for investigating cerebrovascular ageing related brain structural and functional changes.

The DDF metrics were also shown to be associated with age-related cognitive decline in community-dwelling cohorts, demonstrating high correlations with cognitive changes in global cognition and most cognitive domains. Although the DDF metric we developed in the hierarchical regression analysis of UK Biobank explained only a small additional variance in cognition above the traditional DWI measures, the additional association with cognition in MAS was larger. This is likely in part due to differences in the methods between MAS and UK Biobank; face-to-face interviews the participants were carried out in MAS, while computerized tests were used in UK Biobank. The MAS cognitive test battery was therefore more detailed and comprehensive.

An interesting difference between DDF and traditional DWI derived metrics was that DDF had the highest regression coefficient in the binary logistic regression analysis of RCCS, which suggests that $DDF^{MD}_{1000,\ 0.05,0.95}$ had better diagnostic accuracy in differentiating CSVD patients from normal controls. ROC curve analysis also showed higher diagnostic accuracy of $DDF^{MD}_{1000,\ 0.05,0.95}$ than other DWI measures. Taken

Table 8
The associations between regional DDF in eleven tracts and age, cognition in MAS.

	Age (n =402)		Processing speed (n = 387)		Executive (n = 371)		Memory (n = 402)		Language (n =397)		Spatial (n = 394)		Global (n = 368 five domains)		Global [a] (n = 369 no spatial and language)	
	standardized coefficient	p	standardized coefficient	p	standardized coefficient	p	standardized coefficient	p	standardized coefficient	p	standardized coefficient	p	standardized coefficient	p	standardized coefficient	p
ATR	-0.272	<0.001*	0.173	0.001*	0.121	0.024	0.134	0.005	0.003	0.956	0.164	0.001*	0.155	0.002*	0.178	<0.001*
CGC	-0.287	<0.001*	0.107	0.033	0.021	0.674	0.040	0.389	0.110	0.027	0.116	0.018	0.111	0.021	0.080	0.097
CGH	-0.350	<0.001*	0.134	0.011	0.081	0.131	0.166	0.001*	0.134	0.010	0.183	<0.001*	0.184	<0.001*	0.160	0.002*
CST	-0.311	<0.001*	0.162	0.002*	0.151	0.004*	0.061	0.202	0.088	0.086	0.094	0.062	0.155	0.002*	0.161	0.001*
Fmajor	-0.168	0.001*	0.144	0.004*	0.093	0.063	0.022	0.634	0.050	0.307	0.133	0.006	0.114	0.016	0.097	0.042
Fminor	-0.401	<0.001*	0.180	0.001*	0.124	0.025	0.066	0.191	0.116	0.033	0.210	<0.001*	0.190	<0.001*	0.160	0.002*
IFO	-0.338	<0.001*	0.161	0.002*	0.130	0.012	0.047	0.317	0.084	0.098	0.138	0.006	0.155	0.002*	0.146	0.003*
ILF	-0.273	<0.001*	0.089	0.078	0.052	0.310	0.012	0.791	0.066	0.187	0.094	0.056	0.086	0.077	0.069	0.155
SLF	-0.325	<0.001*	0.157	0.003*	0.147	0.006	0.019	0.698	0.077	0.136	0.124	0.014	0.147	0.003*	0.143	0.005
SLFt	-0.053	0.290	0.114	0.018	0.120	0.014	-0.002	0.973	0.085	0.078	0.049	0.299	0.105	0.023	0.102	0.029
UNC	-0.402	<0.001*	0.206	<0.001*	0.144	0.007	0.113	0.020	0.164	0.002*	0.209	<0.001*	0.232	<0.001*	0.205	<0.001*

[a]We computed the global cognition in MAS without spatial and language domains for comparison with UK Biobank which does not have tests on the language and spatial domains.
Abbreviations: ATR, Anterior thalamic radiation; CGC, Cingulum (cingulate gyrus); CGH, Cingulum (hippocampus); CST, Corticospinal tract; Fmajor, Forceps major; Fminor, Forceps minor; IFO, Inferior fronto-occipital fasciculus; ILF, Inferior longitudinal fasciculus; SLF, Superior longitudinal fasciculus; SLFt, Superior longitudinal fasciculus (temporal part); UNC, Uncinate fasciculus;
Two tailed p-value of $p < 0.05$ was considered statistically significant. Reported p-values refer to the raw p-values before Bonferroni correction. *Statistically significant after Bonferroni correction (adjusted p-value: 0.05/11 = 0.0045).

