

**White Matter Microstructural Correlates of Cognitive and Motor Functioning Revealed via
Multimodal Multivariate Analysis**

Zaki Alasmar

A Thesis
in the Department of
Psychology

Presented in Partial Fulfilment of the Requirements

For the Degree of
Master of Art (Psychology) at
Concordia University
Montreal, Quebec, Canada

July 2023

© Zaki Alasmar, 2023

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared:

By: Zaki Alasmar

Entitled: White Matter Microstructural Correlates of Cognitive and Motor Functioning
Revealed via Multimodal Multivariate Analysis

and submitted in partial fulfilment of the requirements for the degree of

Master of Arts (Psychology)

Complies with the regulations of the University and meets the accepted standards with respect to originality and quality

Signed by the final examining committee:

Andrew Chapman, Ph.D. Chair

Virginia Penhune, Ph.D. Examiner

Natalie Phillips, Ph.D. Examiner

Christopher Steele, Ph.D. Supervisor

Approved by

Ph.D. Program Director
Andrew Chapman, Ph.D.

Dean of Faculty
Pascale Sicotte, Ph.D.

Date: 09-August-2023

ABSTRACT

White Matter Microstructural Correlates of Cognitive and Motor Functioning Revealed

via Multimodal Multivariate Analysis

Zaki Alasmar, B.Sc.

Concordia University, 2023

Recent advances in cognitive neuroscience emphasise the importance of healthy white matter (WM) in optimal behavioural functioning. It is now widely accepted that brain connectivity via WM contributes to the emergence of behaviour. However, the association between the microstructure of these fibres and behaviour is poorly understood. This is due to indirect and overlapping methods of assessing microstructure, and the use of simplifying approaches in assessing behaviour. Here, we used the Mahalanobis Distance (D2) to integrate 10 metrics of WM derived from multimodal neuroimaging that have strong ties to microstructure. The D2 metric was chosen because it measures the voxelwise distance between every subject and the average, while also accounting for the covariance between the metrics. To examine WM-behaviour associations, we used multivariate regression to examine the voxelwise correlates of 2 cognitive and 2 motor tasks, which allowed us to compare within and across domains in WM. We observed that behaviour is organised in cognitive, motor, and integrative variables that are widespread in their associations with WM, from frontal to parietal regions. Our results highlight the complex nature of microstructure and behaviour, and show the need for multivariate modelling when examining brain-behaviour associations.

Key words: white matter microstructure, behavioural analysis, multivariate statistics, cognition, motor skills

Acknowledgments

I would like to thank my supervisor, Dr. Christopher Steele, for his continuous support throughout the years. I have been a member of Dr. Steele's team for the past 4 years, and I have learned from him essential skills that I will always use, as well as a level of “*criticality*” that I will always apply to my thinking. I would also like to thank Dr. Virginia Penhune and Dr. Natalie Phillips, for dedicating the time to support this project, and for the helpful feedback in all stages. Thank you to Stefanie Tremblay, my colleague who has worked on research projects from which this project emerged, for preparing the data, contributing to this work, and providing invaluable feedback along the way. I am grateful for the contributions of the co-authors on this work, in particular Dr. Claudine Gauthier and Dr. Yasser Iturria-Medina, who helped shape the way I approach research questions. I would like to thank the members of The Neural Architecture, Behaviour, and Connectivity (NeuralABC) Lab for their support and feedback, and fun times we spent together in the lab. Lastly, I would like to thank my partner, family, and friends for their endless love and support.

I would also like to acknowledge the Human Connectome Project, who provided the great resource of high-quality imaging and behavioural data that we used, and all the participants who contributed to it.

Contributions of Authors

This thesis is composed of a single manuscript in preparation for submission entitled:

White Matter Microstructural Correlates of Cognitive and Motor Functioning Revealed via Multimodal Multivariate Analysis

With the authors:

Zaki Alasmar, Stefanie Tremblay, Tobias R. Baumeister, Amir Pirhadi, Felix Carbonell,
Claudine J. Gauthier, Yasser Iturria-Medina, Christopher J. Steele

The CRediT author statement:

Zaki Alasmar: Conceptualization, Methodology, Data Curation, Formal analysis, Visualization, Original Draft, Review & Editing. **Stefanie Tremblay:** Conceptualization, Methodology, Data Curation, Review & Editing. **Tobias R. Baumeister:** Software. **Amir Pirhadi:** Software. **Felix Carbonell:** Software. **Claudine J. Gauthier:** Conceptualization, Methodology. **Yasser Iturria-Medina:** Methodology, Software. **Christopher J. Steele:** Conceptualization, Methodology, Data Curation, Software, Visualization, Review & Editing.

Table of Contents

List of Figures	VII
1. Introduction	1
2. Method.....	8
a. Participants.....	8
b. Multimodal Neuroimaging Protocols	8
c. Microstructural features from multimodal neuroimaging.....	9
Neuroimaging preprocessing.....	9
Group space coregistration	10
d. Multivariate distance model.....	13
e. Detecting outliers	14
f. Behavioural tasks	14
g. Age and sex correction.....	15
h. Statistical analyses	15
3. Results	17
a. Latent Variable 1.....	19
b. Latent Variable 2.....	22
c. Latent Variable 3.....	25
4. Discussion.....	28
5. References	36
6. Appendix A.....	47
7. Appendix B.....	48

List of Figures

Figure 1: Amount of explained covariance per Latent Variable (LV).....	18
Figure 2: The decomposition results of LV1 on behaviour, WM, and their associations.....	20
Figure 3: The decomposition results of LV1 on behaviour, WM, and their associations.....	23
Figure 4: The decomposition results of LV1 on behaviour, WM, and their associations.....	26

1. Introduction

Recent advances in neuroimaging techniques have allowed for the non-invasive assessments of brain structure and function in health and disease (Calhoun, 2018; Tardif et al., 2016). These techniques have become crucial in understanding the neural underpinnings of several aspects of behaviour, such as cognitive and motor functions. The emergence of these functions is supported via the intricate connectivity of cortical regions through white matter (WM) structural connections (Thiebaut de Schotten et al., 2020; Thiebaut de Schotten & Forkel, 2022). The structural connectivity of WM can be assessed via diffusion weighted imaging (DWI), a magnetic resonance imaging (MRI) sequence tailored to indirectly assess WM by measuring water diffusion in brain tissue. DWI allows for the visualisation of connectivity patterns between cortical areas, and also for the examination of WM microstructure (Duval et al., 2017; Soares et al., 2013). For instance, WM organisation, myelination, and density can be assessed through several computational models. We can then relate the extracted microstructural metrics from these models to observable behaviour, thus assessing microstructure-behaviour associations. Examining these associations is crucial for enhancing our understanding of WM correlates of behaviour, their changes in pathology, and in turn improving treatment strategies (Thiebaut de Schotten & Forkel, 2022).

White matter is composed of axons that vary in their myelination, density, and overall composition between different regions, which are assessed via neuroimaging-derived metrics (Mezer et al., 2013; Stikov et al., 2015). Changes in healthy WM composition are largely due to the refinement of the signal conduction environment that WM provides (Hagmann et al., 2010), leading to observable effects such as enhanced or reduced performance on tasks. For instance, adaptive and maladaptive myelination shape this environment by enhancing signal conduction

(Knowles et al., 2022), and animal models have shown that myelin remodulation is required in motor skills and to maintain healthy cognitive functions (e.g. in memory; McKenzie et al., 2014; Pan et al., 2020). These findings are corroborated by human neuroimaging studies, which emphasised the putative role of WM microstructure in supporting and acquiring motor skills and memory (de Lange et al., 2017; Johansen-Berg et al., 2007; Scholz et al., 2009). This modulation of WM microstructure helps provide the optimal conduction medium for neural signals, which in turn leads to ideal cortical functioning and behaviour (Pajevic et al., 2014). The role of WM microstructure in cortical functioning, and in turn behaviour, constitutes a complex system of precise microstructure-function-behaviour mapping. While there is a large body of work in human neuroimaging focused on cortical function and its link to behaviour, the links between WM microstructure and behaviour have not been adequately examined. Indeed, one of the earliest studies examining WM appeared in 2007, and highlighted the need for a thorough understanding of microstructure (Johansen-Berg et al., 2007). Multiple studies followed showing the importance of healthy WM development in behaviour (Muetzel et al., 2008), and the experience-induced plastic changes that it undergoes (Steele et al., 2013). However, due to much of neuroscientific research being focused on cortical grey matter (GM), the patterns of connectivity between cortical areas through WM have been explored extensively at the expense of microstructure. Assessing how WM microstructure may vary across individuals is important for understanding behaviour since WM enables brain connectivity. A problem that arises in microstructural assessment studies is the use of neuroimaging-derived metrics with overlapping characteristics (Raghavan et al., 2021). Each of these metrics is an indirect measure of underlying tissue structure, providing overlapping complementary information. Thus, using univariate frameworks is suboptimal to comprehensively characterise WM microstructure.

Conversely, the use of multivariate frameworks aids in explaining how behaviour emerges from brain neuroanatomy and functional characteristics, and enhances our understanding of WM correlates of behaviour in health and different populations of disease states (Johansen-Berg, 2010).

One problem when examining brain-behaviour associations using univariate techniques is the inability to explicitly compare across tasks. Reducing a facet of behaviour into simpler components allows for an easier assessment of these components via behavioural tasks. In this paradigm, each component is assessed via one administered task. This approach overlooks the connectedness of these components, since the scores are typically analysed with respect to their association with microstructure in an independent statistical framework (Poldrack, 2010; Varoquaux et al., 2018). For instance, Varoquaux and Poldrack (2019) argue that a memory task does not probe only “memory”, but is also likely to include contributions from visual and motor functions. These problems arise due to the use of single assessments of behavioural measures, which are then compared qualitatively (e.g. by visualisation of the spatial extent of different statistical maps). It is also common to use one specific measure (e.g., grip strength), and then make generalised claims within the entire domain (e.g., the motor system). These problems warranted the development of advanced statistical modelling that incorporated data-driven approaches (Varoquaux & Poldrack, 2019), and integrative behavioural atlases (Varoquaux et al., 2018) in lieu of single-task paradigms. Additionally, using multivariate methods have shown promise in identifying latent factors that are tied to human functioning (Schöttner et al., 2023), which may have associations with WM microstructure that are very dissociable in their spatial locations. Given the accumulating evidence of the complexity of brain-behaviour relationships, recent directives in cognitive neuroscience support the use of multiple measures of behaviour.

This paradigm enables the development of atlases of behavioural functioning (called ontologies), and then the identification of brain correlates of these ontologies (Price & Friston, 2005; Varoquaux et al., 2018). One approach to assess how brain neuroanatomy is related to multifaceted behaviour is through the use of multivariate modelling techniques. Applying these techniques has shown that multiple functionally and structurally connected regions are involved in memory and visually-guided motor function (Pur et al., 2022). Other studies have shown that multivariate techniques can help in extracting the sources of maximum brain-behaviour covariance in cognitive abilities, and ground it in brain structure (Ziegler et al., 2013) and function (Voigt et al., 2023). For instance, combining multiple measures of motor functions with multiple measures of cognition allows us to compare between and across behavioural domains (Schöttner et al., 2023). It is also possible to examine the associations between brain function and the extracted general behavioural domains as a whole by following these multivariate assessments of behaviour, combined with multimodal assessments of brain microstructure.