together, the results demonstrate that DDF metrics have the ability to detect age-related cognitive decline as well as discriminate diseased brains from normal brains. We posit that DDF is a good general framework that can be utilized as a neuroimaging biomarker for investigating white matter health in the ageing process, where the effects of ageing and diseases often overlap and can be additive or interactive or both. In the community-dwelling cohorts of UK Biobank and MAS, ageing was most likely the predominant determinant of cognitive decline, but in CVSD patients of RCCS, vascular burden was likely to have contributed more to cognitive impairment. Development of sensitive imaging markers that can capture this additive and interactive process may help us better understand brain ageing process.

In order to make a straightforward comparison with FA, MD and PSMD, TBSS was applied in our main analysis due to its ability of investigating the entire white matter of the whole brain. It also alleviates the problems caused by less precise registration of white matter fibre tracts and allows for voxel-based analysis. TBSS is convenient for clinical use due to the availability of the computer programs/scripts and methods for fully automated processes. In order to demonstrate the flexibility of our framework, we added the ROI analysis in our study. Interestingly, to our expectation, the ROI results suggested that DDF in all tracts across the brain were significantly correlated with chronological age and cognition especially processing speed. For most of the tracts, DDF did not show significant correlation with memory, which was usually thought to be affected by neurodegenerative factors. For MAS, CGH remained significant after Bonferroni correction, which was consistent with previous ROI analysis of FA and MD in an earlier MAS study of ours (Zhuang et al., 2013). For UK biobank, we also observed the significant relationship between DDF in CGH and memory (p values uncorrected). A study conducted in subjects with subjective memory impairment also reported that only the left CGH had FA decrease and MD increase compared to normal controls (Shao et al., 2019). Our ROI findings demonstrate that our DDF could be a useful tool for examining individual brain structures as well as for the whole brain.

Given that DDF is a flexible framework under which different DWI maps could be calculated, the biological meaning of DDF is mainly the underpinning of the original DWI maps applied. $DDF^{MD}_{1000,\ 0.05,0.95}$ in our study was calculated based on MD distribution. From the mathematical perspective, it is the parameter which describes the shift between the target MD distribution and the MD distribution of a reference. Since it is calculated based on the MD distribution, it reflects the accumulative difference of all the voxels in the MD maps of a subject and a reference. The mechanism of white matter changes in the ageing brain remains unclear; but its aetiology may include different pathologies, such as demyelination and axonal loss, which may lead to the increase of MD values of the ageing brain. DWI technique is well established for its sensitivity of detecting the microstructural impairment of white matter. As a result, many DWI measures are developed to characterize the different physical properties of white matter microstructure, e.g. density, diffusivity and permeability. FA and MD can be used for predicting the white matter changes by estimating the direction and magnitude of water diffusion. However, even these commonly recognized DWI measures are not specific for measuring the pathological changes of the white matter directly (Wozniak and Lim, 2006). As a result, although DWI measures including FA, MD, PSMD and our new developed DDF are highly sensitive to the white matter structure, it is difficult to determine the specific underlying pathologies.

Interestingly, when the DDF framework was applied to the other two commonly used DWI measures AD and RD, i.e. DDF^{AD} and DDF^{RD}, they demonstrated a similar degree of association with age and cognition, which was a significant improvement over the mean AD or RD of the whole brain in this regard. In this study, we describe in detail the results using only DDF^{MD} because we intended to relate and compare our metric with PSMD. On the other hand, DDF^{FA} was not much of an improvement over the mean FA. This poor performance of DDF^{FA} may be explained by the fact that FA is mainly a weighted ratio between the first, second

and third Eigenvalues. Due to the large amount of crossing fibers in the white matter, this ratio would not count as the ratio but the 'mean' of the ratios in the voxel which contains crossing fibers. Given that FA is a measure of how directed the diffusion of water is, it is unable to distinguish between regions of deterioration and crossing fibres. This is problematic because crossing fibres are common in areas where fibre-tracts intersect. MD does not suffer from this pitfall and is thus better able to identify healthy brains.