Another problem is that the microstructural metrics used to assess WM are predominantly derived from non-invasive DWI, and provide indirect measures of myelination and fibre characteristics (Raffelt et al., 2017). For instance, the Diffusion Tensor model (DTI) is the most widely used model to assess WM microstructure (Soares et al., 2013), and can be used to compute multiple metrics from which fibre integrity and direction can be inferred. The Fractional Anisotropy (FA) and Mean, Axial, and Radial Diffusivity (MD, AD, RD) measures from the DTI model are commonly used to assess the integrity of white matter microstructure. In a typical DWI experiment, brain-behaviour associations are usually assessed via univariate statistical analysis (e.g. regression models) to identify relationships between the DTI metrics and behaviour (O'Donnell & Westin, 2011). However, in addition to their physiologically unspecific nature, the

DTI metrics are correlated given their shared model characteristics and their dependence on the diffusion signal (Tremblay et al., 2023¹; Raghavan et al., 2021; Tardif et al., 2016). Therefore, disentangling the contributions of the underlying neurophysiological properties of each of these metrics, and in turn developing a more specific understanding of their links to behaviour, is difficult (Tardif et al., 2016; Zatorre et al., 2012). Other more complex models, such as the Constrained Spherical Deconvolution (CSD) and the Neurite Orientation Dispersion and Density Imaging (NODDI) models can also be computed from DWI with multi-shell high angular resolution protocols (Tournier et al., 2019; Zhang et al., 2012). It is possible to extract measures assessing WM fibres properties via CSD by fitting an orientation distribution function (ODF), which yields measures of fibre morphology (Raffelt et al., 2017). The ODF morphological metrics can measure fibre cross section (i.e. observed thickness or FC) and fibre density (FD) at the macro-level by organising WM in overlapping fibre elements, and computing morphological measures of these fibres. Therefore, the ODF elements are more robust measures when the complexity of fibres increases, such as in areas of crossing fibres (Raffelt et al., 2017). Given that WM fibres are composed of axons, increases in their thickness and density are indicative of increases in number of axons as well. This increase in the number of axons in a fibre bundle in turn leads to an increase in the information carrying capacity of those axons. Raffelt and colleagues (2017) hence describe the total information carrying capacity of the fibre as the modulation of fibre density by their cross section (FDC) in a specific location, giving rise to another more holistic measure of morphology at the macro-level organisation of WM. These metrics provide complementary information to DTI metrics, and a more holistic assessment of WM microstructure, but they still suffer from correlation with each other (Tremblay et al., 2023).

¹ Tremblay et al. 2023 is a co-first authored manuscript in preparation for submission and is included as Appendix B.

Whereas ODF metrics assess the macro-level microstructure of WM, NODDI modelling allows for the assessment of intracellular volumetric spaces at the micro-level. This model takes into account the complexity of white matter fibres, and also uses multi-shell DWI. It is possible to compartmentalise cellular versus extracellular contributions within each voxel, and measure the dispersion of the identified cellular compartments (Zhang et al., 2012). These measures can assess intracellular volumes of WM axons, and their orientation dispersion. Similar to the other two models, the microstructural metrics from NODDI are not independent and they suffer from inter-metric association. Overall, we and others have shown previously that there are associations between metrics within the same model as well as between models (Tremblay et al., 2023; Carter et al., 2022; Figley et al., 2022; Uddin et al., 2019). These associations impose great risks when attempting to examine brain-behaviour associations, because they might signify the same microstructural changes and overlapping information (Tardif et al., 2016; Zatorre et al., 2012)

While most human neuroimaging studies have tended to focus on GM, studies examining WM have been increasing in number. However, most of these studies assess microstructure via univariate analyses of metrics with overlapping contributions of tissue properties (Tardif et al., 2016). The non-specific nature of the metrics makes it difficult to disentangle the contribution of tissue properties to the imaging modalities, and gives rise to shared and overlapping changes in the MR signal. To help overcome this problem, we recently developed an integrative quantitative assessment of WM microstructure using multimodal neuroimaging (Tremblay et al., 2023). We proposed that combining neuroimaging metrics that are highly correlated and represent overlapping physiological tissue properties tie intimately to the underlying microstructure. This is achieved by integrating multiple metrics, while accounting for the covariance between them, in a single multivariate score. We developed *mvComp*, a toolbox to compute voxelwise

Mahalanobis Distance (D2). D2 is a generalisation of the Euclidean distance that measures the distance between a point and a distribution in a multi-dimensional space (Mahalanobis, 1936). It also explicitly incorporates the covariance between all of the dimensions so as to not bias the distance with collinearity. In neuroimaging studies, D2 combining several MRI metrics has been used to characterise the extent of microstructural deviation between a subject and a group average at each voxel. For instance, a D2 framework integrating a single metric (e.g., FA) across several WM tracts has also been applied to study pathologies and was found to be strongly linked to epilepsy duration (Owen et al., 2021), and served as a strong marker of severity in traumatic brain injury (Taylor et al., 2020). Our previous work has shown that D2 provides a representation of WM microstructure in the corpus callosum (CC), and that it is possible to extract subsegments of the CC using machine learning and D2 (Tremblay et al., 2023).

In the present study, we used the Mahalanobis distance (D2) to integrate neuroimaging metrics together and applied a partial least squares multivariate statistical analysis to examine the microstructural correlates of cognitive and motor function. We argue that since one behavioural task recruits multiple behavioural functions, using multiple measures and integrating them will provide a stronger brain-behavioural association. Therefore, the aim of this multivariate approach was to extract latent variables that decompose multifaceted behaviour on WM microstructure. We hypothesised that brain-behaviour associations are stronger when integrating multiple measures of microstructure and behaviour together, and that there will be differential patterns in D2's relationship to behaviour across WM.

2. Method

a. Participants

For this investigation, we used data from the WU-Minn Human connectome project (HCP S1200 release). The HCP is a large cohort of healthy young adults with MRI, DWI, and behavioural measures spanning cognition and motor functioning. The full procedure is described in depth elsewhere (Van Essen et al., 2013). Briefly, our sample consisted of 1001 healthy young adults (M_{age} : 28.76 years, SD_{age} : 3.68 years; 556 females), with no history of psychiatric, neurological, or neuropsychological disorders, and no history of substance abuse. We excluded participants with incomplete imaging data and/or invalid acquisitions.

b. Multimodal Neuroimaging Protocols

All imaging was conducted by the HCP on a custom 3T Connectome Skyra MRI scanner with a 32-channel head coil. Anatomical scans were acquired in the first session, and included T1-weighted as well as T2-weighted protocols (Van Essen et al., 2013). The T1w scans were acquired using a 3D-MPRAGE sequence while a 3D T2-SPACE sequence was used for T2w. Anatomical scans were acquired with a 0.7mm isotropic resolution (FOV=224x224). T1w had a TI=1000, TE=2.14, and TR=2400, while T2w TE was 565 and TR of 3200. The DWI data (TE/TR=89.5/5520 ms, FOV=210×180 mm) were multi-shell with b-values of 1000, 2000 and 3000 s/mm² and a 1.25 mm isotropic resolution, 90 uniformly distributed directions, and 6 B0 volumes. More details on the acquisitions can be found at: <https://www.humanconnectome.org/hcp-protocols-ya-3t-imaging>. The imaging data of 1065 young healthy adults, those who had undergone T1w, T2w and diffusion-weighted imaging, were preprocessed. The data of 64 participants were excluded in the current study due to poor cerebellar coverage.

c. Microstructural features from multimodal neuroimaging

Neuroimaging preprocessing

All MRI scans were preprocessed following the minimal preprocessing procedure by the HCP (Glasser et al., 2013). The HCP preprocessing procedure included intensity normalisation of B0 scans along with corrections for eddy current and susceptibility-induced distortions using the different phase encoding directions. It also included co-registration of DWI to native T1w with a rigid body transformation on the mean B0 scans, which was applied to realign all diffusion directions (Bvecs) as well. Finally, motion and gradient nonlinearity correction were applied. All subsequent data preprocessing steps as well as neuroimaging metric extraction are described in detail in Tremblay and colleagues (2023) and discussed briefly below.

First, the preprocessed multishell diffusion data was bias-field corrected using the ANTs' N4 algorithm via *dwibiascorrect* from MRtrix3 (Tustison et al., 2010). Using MRtrix3's *dwi2tensor* (Tournier et al., 2019), we calculated the diffusion tensor and then extracted its metrics with *tensor2metric*, which yielded voxelwise fractional anisotropy (FA), mean diffusivity (MD), axial diffusivity (AD), and radial diffusivity (RD). DTI metrics maps were transformed into the group template space as described below. We then used the multi-tissue Constrained Spherical Deconvolution (CSD) to estimate a white matter fibre response function, and later WM fibre composition. We first segmented all tissue types (WM, GM, CSF) by applying FSL's *5ttgen* (via the MRtrix3 wrapper, Smith et al., 2012; Tournier et al., 2019) on T1w scans. We computed the response function of each tissue type for all participants from the minimally preprocessed DWI data (without bias field correction) and the five-tissue-type (5tt) image using the *msmt_5tt* algorithm of the *dwi2response* function (Dhollander et al., 2016, 2018; Jeurissen et al., 2014; D. A. Raffelt et al., 2017). The response function represents the diffusion

profile of a specific fibre population and is used to estimate the Orientation Distribution Functions (ODFs) with CSD. The WM, GM and CSF response functions were then averaged across all participants, resulting in a single response function for each of the three tissue types. Multi-shell multi-tissue CSD was then performed in each individual with the average response functions to obtain an estimation of orientation distribution functions (ODFs) for each tissue type (Jeurissen et al., 2014). This step is performed using the *dwi2fod msmt_csd* function of MRtrix3 within a brain mask. Bias field correction and global intensity normalisation, which normalise signal amplitudes to make subjects comparable, were performed on the ODFs using the *mtnormalise* function in MRtrix3 (Dhollander et al., 2016, 2018; Jeurissen et al., 2014; D. A. Raffelt et al., 2017).

Group space coregistration

While most registration approaches are focused on optimising GM alignment, between different subjects and across modalities, we opted to perform a WM-focused registration. For this, we used a multi-contrast registration that was primarily driven by WM FODs but also included information about GM. We created population templates for WM, GM, and CSF FODs based on a subset of 200 participants using MRtrix3's *population_template* function, with the following parameters: *nl_update_smooth= 1.0*, *nl_disp_smooth= 0.75* to apply a gaussian smoothing kernel on the gradient and displacement field, respectively. We subsequently computed the warps between all subjects and our population template using MRtrix3's *mrregister* function with identical regularisation parameters. These warps were subsequently applied to the brain masks extracted by using Brain Extraction Tool (*BET*) on T1w, WM FODs, DTI metrics (FA, MD, AD, and RD), T1-weighted (T1w), and T2-weighted (T2w) images through *mrtransform* (Raffelt et al., 2012). During this step, the WM FODs were transformed

without reorientation, aligning the image voxels but not the "fibre bundle elements" (fixels; see below and Raffelt et al., 2015). To create a template mask encompassing only the voxels with data from all subjects, we computed the intersection of all warped brain masks using the *mrmath* min function. Additionally, we warped the WM probability images from the five-tissue-type (5tt) segmentation to the group template space in order to generate a WM mask, and computed the group average probability of WM at each voxel. Lastly, we generated an FOD template average and retained the T1w and T2w images in their native resolution (0.7mm) and calculated the T1w/T2w ratio to generate a proxy myelin map, which was then warped to the FOD template (Glasser & Van Essen, 2011).