The development of a new mathematical framework for computing DWI-derived metrics is a key strength of this study. The DDF's mathematical property of capturing a comprehensive difference of two distributions show the advantage of DDF over PSMD or other DWI measures. Such a general framework provides researchers with flexible means to apply different DDF measures to target the individual aims of their studies using different DWI maps, e.g. FA, MD, AD and RD in different forms. The computation of the DDF metrics is automated and the code is publicly available. This study was based on the large exploratory cohort (UK Biobank) and we made replications in two discordant cohorts (MAS is also community-dwelling population and RCCS is a CSVD cohort).

Potential limitations of this study should be noted. All three cohorts in our study were cross-sectional, which did not allow us to examine the trajectories of our DDF metrics in relation to age or cognition. Even though we consider that DDF metrics reflect the cerebrovascular burden because they were computed using DWI properties of brain white matter, it is not possible for us to determine the underlying etiology using these measures because cerebrovascular burden is markedly pleomorphic and our DDF metrics are blind to the direction. Moreover, although the significant explorative results in UK Biobank might be due to its large sample size, the effect was validated in MAS cohort, the sample size of which is relatively much smaller.

In conclusion, this study introduces a novel DWI-derived metric DDF and the mathematical framework for calculating it, using MD and other DWI maps. The sensitivity of tracking the brain ageing process allows DDF to be a biomarker applied for monitoring the structural and functional changes in both healthy and diseased brains.

Credit author statement

Contributions: Writing – JD, FK, AX, JJ, PSS, WW. Conception of idea – WW, FK, AX, JD. Mathematics – AX, FK, WW. Computation and coding– FK. Statistics/Analysis – JD, JC, BL, FK, AT, JJ. Cognition – JD, JC, BL, TL, NK, CFR, WW. Cohorts funding –HB, PSS (MAS), QX (RCCS). Comments/edits – all.

Data and code availability statement

Data for UK Biobank can be accessed by formal application. Deidentified data for MAS can be obtained upon application (MED CHeBA Data: chebadata@unsw.edu.au). Data for RCCS is not publicly available. Code for the computation of DDF is online (https://github.com/ForrestCKoch/DCDF) for the general public.

Declaration of Competing Interest

The authors declare no conflict of interest.

Acknowledgements

The authors thank the participants and the research teams for UK Biobank, Sydney MAS and RCCS. This research has been conducted using the UK Biobank Resource under Application ID 45262 and 37103. Also, we thank Angie Russell for her assistance in the preparation of the manuscript. This research was undertaken with the assistance of resources and services from the National Computational Infrastructure (NCI), which is supported by the Australian Government.

Study fundings

MAS cohort was supported by National Health and Medical Research Council (NHMRC) Australia Project Program Grants ID350833, ID568969 and ID1093083. RCCS was supported by grants from the National Key Research and Development Program of China (2016YFC1300600), Shanghai Science and Technology Committee Project (Natural Science Funding; No. 19ZR1430500), the Project of Collaborative Innovation Centre of Translational Medicine (TM201808) and SJTU-UNSW Collaborative Research Fund. The work was also supported by NHMRC Project Grant ID630593 and John Holden Family Foundation.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2021.118381.

References

Abe, O., Aoki, S., Hayashi, N., Yamada, H., Kunimatsu, A., Mori, H., Yoshikawa, T., Okubo, T., Ohtomo, K., 2002. Normal aging in the central nervous system: quantitative MR diffusion-tensor analysis. Neurobiol. Aging 23, 433–441.

Alfaro-Almagro, F., Jenkinson, M., Bangerter, N.K., Andersson, J.L.R., Griffanti, L., Douaud, G., Sotiropoulos, S.N., Jbabdi, S., Hernandez-Fernandez, M., Vallee, E., Vidaurre, D., Webster, M., McCarthy, P., Rorden, C., Daducci, A., Alexander, D.C., Zhang, H., Dragonu, I., Matthews, P.M., Miller, K.L., Smith, S.M., 2018. Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank. Neuroimage 166, 400–424.