The FOD template was then segmented to extract the WM fixel mask. This fixel mask determines the fibre bundle elements (fixels) within each voxel. Fixel segmentation is then performed on the WM FODs of each subject. The fixels are aligned with the template using subject-to-template warps (*fixelreorient* function) and mapped to the corresponding fixels in the fixel mask (*fixelcorrespondence* function). This ensures a consistent set of fixel directions for all subjects. We computed 2 fixel metrics in addition to the computation of the apparent fiber density (FD) of each fixel. First, the fibre bundle cross-section (FC) metric was computed to measure the expansion or contraction required for the fibre bundle to fit the fixel template, and the fibre density and cross-section (FDC) metric was calculated by multiplying the FD and FC metrics, representing the overall capacity of a fibre bundle to carry information.

To incorporate all metrics into a unified multi-modal model, the fixel metric maps were summarised into voxelwise maps. Instead of deriving the measure of total fiber density (FD_{total}) per voxel from fibre-specific FD, we used the sum of FOD lobe integrals. This summation yielded more reproducible estimates, as demonstrated in previous studies (Calamante et al.,

2015). For the fibre cross-section voxel aggregate measure, we computed the weighted mean of FC using the *fixel2voxel* function's mean option. This measure represents the average expansion or contraction required to align fibre bundles within a voxel to the fixels in the template. The weighting factor was based on FD (fibre density), ensuring that bundles with higher density exerted a greater influence on the voxelwise FC value compared to those with lower density. Finally, to assess the overall information-carrying capacity at each voxel, we computed the voxelwise sum of FDC (fibre density and cross-section) using the *fixel2voxel* function's sum option. This measure represents the cumulative capacity to carry information within a voxel (Raffelt et al., 2017).

The bias field-corrected DWI data was also fit with the neurite orientation dispersion and density imaging (NODDI) model using the python implementation of Accelerated Microstructure Imaging via Convex Optimization (AMICO) (Daducci et al., 2015; Zhang et al., 2012). Initially, small variations in b values were addressed by assigning the nearest target b value (0, 1000, 2000, or 3000) to each value in the b-values file. This step aimed to prevent the fitting algorithm from considering slightly different b values as distinct diffusion shells, given that the b-vals from the HCP exhibit slight variations. Subsequently, a diffusion gradient scheme file was created based on the b-vectors and the modified b-values file. The response functions were computed for all compartments, and the fitting procedure was performed on the unbiased DWI volumes, specifically within the brain mask excluding non-brain voxels. The resulting parameters obtained from the fitting process were the intracellular volume fraction (ICVF) assessing neurite density (i.e. interpreted as axonal density in WM) and orientation dispersion index (OD) assessing the extent of dispersion around the mean orientation (Zhang et al., 2012).

In total, 10 voxelwise microstructural measures were computed for each of the 1001 subjects to use as features in our analyses. These features were 4 DTI measures (FA, MD, RD, AD), 3 ODF measures (FC, FD, FDC), and 2 NODDI measures (ICVF & OD) and T1w/T2w. We then computed voxelwise averages for each of these features and the covariance between each pair across all of WM voxels.

d. Multivariate distance model

We computed the Mahalanobis Distance (D2); a voxelwise multivariate distance to incorporate all 10 microstructural features while accounting for the covariance between them. Due to the high covariance between imaging features that capture shared and overlapping biological and physiological mechanisms (Raghavan et al., 2021; Tardif et al., 2016), accounting for it in an explicit statistical model is a more precise way to intimately tie derived metrics to microstructure (Tremblay et al., 2023). The voxelwise approach permits group-level analysis since it results in individual subject maps that are direct assessments of the multivariate differences from the average microstructure. To avoid biasing the average with each subject's data when distances between the subject and the average is computed, a leave-one-subject-out approach was implemented. Here, each subject was removed from the calculated metric averages when that subject's voxelwise D2 scores were computed.

The distance between each subject's microstructural features (i.e., the 10 features) and the average of those features at a given voxel is first calculated resulting in a subject distance vector of the shape 10x1. Then, the distance vector is divided by the covariance between features (in our MVComp approach, this is achieved by multiplication with the pseudoinverse of the covariance matrix; Tremblay et al., 2023). The same procedure is conducted for each voxel in every subject, resulting in voxelwise D2 maps for each of them.

e. Detecting outliers

We computed D2 for all 1001 subjects in our sample. Due to the large number of voxels and features used per subject, outliers were observed in these individualised D2 maps. Therefore, we removed every participant with extensive voxelwise outliers, based on a threshold of 5 standard deviations from their average D2 value. That is, if a participant's D2 map contained 55 or more voxels with D2 values larger than 5 SD (i.e. $> 0.07\%$ voxels) of their own voxelwise average, that participant was dropped from the analysis. This value was chosen as the optimal tradeoff between outlier removal and retaining a large sample - further increasing the cutoff results in an $\sim 50\%$ drop in sample size. As D2 is computed relative to the average, we reran the D2 computations on the resulting final smaller sample of 735 subjects. We then applied a power transformation at each voxel (across all subjects) to normalise the D2 distribution and z-scored for standardisation. Lastly, standardised D2 z-scores with probabilities that exceeded 99.7% in each tail of the distribution were transformed to the closest value on a per-voxel basis since they were not deemed as outliers but rather large values (i.e. values larger than $\sim 3SD$ were transformed to $\sim 3SD$).

f. Behavioural tasks

We used 4 tasks to assess cognition and motor functioning from the NIH toolbox, which were included with the HCP (Reuben et al., 2013; Weintraub et al., 2013). To assess motor functioning, we used the Grip Strength task (GST) and the 9-Hole Pegboard task (9-HPT), examining strength and dexterity respectively. In the GST, subjects were asked to squeeze a hand dynamometer as hard as they can, while in the 9-HPT, they are timed as they place 9 pegs in holes on a board and then remove them as quickly as possible. The tasks chosen to assess cognition were the card sorting task (CST) and the list sorting task (LST), both well validated

and standardised tasks widely used in examining cognitive flexibility and working memory respectively. All behavioural measures were also power-transformed and z-scored. The voxelwise correlations between these tasks and D2 are shown in supplementary figure 1.

g. Age and sex correction

To account for the potentially confounding effects of age and sex in on neuroimaging metrics (Weber et al., 2022), each of the behavioural data as well as the voxelwise D2 values were fitted with a robust linear regression ($\text{task} \sim \text{age} + \text{sex}$, $\text{D2} \sim \text{age} + \text{sex}$) and the residuals retained for subsequent analysis to statistically remove their effects. As with the imaging data, age and sex measures were power-transformed and z-scored prior to inclusion in the analyses.

h. Statistical analyses

A multivariate partial least squares-singular value decomposition (PLS-SVD) analysis was used to examine the association between D2 and the 4 behavioural measures (GST, 9-HPT, CST, and LST). This approach applies a singular value decomposition (SVD) to the brain-behaviour covariance matrix to extract latent variables (LV) (McIntosh & Mišić, 2013). Given a matrix M of voxel and task associations, SVD decomposes M into USV^T , such that U has the shape (voxel \times LV), whereas V^T is an (LV \times task) matrix. Decomposing the covariance matrix allows for the extraction of latent variables that have the largest amount of explained brain-behaviour covariance, along with the weight (importance) of each task and voxel on each latent variable. The weight of each voxel is represented by the columns of U , while the weight of each task is represented by the row of V^T . Statistical significance of the LVs was assessed with a 1000 permutation of the PLS-SVD decomposition. In other words, we conducted 1000 PLS-SVD iterations, shuffling the behavioural input each time and extracting the explained covariance from each iteration. After extracting all the explained covariance values from each

iteration, those that were larger in the original decomposition than the shuffled iterations were deemed statistically significant. We conducted 1000 bootstrap iterations for each voxel and task of every latent variable to assess statistical significance of that voxel or task (by bootstrapping the U and V^T matrices). For each bootstrapping iteration, behavioural as well as D2 data were resampled with replacement, and for each resampled set the PLS-SVD decomposition was rerun. This allowed us to calculate confidence intervals around each task/voxel weight. For permutation and bootstrap testing, the significance level was set to a family-wise error corrected $p < 0.05$. The outcomes from this analysis were the brain-behaviour explained covariance, behavioural and microstructural (i.e. D2) PLS scores for each subject, and task and voxel weights for each LV.

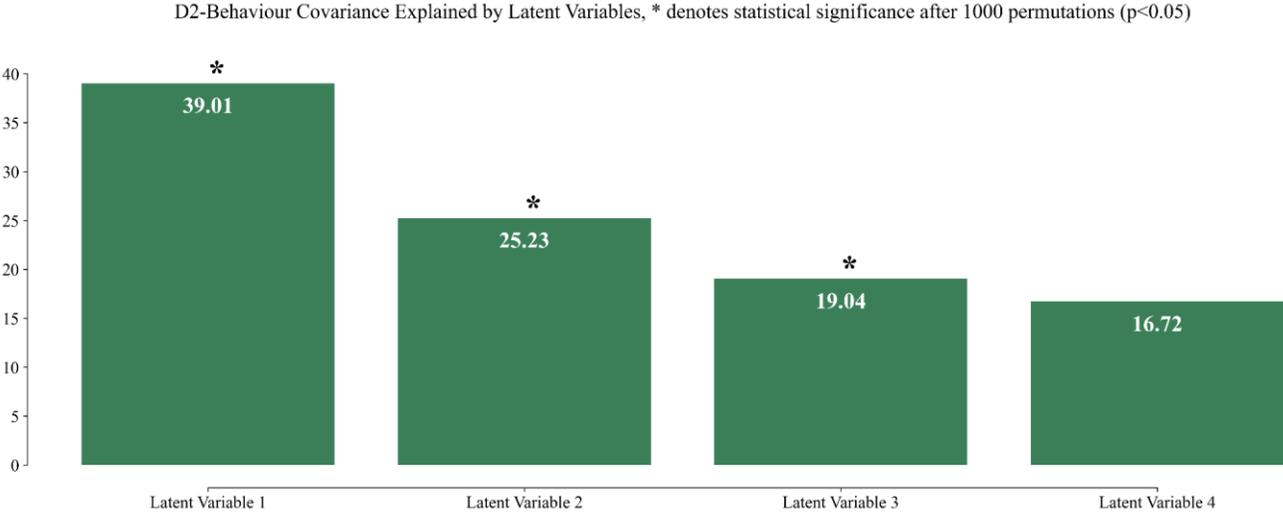
Significant findings were further assessed to identify their WM structural connectivity profiles using a normative model of whole-brain connectivity. For each significant LV, we chose a maximum of 3 regions of spatially connected voxels with large weights in order to individually identify the set of streamlines passing through them. This approach is identical to that used in identifying the network effects of brain lesions (Karnath et al., 2018; Talozzi et al., 2023; Thiebaut de Schotten et al., 2020; Zayed et al., 2020). Here, we used a previously constructed streamline connectivity model that consisted of 10 million representative streamlines between all GM regions in our template space. After extracting the streamlines passing through the clusters, the GM connectivity profile of the WM passing through these clusters was summarised using the Automated anatomical labelling atlas 3 (AAL; Rolls et al., 2015) and using the `tck2connectome` function of MRTrix3 (Tournier et al., 2019).

We used the python package mvComp to compute voxelwise D2 scores for each subject, and to perform all data correction on D2 and behavioural data. We used MATLAB (R2021a) to run the PLS-SVD analysis as well the permutation and bootstrapping iterations.

3. Results

We used D2 and PLS-SVD to examine the multivariate relationship between microstructure and cognitive and motor behaviour. After 1000 permutations, we found that the first 3 latent variables extracted via PLS-SVD were statistically significant, and explained 39.01%, 25.23%, and 19.04% of brain-behaviour covariance, respectively (Figure 1). The PLS scores of D2 and Behaviour were positively correlated in all LVs ($r=0.39$, $r=0.69$, $r=0.70$ for LV1, LV2 and Lv3, respectively; Panel a in fig 2,3,4). After 1000 bootstraps, we found that LV1 was characterised by positive behavioural weights on both cognitive tasks (cognitive flexibility and working memory) and one motor task (manual dexterity) (Figure 2b). It also showed mostly negative weight of voxels that are distributed in WM underlying frontal and parietal regions, as well as cerebellar WM (Figure 2c). In contrast, grip strength had a strong positive weight on the second LV and cognitive flexibility weighed less heavily, while WM voxels showed positive and negative values in parietal and cerebellar WM (Figure 3b,c). Lastly, LV3 showed negative weights of working memory while manual dexterity weighed positively (Figure 4b). There were strong negative and positive weights in adjacent regions underlying frontal WM as well as in the corpus callosum (Figure 4c). Each of the next 3 sections discusses the results of one of the LVs.

Figure 1. Explained brain-behaviour covariance per extracted latent variable.



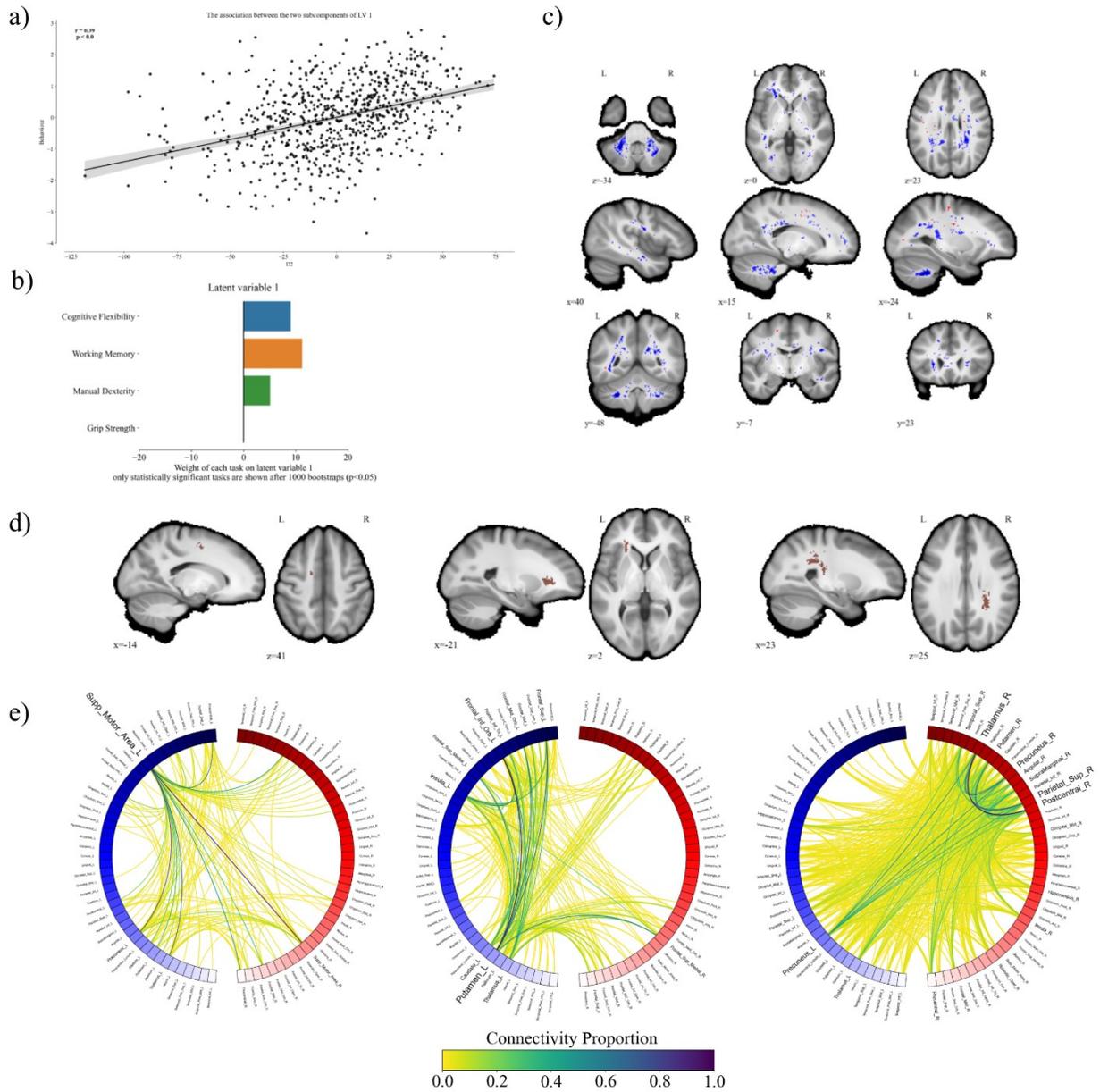
Note. The first latent variable accounts for the largest amount of covariance (39.01%), followed by the second LV (25.23%), the third LV (19.04%), and lastly the fourth LV (16.72%). (*) denotes statistical significance after 1000 permutations (extracting all the explained covariance values from 1000 reshuffled iterations and comparing to the original decomposition). Statistical significance was assessed at family-wise error corrected p-value < 0.05.

a. Latent Variable 1

The first LV revealed a moderate association between transformed D2 and behavioural scores ($r = 0.39$, $p < 0.001$; Figure 2a). We found that cognitive flexibility, working memory, and manual dexterity were statistically significant in their positive weights on LV1 (Figure 2b). White matter voxels adjacent to frontal, temporal, and cerebellar regions had negative weights, while parietal regions were either negatively weighted (e.g. corticospinal tract) or positively weighted (only in WM underlying the pericentral sulcal regions), as shown in Figure 2c. Put differently, these distributed negative voxels across the WM along with the positively weighted tasks are negatively correlated given their opposite signs, while the positively weighted voxels are positively correlated with behaviour.

To examine the connectivity profile of some of the large voxel weights that are spatially connected, we extracted 3 clusters situated in WM underlying the left pericentral sulcus and left prefrontal WM, and in the right corticospinal tract. The connectivity profile of each is shown in Figure 2e. We observed that for the first cluster in Figure 2d, the left supplementary motor area (left SMA) had the most proportion of connections passing through it, followed by the right SMA. For the cluster in prefrontal WM, the left putamen is maximally connected through this cluster, followed by the left frontal inferior orbital cortex. Lastly, the cluster in the corticospinal tract showed that the right thalamus, followed by right parietal cortices, then the right precuneus are maximally connected through it.

Figure 2. Decomposition results of the first latent variable that explained 39.01% of brain-behaviour covariance.



Note. (a) PLS score correlation showed strong association between microstructure (D2) and behaviour ($r = 0.39$, $p < 0.001$). After 1000 bootstraps we observed that, (b) weights of statistically significant tasks reflected cognitive representations on this LV, and (c) weights of statistically significant voxels were mostly negative and distributed in frontal, parietal, and cerebellar WM.

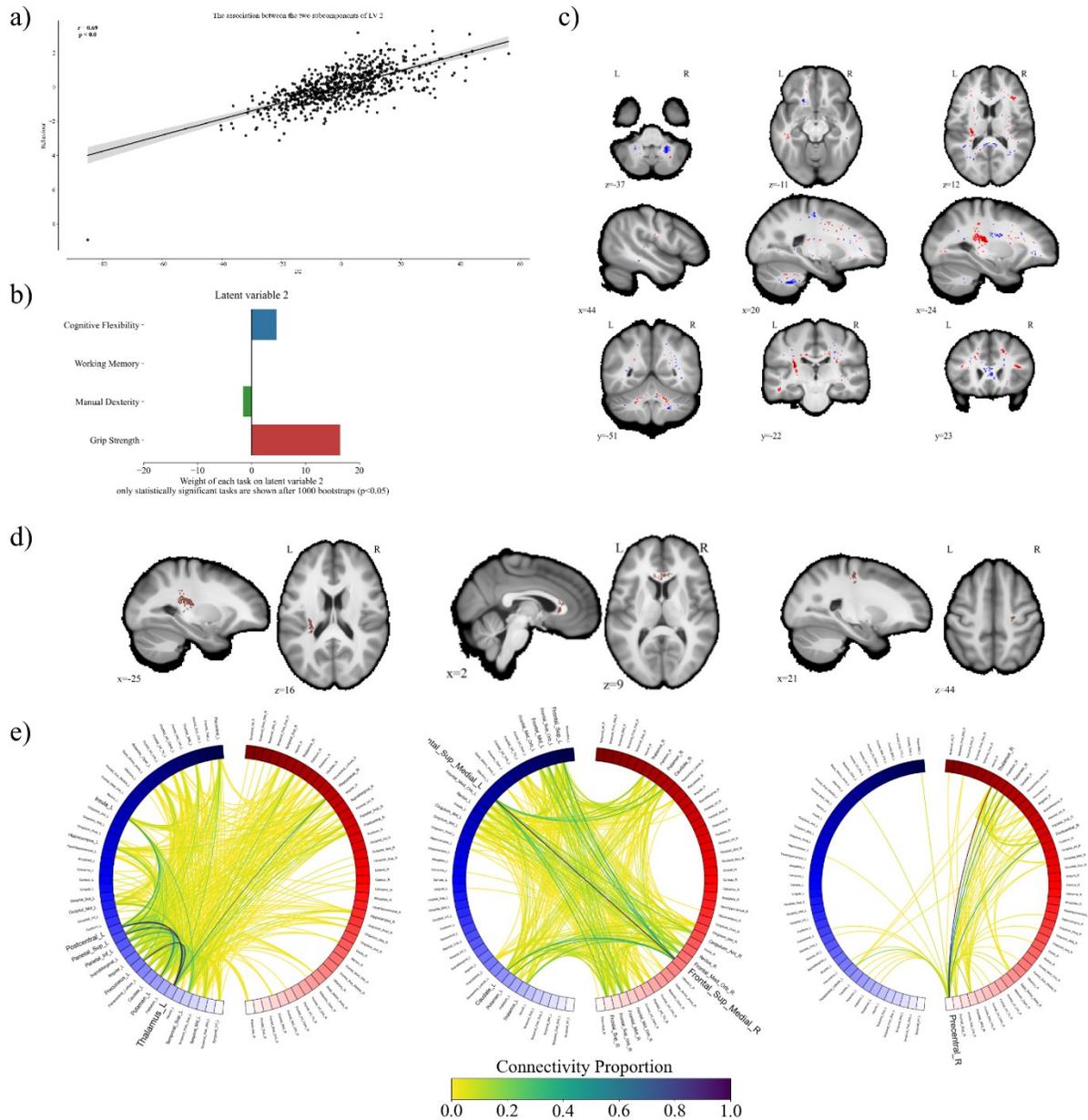
(d) We extracted 3 clusters of largest weights that served as regions of interest in our connectivity analysis, in posterior frontal (left), prefrontal (middle), and parietal (right). (e) connectivity profile of clusters shown in (d) based on automated anatomical atlas (AAL) that showed the involvement of high-order cognitive cortical regions as well premotor ones. Statistical significance was assessed at family-wise error corrected p-value < 0.05 .

b. Latent Variable 2

The second LV revealed a stronger association between PLS scores (D2 and tasks; $r = 0.69$, $p < 0.05$; Figure 3a). Two tasks were positively weighted on this LV; cognitive flexibility and grip strength. However, manual dexterity had a negative weight. This LV was predominantly weighted by grip strength, while working memory weight was not statistically significant after 1000 bootstraps (Figure 3b). Unlike LV1, there were positively- and negatively- weighted voxels distributed in WM (Figure 3c). The positively weighted voxels were in the left parietal, right prefrontal, right temporal, and right superior cerebellar WM, whereas negatively weighted voxels were scattered in left frontal, colossal, and inferior cerebellar WM. The positive voxels in parietal and frontal regions are maximally positively correlated with cognitive flexibility and grip strength, but they are negatively correlated to manual dexterity. On the other hand, the negative voxels in frontal and colossal WM are negatively associated with cognitive flexibility and grip strength, but positively associated with manual dexterity.