Basser, P.J., Pierpaoli, C., 1996. Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. J. Magn. Reson. B 111, 209–219.

Baykara, E., Gesierich, B., Adam, R., Tuladhar, A.M., Biesbroek, J.M., Koek, H.L., Ropele, S., Jouvent, E., Alzheimer's Disease Neuroimaging, I., Chabriat, H., Ertl-Wagner, B., Ewers, M., Schmidt, R., de Leeuw, F.E., Biessels, G.J., Dichgans, M., Duering, M., 2016. A Novel Imaging Marker for Small Vessel Disease Based on Skeletonization of White Matter Tracts and Diffusion Histograms. Ann. Neurol. 80, 581–592.

Benton, A, 1989. Multilingual aphasia examination. AJA Associates.

Cole, J.H., Marioni, R.E., Harris, S.E., Deary, I.J., 2019. Brain age and other bodily 'ages': implications for neuropsychiatry. Mol. Psychiatry 24, 266–281.

Cole, J.H., Ritchie, S.J., Bastin, M.E., Valdés Hernández, M.C., Muñoz Maniega, S., Royle, N., Corley, J., Pattie, A., Harris, S.E., Zhang, Q., Wray, N.R., Redmond, P., Marioni, R.E., Starr, J.M., Cox, S.R., Wardlaw, J.M., Sharp, D.J., Deary, I.J., 2018. Brain age predicts mortality. Mol. Psychiatry 23, 1385–1392.

Collins, R., 2012. What makes UK Biobank special? Lancet 379, 1173–1174.

Cox, S.R., Ritchie, S.J., Fawns-Ritchie, C., Tucker-Drob, E.M., Deary, I.J., 2019. Structural brain imaging correlates of general intelligence in UK Biobank. Intelligence 76, 101376.

Croall, I.D., Lohner, V., Moynihan, B., Khan, U., Hassan, A., O'Brien, J.T., Morris, R.G., Tozer, D.J., Cambridge, V.C., Harkness, K., Werring, D.J., Blamire, A.M., Ford, G.A., Barrick, T.R., Markus, H.S., 2017. Using DTI to assess white matter microstructure in cerebral small vessel disease (SVD) in multicentre studies. Clin. Sci. (Lond.) 131, 1361–1373.

Du, J., Wang, Y., Zhi, N., Geng, J., Cao, W., Yu, L., Mi, J., Zhou, Y., Xu, Q., Wen, W., Sachdev, P., 2019. Structural brain network measures are superior to vascular burden scores in predicting early cognitive impairment in post stroke patients with small vessel disease. Neuroimage Clin. 22, 101712.

Fazekas, F, J.C., A Alavi, Hurtig, HI, Zimmerman, RA, 1987. MR signal abnormalities at 1.5 T in Alzheimer's dementia and normal aging. Am. J. Roentgenol. 149, 351–356.

Fawns-Ritchie, C., Deary, I.J., 2020. Reliability and validity of the UK Biobank cognitive tests. PLoS One 15, e0231627.

Franke, K., Ziegler, G., Kloppel, S., Gaser, C., Alzheimer's Disease Neuroimaging, I., 2010. Estimating the age of healthy subjects from T1-weighted MRI scans using kernel methods: exploring the influence of various parameters. Neuroimage 50, 883–892.

Grieve, S.M., Williams, L.M., Paul, R.H., Clark, C.R., Gordon, E., 2007. Cognitive aging, executive function, and fractional anisotropy: a diffusion tensor MR imaging study. AJNR Am. J. Neuroradiol. 28, 226–235.

Guo, QH, H.Z., Shi WX, Sun, YM, Lu, CZ, 2006. Boston naming test using by Chinese elderly patient with mild cognitive impairment and Alzheimer's dementia. Chin. Mental Health J. 81–84.

Hamilton, M., 1960. A rating scale for depression. J. Neurol. Neurosurg. Psychiatry 23, 56–62.

Jiang, J., Liu, T., Zhu, W., Koncz, R., Liu, H., Lee, T., Sachdev, P.S., Wen, W., 2018. UBO Detector - A cluster-based, fully automated pipeline for extracting white matter hyperintensities. Neuroimage 174, 539–549.