Again, to examine the connectivity profile of some of the statistically significant regions, we extracted 3 clusters situated in WM of the left corticospinal tract, genu of the corpus callosum, and WM underlying th in an k toe right primary motor area (Figure 3d). Similarly to what was observed in LV1, the connectivity profile of the first cluster was dominated by left thalamic and parietal regions. The genu of the corpus callosum was the main hub of interhemispheric connectivity, mainly the frontal cortices (e.g. superior medial cortex) and left and right caudate nuclei. As expected, the underlying WM of M1 largely affected the connectivity of M1, followed by the right thalamus and the putamen.

Figure 3. Decomposition results of the second latent variable that explained 25.23% of brain-behaviour covariance.



Note. (a) PLS score correlation showed strong association between microstructure (D2) and behaviour ($r=0.69$, $p<0.001$). After 1000 bootstraps we observed that, (b) weights of statistically significant tasks reflected gross vs fine motor representations on this LV, and (c) weights of statistically significant voxels were distributed in parietal, callosal, and cerebellar WM. (d) We

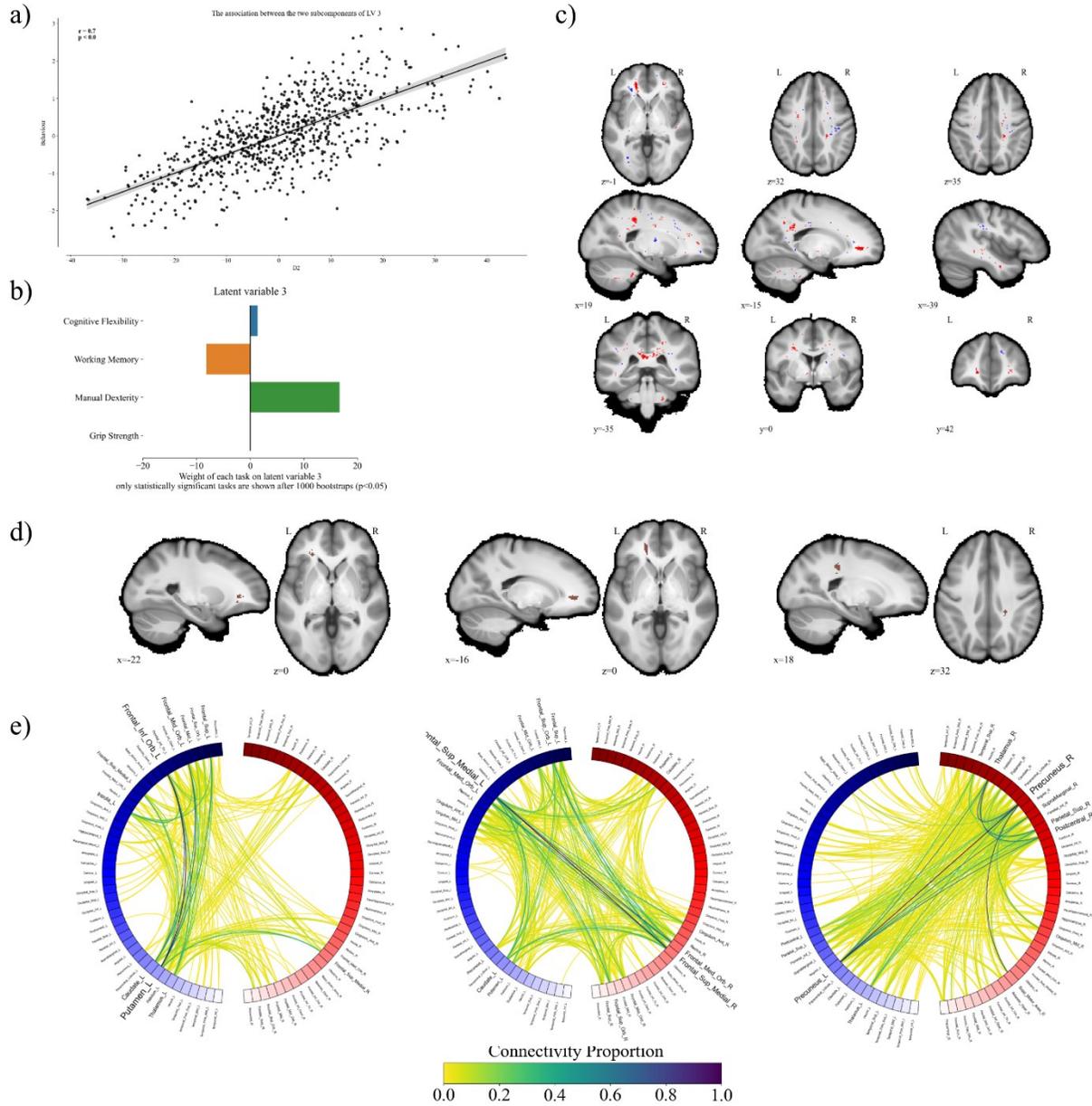
extracted 3 clusters of largest weights that served as regions of interest in our connectivity analysis, in parietal (left), genu of the corpus callosum (middle), and posterior frontal (right). (e) connectivity profile of clusters shown in (d) based on automated anatomical atlas (AAL) that showed the involvement of sensorimotor cortical and subcortical regions, frontal regions, and the motor strip. Statistical significance was assessed at family-wise error corrected p-value < 0.05 .

c. Latent Variable 3

The third latent variable (LV) showed a similar correlation of transformed brain-behaviour scores to LV2, with a correlation coefficient of 0.70 ($p < 0.05$; Figure 4a). Working memory and manual dexterity weighed heavily on this LV, albeit in opposite directions. Memory task was negatively weighted whereas dexterity weighed positively (Figure 4b). There were negatively and positively weighted voxels distributed in WM. Significant positive voxels were in prefrontal WM, adjacent to negative voxels. There were also positively weighted voxels in the right CST and in the splenium of the corpus callosum (Figure 4c). Manual dexterity was positively associated with the positive clusters in the left frontal, CST, and colossal WM while working memory was negatively associated with these regions, and vice versa. On the other hand,

We identified 3 clusters of statistically significant weights (Figure 4d), 2 of which were in prefrontal WM and one in the CST. The negative prefrontal cluster showed streamlines connecting the left frontal inferior orbital cortex. The adjacent positive cluster contained streamlines of the left frontal superior medial cortex. Lastly, the positive cluster in the CST was the passage of streamlines connecting the right precuneus, thalamus, and parietal cortices.

Figure 4. Decomposition results of the third latent variable that explained 19.04% of brain-behaviour covariance.



Note. (a) PLS score correlation showed strong association between microstructure (D2) and behaviour ($r=0.70$, $p<0.001$). After 1000 bootstraps we observed that, (b) weights of statistically

significant tasks reflected cognitive and motor integration on this LV, and (c) weights of statistically significant voxels were distributed in prefrontal and parietal WM. (d) We extracted 3 clusters of largest weights that served as regions of interest in our connectivity analysis, in prefrontal (left & middle), and parietal (right). (e) connectivity profile of clusters shown in (d) based on automated anatomical atlas (AAL) that showed the involvement of frontal and subcortical regions, as well as parietal cortex. Statistical significance was assessed at family-wise error corrected p-value < 0.05.

4. Discussion

The goal of the current study was to examine the association between white matter microstructure and behaviour using multivariate analysis, incorporating 10 neuroimaging metrics and 4 behavioural tasks. By integrating these metrics, we were able to tie non-invasive neuroimaging to WM microstructure, and quantitatively compare across healthy cognitive and motor function. We used the MVComp toolbox to compute voxelwise D2 scores from neuroimaging metrics assessing WM myelination, density, and fibre thickness while accounting for their covariance. We then extracted latent variables of the spatial pattern of association between D2 and 4 tasks spanning motor and cognitive function. We sought to expand recent directives in neuroscience that emphasised the need for holistic mappings of behaviour on the brain (Varoquaux et al., 2018) to WM analysis. In order to do so, we used multivariate statistical integration of brain and behaviour which have shown enhanced validity and reliability compared to traditionally used techniques (Yoo et al., 2019). We extracted multidomain behavioural variables and found their WM microstructural correlates in a large sample of healthy young adults. We observed 3 statistically significant latent variables that explain a total of 83% of the brain-behaviour associations (Figure 1), and decompose behavioural functioning into higher order cognitive and premotor functioning (Figure 2), motor functioning (Figure 3), and integrative cognitive-motor functioning (Figure 4).

We used D2 to integrate neuroimaging-derived metrics which have been shown to be linked strongly to WM microstructure as described in Tremblay and colleagues (2023). There have been recent implementations which suggest that D2 is a more sensitive biomarker in epilepsy, traumatic brain injury, and in detecting individual variability in WM connectivity (Guerrero-Gonzalez et al., 2022; Owen et al., 2021; Taylor et al., 2020). However, our approach

differs from these studies in a number of ways. First, rather than integrating multiple correlated neuroimaging metrics, these studies integrated the same measure across multiple WM tracts. This approach results in a single D2 score per participant, which could be used as a severity, classification, or prediction score at the expense of spatial information. That is, it is not possible to determine the location of differences within WM for each subject. Alternatively, our results showed a whole-WM differential pattern by using multiple microstructural metrics at the voxel level, which was both specific to the subject and the location within WM. Given the specificity of our approach, we also performed multivariate statistical analysis to investigate the association between voxelwise D2 scores and 4 behavioural tasks. We used partial least squares techniques, which is increasing in popularity in neuroimaging research due to its robustness and sensitivity (McIntosh & Lobaugh, 2004; McIntosh & Mišić, 2013). This technique allowed us to provide a WM-behaviour mapping and to explore the holistic nature of the relationship between microstructure and cognitive and motor functions that would not have been possible with other approaches.