Jiang, J., Trollor, J.N., Brown, D.A., Crawford, J.D., Thalamuthu, A., Smith, E., Breit, S.N., Liu, T., Brodaty, H., Baune, B.T., Sachdev, P.S., Wen, W., 2015. An inverse relationship between serum macrophage inhibitory cytokine-1 levels and brain white matter integrity in community-dwelling older individuals. Psychoneuroendocrinology 62, 80–88.

Kendall, K.M., Rees, E., Escott-Price, V., Einon, M., Thomas, R., Hewitt, J., O'Donovan, M.C., Owen, M.J., Walters, J.T.R., Kirov, G., 2017. Cognitive performance

among carriers of pathogenic copy number variants: analysis of 152,000 UK biobank subjects. Biol. Psychiatry 82, 103–110.
Lakatta, E.G., Levy, D., 2003. Arterial and cardiac aging: major shareholders in cardiovascular disease enterprises: Part I: aging arteries: a "set up" for vascular disease. Circulation 107, 139–146.
Lam, B.Y.K., Leung, K.T., Yiu, B., Zhao, L., Biesbroek, J.M., Au, L., Tang, Y., Wang, K., Fan, Y., Fu, J.H., Xu, Q., Song, H., Tian, X., Chu, W.C.W., Abrigo, J., Shi, L., Ko, H., Lau, A., Duering, M., Wong, A., Mok, V.C.T., 2019. Peak width of skeletonized mean diffusivity and its association with age-related cognitive alterations and vascular risk factors. Alzheimers Dement (Amst) 11, 721–729.
Lebel, C., Gee, M., Camicioli, R., Wieler, M., Martin, W., Beaulieu, C., 2012. Diffusion tensor imaging of white matter tract evolution over the lifespan. Neuroimage 60, 340–352.
Li, Y., Feng, F., Lin, P., Huang, Z.G., Liu, T., Zhou, B., Yao, H., Zheng, L., Li, C., Wang, P., Zhang, Z., Guo, Y., Wang, L., An, N., Zhu, X., Zhang, X., Wang, J., 2018. Cognition-related white matter integrity dysfunction in Alzheimer's disease with diffusion tensor image. Brain Res. Bull. 143, 207–216.
Lipnicki, D.M., Sachdev, P.S., Crawford, J., Reppermund, S., Kochan, N.A., Trollor, J.N., Draper, B., Slavin, M.J., Kang, K., Lux, O., Mather, K.A., Brodaty, H., 2013. Risk factors for late-life cognitive decline and variation with age and sex in the Sydney Memory and Ageing Study. PLoS One 8, e65841.
Miller, K.L., Alfaro-Almagro, F., Bangerter, N.K., Thomas, D.L., Yacoub, E., Xu, J., Bartsch, A.J., Jbabdi, S., Sotiropoulos, S.N., Andersson, J.L., Griffanti, L., Douaud, G., Okell, T.W., Weale, P., Dragonu, I., Garratt, S., Hudson, S., Collins, R., Jenkinson, M., Matthews, P.M., Smith, S.M., 2016. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. Nat. Neurosci. 19, 1523–1536.
Mwangi, B., Hasan, K.M., Soares, J.C., 2013. Prediction of individual subject's age across the human lifespan using diffusion tensor imaging: a machine learning approach. Neuroimage 75, 58–67.
Partridge, L., Deelen, J., Slagboom, P.E., 2018. Facing up to the global challenges of ageing. Nature 561, 45–56.
Qihao Guo, Q.Z., Chen, Meirong, Ding, Ding, Hong, Zhen, 2009. A comparison study of mild cognitive impairment with 3 memory tests among Chinese individuals. Alzheimer Dis. Assoc. Disord..
Sachdev, P.S., Brodaty, H., Reppermund, S., Kochan, N.A., Trollor, J.N., Draper, B., Slavin, M.J., Crawford, J., Kang, K., Broe, G.A., Mather, K.A., Lux, O., Memory, Ageing, Study, T., 2010. The Sydney Memory and Ageing Study (MAS): methodology and baseline medical and neuropsychiatric characteristics of an elderly epidemiological non-demented cohort of Australians aged 70-90 years. Int. Psychogeriatr. 22, 1248–1264.
Satoh, A., Imai, S.I., Guarente, L., 2017. The brain, sirtuins, and ageing. Nat. Rev. Neurosci. 18, 362–374.
Shao, W., Li, X., Zhang, J., Yang, C., Tao, W., Zhang, S., Zhang, Z., Peng, D., 2019. White matter integrity disruption in the pre-dementia stages of Alzheimer's disease: from subjective memory impairment to amnestic mild cognitive impairment. Eur. J. Neurol. 26, 800–807.
Shin, M.S., Park, S.Y., Park, S.R., Seol, S.H., Kwon, J.S., 2006. Clinical and empirical applications of the Rey-Osterrieth complex figure test. Nat. Protoc. 1, 892–899.
Smith, S.M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T.E., Mackay, C.E., Watkins, K.E., Ciccarelli, O., Cader, M.Z., Matthews, P.M., Behrens, T.E., 2006. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. Neuroimage 31, 1487–1505.
Stroop, J.R., 1935. Studies of interference in serial verbal reactions. J. Experiment. Psychol. 18, 643–662.
Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., Liu, B., Matthews, P., Ong, G., Pell, J., Silman, A., Young, A., Sprosen, T., Peakman, T., Collins, R., 2015. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. 12, e1001779.
Terence, C., Chua, W.W., Slavin, Melissa J., Sachdev, Perminder S., 2008. Diffusion tensor imaging in mild cognitive impairment and Alzheimer's disease: a review. Curr. Opin. Neurol. 21, 83–92.
Thom, T., Haase, N., Rosamond, W., Howard, V.J., Rumsfeld, J., Manolio, T., Zheng, Z.J., Flegal, K., O'Donnell, C., Kittner, S., Lloyd-Jones, D., Goff Jr., D.C., Hong, Y., Adams, R., Friday, G., Furie, K., Gorelick, P., Kissela, B., Marler, J., Meigs, J., Roger, V., Sidney, S., Sorlie, P., Steinberger, J., Wasserthiel-Smoller, S., Wilson, M., Wolf, P., American Heart Association Statistics, C., Stroke Statistics, S., 2006. Heart disease and stroke statistics–2006 update: a report from the American heart association statistics committee and stroke statistics subcommittee. Circulation 113, e85–151.
Villani, C., 2009. The Wasserstein distances. Optimal Transport 93–111.
Villani, C.d., 2003. Topics in Optimal Transportation. American Mathematical Society, Providence, RI.
Wakana, S., Caprihan, A., Panzenboeck, M.M., Fallon, J.H., Perry, M., Gollub, R.L., Hua, K., Zhang, J., Jiang, H., Dubey, P., Blitz, A., van Zijl, P., Mori, S., 2007. Reproducibility of quantitative tractography methods applied to cerebral white matter. Neuroimage 36, 630–644.
Wardlaw, J.M., Smith, E.E., Biessels, G.J., Cordonnier, C., Fazekas, F., Frayne, R., Lindley, R.I., O'Brien, J.T., Barkhof, F., Benavente, O.R., Black, S.E., Brayne, C., Breteler, M., Chabriat, H., DeCarli, C., de Leeuw, F.-E., Doubal, F., Duering, M., Fox, N.C., Greenberg, S., Hachinski, V., Kilimann, I., Mok, V., Oostenbrugge, R.v., Pantoni, L., Speck, O., Stephan, B.C.M., Teipel, S., Viswanathan, A., Werring, D., Chen, C., Smith, C., van Buchem, M., Norrving, B., Gorelick, P.B., Dichgans, M., 2013. Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. Lancet Neurol. 12, 822–838.
Wozniak, J.R., Lim, K.O., 2006. Advances in white matter imaging: a review of in vivo magnetic resonance methodologies and their applicability to the study of development and aging. Neurosci. Biobehav. Rev. 30, 762–774.
Wyss-Coray, T., 2016. Ageing, neurodegeneration and brain rejuvenation. Nature 539, 180–186.
Zhuang, L., Sachdev, P.S., Trollor, J.N., Reppermund, S., Kochan, N.A., Brodaty, H., Wen, W., 2013. Microstructural white matter changes, not hippocampal atrophy, detect early amnestic mild cognitive impairment. PLoS One 8, e58887.