The association between transformed scores of microstructure and behaviour were statistically significant for three of the four latent variables (panel a in Figure 2,3,4). This finding indicated that the latent variables we identified were stronger at explaining microstructural differences in our sample than each of the tasks separately (Supplementary Figure 1). That is, combining multiple cognitive and motor tasks allowed us to extract the microstructural correlates of behaviour that were stronger and more holistic than univariate techniques. In all cases, we observed that better performance on cognitive (Figure 2a), motor (Figure 3a), and integrative (Figure 4a) behaviours was correlated with greater distances (D2) in white matter microstructure. This is in agreement with the current frameworks in cognitive neuroscience that focus on the

complex modelling of behaviour, starting with Price and Friston (2005), and more recently expanded on in Varoquaux and colleagues (2018) and Varoquaux and Poldrack (2019). These authors suggested that past and current neuroscientific evidence is overly reductionist, which limits its interpretability and generalizability across behavioural domains. Our method avoided oversimplifying behaviour in a reductionist way, and rather extracted latent variables that ground behaviour in microstructure. Thus it provided a path for the development of ontologies beyond the cortex and into WM. Overall, the effects we observed could be due to plastic changes in WM that lead to enhanced performance, or practice-induced changes in microstructure. Although it is not possible to assess which one precedes the other (Price & Friston, 2005), our technique shows that combining multiple measures of behaviour with multiple measures of microstructure was a stronger approach than independent univariate approaches.

Based on the behavioural weights of each latent variable, we showed that the first LV represented higher order cognitive and premotor skills, the second LV represented motor skills, and the third LV represented integrative functioning (panel b, Figure 2,3,4). These findings emphasised the connectedness and interplay of different components of behaviour during a task (Schöttner et al., 2023). For instance, while cognitive tasks (i.e. cognitive flexibility and working memory) loaded strongly on LV1, the dexterity task weighed similarly. This is likely due to the premotor execution functions required in the task, which are the planning of the reaching and grasping motions (Figure 2b). Dexterity had been previously grouped with cognition tasks via experimental manipulation and when behavioural latent spaces were explored via machine learning (Rodríguez-Aranda et al., 2016; Schöttner et al., 2023). Previous studies examining the neural correlates of fine motor skills suggest that many cortical regions are involved in this complex task, including visuospatial and somatosensory areas (Sobinov & Bensmaia, 2021).

This is contrasted with the second LV, where dexterity and grip strength weighed strongly, in addition to the cognitive flexibility task (Figure 3b). Of note, the working memory task is the only task in this study that did not require the participants to use their arms to perform it (Weintraub et al., 2013). This was supported by the lack of statistically significant weight of this task, and the positive weight of the cognitive flexibility task. Indeed, meta-analytical evidence from functional imaging during cognitive flexibility tasks showed activation in cognitive and motor areas of the cortex (Buchsbaum et al., 2005). Furthermore, we observed that tasks that require gross motor movement are dissociable from the fine motor ones. This was supported by the positive weights of cognitive flexibility and grip strength, and the negative weight of dexterity (Figure 3b). Lastly, the third LV showed strong positive and negative weights of cognitive and motor tasks, reflecting the integrative nature of this variable. This integrative nature could only be extracted using multivariate techniques, and the use of multiple tasks spanning multiple behavioural domains. However, this result did not match the clustering of Schöttner and colleagues (2023), which focused on similarity between the loadings of behavioural tasks. One potential reason for this mismatch between our work and theirs is that our decomposition aimed to extract latent brain-behavioural associations rather than behavioural associations alone. Thus, the incorporation of WM microstructure in behavioural analysis provided a grounding for behaviour in neuroanatomy. There is evidence for disrupted cognitive-motor integration in traumatic brain injury in relation to cortical functioning (Sergio et al., 2020) and post-concussion temporal WM (Hurtubise et al., 2020). However, to our knowledge this was the first exploration of healthy integrative functioning in WM with such a multivariate approach. Therefore, considering our findings, further research is required to assess the effects of incorporating microstructure with behavioural findings.

When examining the voxel weights, we observed widespread shared and overlapping patterns across WM in all LVs (panel c, Figure 2,3,4). These findings could expand our understanding of the complexity of brain-behaviour associations, emphasising the active role of white matter microstructure in supporting higher-order cognitive processes, motor skills, and the integration of cognitive and motor functions. This was evident in voxel weights that spanned traditionally labelled regions as cognitive, motor, and higher-order integration. Namely, frontal and parietal regions for cognitive functions, parietal, cerebellar, and frontal for motor skills, and prefrontal WM for integrative functioning. An interesting pattern of voxel weights emerged between LVs. First, the cognitive LV (LV1) was mostly negatively distributed indicating that better performance on cognitive tasks was associated with smaller deviations of WM microstructure (i.e. larger D2). This pattern was observed for anterior and posterior frontal in addition to parietal WM. Our results were concordant with stroke disconnectome research, which shows cognitive and dexterous impairments with stroke-related disconnection in these regions (Talozzi et al., 2023). Moreover, cortical regions involved in premotor functions, visuospatial imagery, and executive skills were connected through these WM regions (Krüger et al., 2020; Sobinov & Bensmaia, 2021; Suchan et al., 2002). Specifically, the left and right supplementary motor area, the left and right precuneus, and the left putamen and frontal cortices, respectively. Therefore, it is possible that an interplay between microstructure and connectivity within WM gives rise to these behaviours. However, since our results came from a large sample of healthy individuals, we demonstrated that it is possible to localise cognitive functioning in WM prior to injury with an individualised differences approach in assessing microstructure alone. This could aid in symptom prediction and personalised treatment following tissue damage, given the inability to extract connectivity profiles in these cases. Second, the motor LV was positively and

negatively weighted differentially across all of WM. Unsurprisingly, the second LV's positive voxels were highly associated with motor function, and were distributed in WM underlying left motor regions. This effect was largely due to the recruitment of these regions in signal conduction, and the modulation of the microstructure for optimal motor performance (Sampaio-Baptista et al., 2013). The corticospinal tract was positively associated with gross motor performance, and negatively with fine motor skills. Here, all tasks required motor execution, and a simple motor task such as grip strength had the largest weight. This was also supported by the cortices most connected through these clusters. For instance, we observed maximal connectivity of the left thalamus, left and right peri-central and frontal superior cortices in these clusters. For the third latent variable, integration of motor functioning with cognitive abilities has been shown to involve prefrontal and parietal WM as well and the splenium of the corpus callosum. These are consistent findings linking motor, executive and working memory functioning to prefrontal areas (Thiebaut de Schotten & Forkel, 2022). Microstructural differences in these regions may be part of a larger network of cortical regions connected through WM, since the extracted cortical regions of prefrontal GM were involved in integrative functioning (Seidler et al., 2012). Therefore, our observations in the prefrontal WM indicated a modulation of microstructure to support cortical functioning.

This study was not without limitations. We used 4 tasks spanning cognition and motor behaviours to provide a holistic brain-behaviour mapping. However, in order to provide a complete behavioural atlas grounded in WM microstructure, more tasks should be incorporated. Schöttner and colleagues (2023) showed that using dimensionality reduction on a large battery of tasks and extracting 4 factors provided a good trade-off between variance, reliability, and validity in behavioural assessments. Also, there are other functions for each of the measures we

used that are not accounted for with only 4 tasks, such as sensory and emotional behaviour (see Varoquaux & Poldrack, 2019). Therefore, future examinations should use similar dimensionality reduction techniques to Schöttner and colleagues (2023) while also incorporating emotional, language, and sensory measures. This will provide a more complete atlas of human behaviour, and help disentangle the contribution of microstructural differences across WM to healthy functioning. Moreover, the Mahalanobis distance suffers from two main issues. First, it is a squared metric that does not distinguish the directionality of deviation from the average.

Therefore, it was not possible using our method to assess whether there has been an increase or a decrease from the average in WM microstructure. The directionality could be an important factor in assessing longitudinal change, where the average is substituted as the baseline. Second, given that it is a distance to the average, voxelwise D2 values are arbitrary units of the differences in microstructure for a given subject. Guerrero-Gonzalez and colleagues (2022) applied the Wilk's criterion in order to define a statistical significance threshold for D2 scores (Penny, 1996), which would help in identifying the largest distance in a sample. However, given that D2 distributions require normality (Mahalanobis, 1936), computing a threshold of significance for this metric may be biased in non-normal large samples. In future studies, permutation approaches may prove more beneficial in identifying robust and significant D2 values, while also identifying small deviations in order to remove them. Future studies should also use longitudinal designs to explore within subject microstructural changes and their relationship to behaviour. The MVComp toolbox contains specific functions to perform within-subject analysis while also incorporating several microstructural metrics.

In conclusion, we provided a holistic mapping between cognitive and motor behaviours and white matter microstructure. We achieved this goal by using the Mahalanobis distance (D2)

computed from 10 microstructural metrics, and 4 behavioural measures assessing cognitive flexibility, working memory, dexterity, and strength. We showed that using multivariate statistical analysis provides stronger brain-behaviour associations and is able to extract latent variables of cognitive, motor and integrative function. We also provided mappings between these 3 domains of behaviour and WM microstructure displayed as the strongest associations across WM. We extended our analysis to the cortical GM, and showed that there exists an interplay between connectivity and microstructure in order to give rise to and sustain healthy human behavioural functioning.

5. References

- Buchsbaum, B. R., Greer, S., Chang, W., & Berman, K. F. (2005). Meta-analysis of neuroimaging studies of the Wisconsin Card-Sorting task and component processes. *Human Brain Mapping, 25*(1), 35–45. <https://doi.org/10.1002/hbm.20128>
- Calhoun, V. (2018). Data-driven approaches for identifying links between brain structure and function in health and disease. *Dialogues in Clinical Neuroscience, 20*(2), 87–99. <https://doi.org/10.31887/DCNS.2018.20.2/vcalhoun>
- Carter, F., Anwander, A., Goucha, T., Adamson, H., Friederici, A. D., Lutti, A., Gauthier, C. J., Weiskopf, N., Bazin, P.-L., & Steele, C. J. (2022). *Assessing Quantitative MRI Techniques using Multimodal Comparisons* (p. 2022.02.10.479780). bioRxiv. <https://doi.org/10.1101/2022.02.10.479780>
- de Lange, A.-M. G., Bråthen, A. C. S., Rohani, D. A., Grydeland, H., Fjell, A. M., & Walhovd, K. B. (2017). The effects of memory training on behavioral and microstructural plasticity in young and older adults. *Human Brain Mapping, 38*(11), 5666–5680. <https://doi.org/10.1002/hbm.23756>
- Dhollander, T., Raffelt, D., & Connelly, A. (2016, September 11). *Unsupervised 3-tissue response function estimation from single-shell or multi-shell diffusion MR data without a co-registered T1 image*.
- Dhollander, T., Raffelt, D., & Connelly, A. (2018). *Accuracy of response function estimation algorithms for 3-tissue spherical deconvolution of diverse quality diffusion MRI data*. International Society for Magnetic Resonance in Medicine (ISMRM). <https://archive.ismrm.org/2018/1569.html>

- Duval, T., Stikov, N., & Cohen-Adad, J. (2017). Modeling white matter microstructure. *Functional Neurology*, *31*(4), 217–228. <https://doi.org/10.11138/FNeur/2016.31.4.217>
- Figley, C. R., Uddin, M. N., Wong, K., Kornelsen, J., Puig, J., & Figley, T. D. (2022). Potential Pitfalls of Using Fractional Anisotropy, Axial Diffusivity, and Radial Diffusivity as Biomarkers of Cerebral White Matter Microstructure. *Frontiers in Neuroscience*, *15*. <https://www.frontiersin.org/articles/10.3389/fnins.2021.799576>
- Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J. R., Van Essen, D. C., & Jenkinson, M. (2013). The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage*, *80*, 105–124. <https://doi.org/10.1016/j.neuroimage.2013.04.127>
- Guerrero-Gonzalez, J. M., Yeske, B., Kirk, G. R., Bell, M. J., Ferrazzano, P. A., & Alexander, A. L. (2022). Mahalanobis distance tractometry (MaD-Tract) – a framework for personalized white matter anomaly detection applied to TBI. *NeuroImage*, *260*, 119475. <https://doi.org/10.1016/j.neuroimage.2022.119475>
- Hagmann, P., Sporns, O., Madan, N., Cammoun, L., Pienaar, R., Wedeen, V. J., Meuli, R., Thiran, J.-P., & Grant, P. E. (2010). White matter maturation reshapes structural connectivity in the late developing human brain. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(44), 19067–19072. <https://doi.org/10.1073/pnas.1009073107>
- Hurtubise, J. M., Gorbet, D. J., Hynes, L. M., Macpherson, A. K., & Sergio, L. E. (2020). White Matter Integrity and Its Relationship to Cognitive-Motor Integration in Females with and without Post-Concussion Syndrome. *Journal of Neurotrauma*, *37*(13), 1528–1536. <https://doi.org/10.1089/neu.2019.6765>

- Jeurissen, B., Tournier, J.-D., Dhollander, T., Connelly, A., & Sijbers, J. (2014). Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion MRI data. *NeuroImage*, *103*, 411–426. <https://doi.org/10.1016/j.neuroimage.2014.07.061>
- Johansen-Berg, H. (2010). Behavioural relevance of variation in white matter microstructure. *Current Opinion in Neurology*, *23*(4), 351–358. <https://doi.org/10.1097/WCO.0b013e32833b7631>
- Johansen-Berg, H., Della-Maggiore, V., Behrens, T. E., Smith, S. M., & Paus, T. (2007). Integrity of white matter in the corpus callosum correlates with bimanual co-ordination skills. *NeuroImage*, *36*(Suppl 2), T16–T21. <https://doi.org/10.1016/j.neuroimage.2007.03.041>
- Karnath, H. O., Sperber, C., & Rorden, C. (2018). Mapping human brain lesions and their functional consequences. *NeuroImage*, *165*, 180–189. <https://doi.org/10.1016/j.neuroimage.2017.10.028>
- Knowles, J. K., Batra, A., Xu, H., & Monje, M. (2022). Adaptive and maladaptive myelination in health and disease. *Nature Reviews Neurology*, *18*(12), Article 12. <https://doi.org/10.1038/s41582-022-00737-3>
- Krüger, B., Hettwer, M., Zabicki, A., de Haas, B., Munzert, J., & Zentgraf, K. (2020). Practice modality of motor sequences impacts the neural signature of motor imagery. *Scientific Reports*, *10*, 19176. <https://doi.org/10.1038/s41598-020-76214-y>
- Mahalanobis, P. C. (1936). On the Generalised Distance in Statistics. *Proceedings of the National Institute of Sciences of India*, *2*(1), 49–55.
- McIntosh, A. R., & Lobaugh, N. J. (2004). Partial least squares analysis of neuroimaging data: Applications and advances. *NeuroImage*, *23*, S250–S263. <https://doi.org/10.1016/j.neuroimage.2004.07.020>

- McIntosh, A. R., & Mišić, B. (2013). Multivariate Statistical Analyses for Neuroimaging Data. *Annual Review of Psychology*, *64*(1), 499–525. <https://doi.org/10.1146/annurev-psych-113011-143804>
- McKenzie, I. A., Ohayon, D., Li, H., Paes de Faria, J., Emery, B., Tohyama, K., & Richardson, W. D. (2014). Motor skill learning requires active central myelination. *Science*, *346*(6207), 318–322. <https://doi.org/10.1126/science.1254960>
- Mezer, A., Yeatman, J. D., Stikov, N., Kay, K. N., Cho, N., Dougherty, R. F., Perry, M. L., Parvizi, J., Hua, L. H., Butts-Pauly, K., & Wandell, B. (2013). Quantifying the local tissue volume and composition in individual brains with MRI. *Nature Medicine*, *19*(12), 1667–1672. <https://doi.org/10.1038/nm.3390>
- Muetzel, R. L., Collins, P. F., Mueller, B. A., M. Schissel, A., Lim, K. O., & Luciana, M. (2008). The development of corpus callosum microstructure and associations with bimanual task performance in healthy adolescents. *NeuroImage*, *39*(4), 1918–1925. <https://doi.org/10.1016/j.neuroimage.2007.10.018>
- O'Donnell, L. J., & Westin, C.-F. (2011). An introduction to diffusion tensor image analysis. *Neurosurgery Clinics of North America*, *22*(2), 185–viii. <https://doi.org/10.1016/j.nec.2010.12.004>
- Owen, T. W., de Tisi, J., Vos, S. B., Winston, G. P., Duncan, J. S., Wang, Y., & Taylor, P. N. (2021). Multivariate white matter alterations are associated with epilepsy duration. *The European Journal of Neuroscience*, *53*(8), 2788–2803. <https://doi.org/10.1111/ejn.15055>
- Pajevic, S., Basser, P. J., & Fields, R. D. (2014). Role of Myelin Plasticity in Oscillations and Synchrony of Neuronal Activity. *Neuroscience*, *276*, 135–147. <https://doi.org/10.1016/j.neuroscience.2013.11.007>

- Pan, S., Mayoral, S. R., Choi, H. S., Chan, J. R., & Kheirbek, M. A. (2020). Preservation of a remote fear memory requires new myelin formation. *Nature Neuroscience*, 23(4), 487–499. <https://doi.org/10.1038/s41593-019-0582-1>
- Penny, K. I. (1996). Appropriate Critical Values When Testing for a Single Multivariate Outlier by Using the Mahalanobis Distance. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 45(1), 73–81. <https://doi.org/10.2307/2986224>
- Poldrack, R. A. (2010). Mapping mental function to brain structure: How can cognitive neuroimaging succeed? *Perspectives on Psychological Science : A Journal of the Association for Psychological Science*, 5(5), 753–761. <https://doi.org/10.1177/1745691610388777>
- Price, C. J., & Friston, K. J. (2005). Functional ontologies for cognition: The systematic definition of structure and function. *Cognitive Neuropsychology*, 22(3–4), 262–275. <https://doi.org/10.1080/02643290442000095>
- Pur, D. R., Preti, M. G., de Ribaupierre, A., Van De Ville, D., Eagleson, R., Mella, N., & de Ribaupierre, S. (2022). Mapping of Structure-Function Age-Related Connectivity Changes on Cognition Using Multimodal MRI. *Frontiers in Aging Neuroscience*, 14, 757861. <https://doi.org/10.3389/fnagi.2022.757861>
- Raffelt, D. A., Smith, R. E., Ridgway, G. R., Tournier, J.-D., Vaughan, D. N., Rose, S., Henderson, R., & Connelly, A. (2015). Connectivity-based fixel enhancement: Whole-brain statistical analysis of diffusion MRI measures in the presence of crossing fibres. *Neuroimage*, 117, 40–55. <https://doi.org/10.1016/j.neuroimage.2015.05.039>
- Raffelt, D. A., Tournier, J.-D., Smith, R. E., Vaughan, D. N., Jackson, G., Ridgway, G. R., & Connelly, A. (2017). Investigating white matter fibre density and morphology using fixel-

based analysis. *NeuroImage*, 144, 58–73.

<https://doi.org/10.1016/j.neuroimage.2016.09.029>

Raffelt, D., Tournier, J.-D., Rose, S., Ridgway, G. R., Henderson, R., Crozier, S., Salvado, O., & Connelly, A. (2012). Apparent Fibre Density: A novel measure for the analysis of diffusion-weighted magnetic resonance images. *NeuroImage*, 59(4), 3976–3994.

<https://doi.org/10.1016/j.neuroimage.2011.10.045>

Raghavan, S., Reid, R. I., Przybelski, S. A., Lesnick, T. G., Graff-Radford, J., Schwarz, C. G., Knopman, D. S., Mielke, M. M., Machulda, M. M., Petersen, R. C., Jack, C. R., & Vemuri, P. (2021). Diffusion models reveal white matter microstructural changes with ageing, pathology and cognition. *Brain Communications*, 3(2), fcab106.

<https://doi.org/10.1093/braincomms/fcab106>

Reuben, D. B., Magasi, S., McCreath, H. E., Bohannon, R. W., Wang, Y. C., Bubela, D. J., Rymer, W. Z., Beaumont, J., Rine, R. M., Lai, J. S., & Gershon, R. C. (2013). Motor assessment using the NIH Toolbox. *Neurology*, 80(11 Suppl 3), S65–S65.

<https://doi.org/10.1212/wnl.0b013e3182872e01>

Rodríguez-Aranda, C., Mittner, M., & Vasylenko, O. (2016). Association Between Executive Functions, Working Memory, and Manual Dexterity in Young and Healthy Older Adults: An Exploratory Study. *Perceptual and Motor Skills*, 122(1), 165–192.

<https://doi.org/10.1177/0031512516628370>

Rolls, E. T., Joliot, M., & Tzourio-Mazoyer, N. (2015). Implementation of a new parcellation of the orbitofrontal cortex in the automated anatomical labeling atlas. *NeuroImage*, 122, 1–5.

<https://doi.org/10.1016/j.neuroimage.2015.07.075>

- Sampaio-Baptista, C., Khrapitchev, A. A., Foxley, S., Schlagheck, T., Scholz, J., Jbabdi, S., DeLuca, G. C., Miller, K. L., Taylor, A. E., Thomas, N., Kleim, J. A., Sibson, N. R., Bannerman, D. M., & Johansen-Berg, H. (2013). Motor Skill Learning Induces Changes in White Matter Microstructure and Myelination. *The Journal of Neuroscience*, *33*(50), 19499–19503. <https://doi.org/10.1523/jneurosci.3048-13.2013>
- Scholz, J., Klein, M. C., Behrens, T. E. J., & Johansen-Berg, H. (2009). Training induces changes in white-matter architecture. *Nature Neuroscience*, *12*(11), 1370–1371. <https://doi.org/10.1038/nn.2412>
- Schöttner, M., Bolton, T. A. W., Patel, J., Nahálka, A. T., Vieira, S., & Hagmann, P. (2023). Exploring the latent structure of behavior using the Human Connectome Project’s data. *Scientific Reports*, *13*(1), Article 1. <https://doi.org/10.1038/s41598-022-27101-1>
- Seidler, R. D., Bo, J., & Anguera, J. A. (2012). Neurocognitive Contributions to Motor Skill Learning: The Role of Working Memory. *Journal of Motor Behavior*, *44*(6), 445–453. <https://doi.org/10.1080/00222895.2012.672348>
- Sergio, L. E., Gorbet, D. J., Adams, M. S., & Dobney, D. M. (2020). The Effects of Mild Traumatic Brain Injury on Cognitive-Motor Integration for Skilled Performance. *Frontiers in Neurology*, *11*. <https://www.frontiersin.org/articles/10.3389/fneur.2020.541630>
- Smith, R. E., Tournier, J.-D., Calamante, F., & Connelly, A. (2012). Anatomically-constrained tractography: Improved diffusion MRI streamlines tractography through effective use of anatomical information. *NeuroImage*, *62*(3), 1924–1938. <https://doi.org/10.1016/j.neuroimage.2012.06.005>
- Soares, J., Marques, P., Alves, V., & Sousa, N. (2013). A hitchhiker’s guide to diffusion tensor imaging. *Frontiers in Neuroscience*. <http://dx.doi.org/10.3389/fnins.2013.00031>

- Sobinov, A. R., & Bensmaia, S. J. (2021). The neural mechanisms of manual dexterity. *Nature Reviews. Neuroscience*, 22(12), 741–757. <https://doi.org/10.1038/s41583-021-00528-7>
- Steele, C. J., Bailey, J. A., Zatorre, R. J., & Penhune, V. B. (2013). Early musical training and white-matter plasticity in the corpus callosum: Evidence for a sensitive period. *Journal of Neuroscience*, 33(3), 1282–1290. <https://doi.org/10.1523/JNEUROSCI.3578-12.2013>
- Stikov, N., Campbell, J. S. W., Stroh, T., Lavelée, M., Frey, S., Novek, J., Nuara, S., Ho, M.-K., Bedell, B. J., Dougherty, R. F., Leppert, I. R., Boudreau, M., Narayanan, S., Duval, T., Cohen-Adad, J., Picard, P.-A., Gasecka, A., Côté, D., & Pike, G. B. (2015). In vivo histology of the myelin g-ratio with magnetic resonance imaging. *NeuroImage*, 118, 397–405. <https://doi.org/10.1016/j.neuroimage.2015.05.023>
- Suchan, B., Yágüez, L., Wunderlich, G., Canavan, A. G. M., Herzog, H., Tellmann, L., Hömberg, V., & Seitz, R. J. (2002). Neural correlates of visuospatial imagery. *Behavioural Brain Research*, 131(1–2), 163–168. [https://doi.org/10.1016/S0166-4328\(01\)00373-4](https://doi.org/10.1016/S0166-4328(01)00373-4)
- Talozzi, L., Forkel, S. J., Pacella, V., Nozais, V., Allart, E., Piscicelli, C., Pérennou, D., Tranel, D., Boes, A., Corbetta, M., Nachev, P., & Thiebaut de Schotten, M. (2023). Latent disconnectome prediction of long-term cognitive-behavioural symptoms in stroke. *Brain*, 146(5), 1963–1978. <https://doi.org/10.1093/brain/awad013>
- Tardif, C. L., Gauthier, C. J., Steele, C. J., Bazin, P.-L., Schäfer, A., Schaefer, A., Turner, R., & Villringer, A. (2016). Advanced MRI techniques to improve our understanding of experience-induced neuroplasticity. *NeuroImage*, 131, 55–72. <https://doi.org/10.1016/j.neuroimage.2015.08.047>

- Taylor, P. N., Silva, N. M. da, Blamire, A., Wang, Y., & Forsyth, R. (2020). Early deviation from normal structural connectivity: A novel intrinsic severity score for mild TBI. *Neurology*, *94*(10), e1021–e1026. <https://doi.org/10.1212/WNL.00000000000008902>
- Thiebaut de Schotten, M., & Forkel, S. J. (2022). The emergent properties of the connected brain. *Science*, *378*(6619), 505–510. <https://doi.org/10.1126/science.abq2591>
- Thiebaut de Schotten, M., Foulon, C., & Nachev, P. (2020). Brain disconnections link structural connectivity with function and behaviour. *Nature Communications*, *11*, 5094. <https://doi.org/10.1038/s41467-020-18920-9>
- Tournier, J.-D., Smith, R., Raffelt, D., Tabbara, R., Dhollander, T., Pietsch, M., Christiaens, D., Jeurissen, B., Yeh, C.-H., & Connelly, A. (2019). MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation. *NeuroImage*, *202*, 116137. <https://doi.org/10.1016/j.neuroimage.2019.116137>
- Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., & Gee, J. C. (2010). N4ITK: Improved N3 Bias Correction. *IEEE Transactions on Medical Imaging*, *29*(6), 1310–1320. <https://doi.org/10.1109/TMI.2010.2046908>
- Uddin, Md. N., Figley, T. D., Solar, K. G., Shatil, A. S., & Figley, C. R. (2019). Comparisons between multi-component myelin water fraction, T1w/T2w ratio, and diffusion tensor imaging measures in healthy human brain structures. *Scientific Reports*, *9*, 2500. <https://doi.org/10.1038/s41598-019-39199-x>
- Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E. J., Yacoub, E., & Ugurbil, K. (2013). The WU-Minn Human Connectome Project: An overview. *NeuroImage*, *80*, 62–79. <https://doi.org/10.1016/j.neuroimage.2013.05.041>

- Varoquaux, G., & Poldrack, R. A. (2019). Predictive models avoid excessive reductionism in cognitive neuroimaging. *Current Opinion in Neurobiology*, *55*, 1–6.
<https://doi.org/10.1016/j.conb.2018.11.002>
- Varoquaux, G., Schwartz, Y., Poldrack, R. A., Gauthier, B., Bzdok, D., Poline, J.-B., & Thirion, B. (2018). Atlases of cognition with large-scale human brain mapping. *PLOS Computational Biology*, *14*(11), e1006565. <https://doi.org/10.1371/journal.pcbi.1006565>
- Voigt, K., Liang, E. X., Masic, B., Ward, P. G. D., Egan, G. F., & Jamadar, S. D. (2023). Metabolic and functional connectivity provide unique and complementary insights into cognition-connectome relationships. *Cerebral Cortex (New York, N.Y.: 1991)*, *33*(4), 1476–1488. <https://doi.org/10.1093/cercor/bhac150>
- Weber, K. A., Teplin, Z. M., Wager, T. D., Law, C. S. W., Prabhakar, N. K., Ashar, Y. K., Gilam, G., Banerjee, S., Delp, S. L., Glover, G. H., Hastie, T. J., & Mackey, S. (2022). Confounds in neuroimaging: A clear case of sex as a confound in brain-based prediction. *Frontiers in Neurology*, *13*, 960760. <https://doi.org/10.3389/fneur.2022.960760>
- Weintraub, S., Dikmen, S. S., Heaton, R. K., Tulsky, D. S., Zelazo, P. D., Bauer, P. J., Carlozzi, N. E., Slotkin, J., Blitz, D., Wallner-Allen, K., Fox, N. A., Beaumont, J. L., Mungas, D., Nowinski, C. J., Richler, J., Deocampo, J. A., Anderson, J. E., Manly, J. J., Borosh, B., ... Gershon, R. C. (2013). Cognition assessment using the NIH Toolbox. *Neurology*, *80*(11 Suppl 3), S54. <https://doi.org/10.1212/WNL.0b013e3182872ded>
- Yoo, K., Rosenberg, M. D., Noble, S., Scheinost, D., Constable, R. T., & Chun, M. M. (2019). Multivariate approaches improve the reliability and validity of functional connectivity and prediction of individual behaviors. *NeuroImage*, *197*, 212–223.
<https://doi.org/10.1016/j.neuroimage.2019.04.060>

Zatorre, R. J., Fields, R. D., & Johansen-Berg, H. (2012). Plasticity in Gray and White. *Nature Neuroscience*, *15*(4), 528–536. <https://doi.org/10.1038/nn.3045>

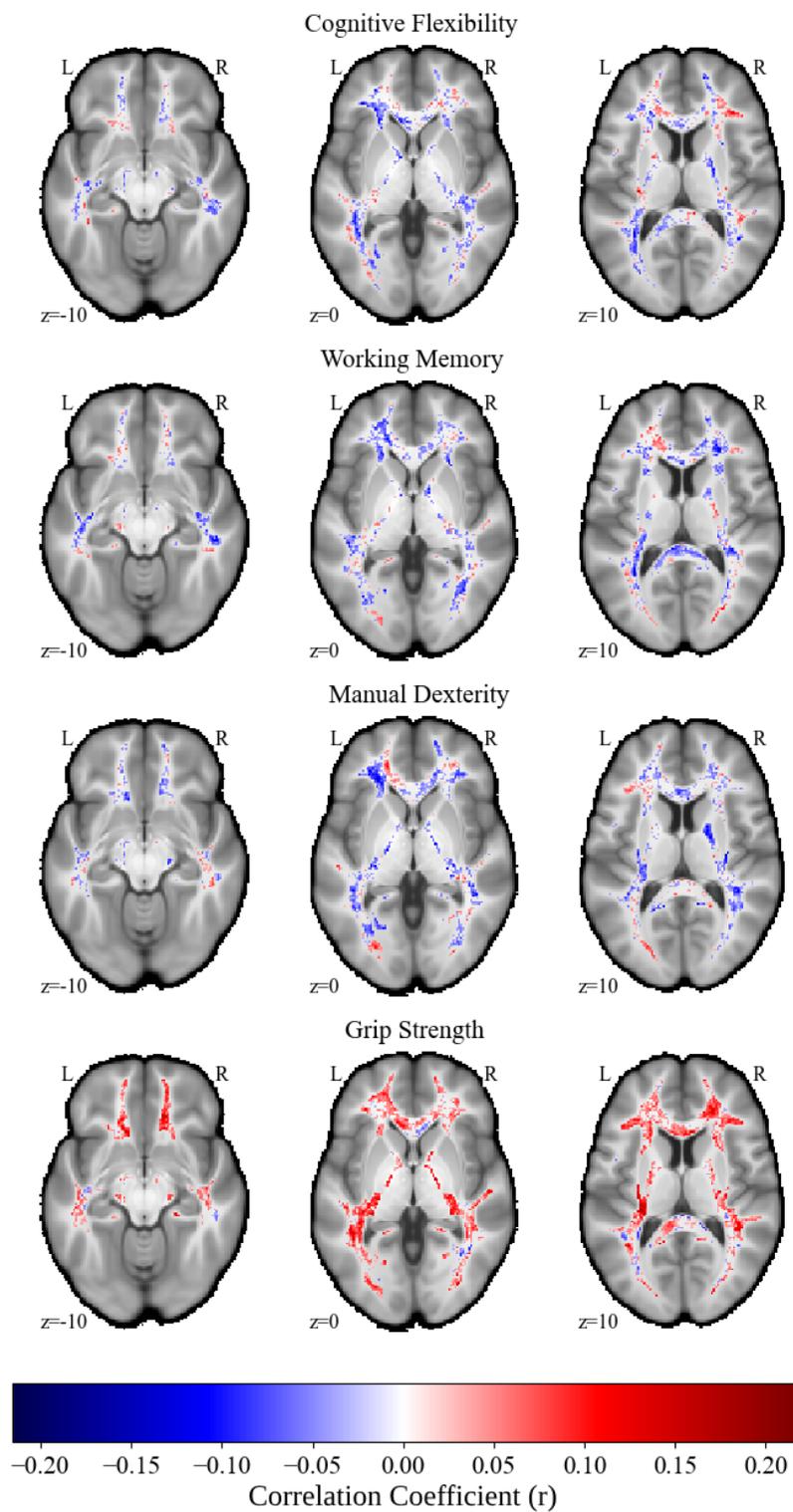
Zayed, A., Iturria-Medina, Y., Villringer, A., Sehm, B., & Steele, C. J. (2020). Rapid Quantification of White Matter Disconnection in the Human Brain. *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 1701–1704. <https://doi.org/10.1109/EMBC44109.2020.9176229>

Zhang, H., Schneider, T., Wheeler-Kingshott, C. A., & Alexander, D. C. (2012). NODDI: Practical in vivo neurite orientation dispersion and density imaging of the human brain. *NeuroImage*, *61*(4), 1000–1016. <https://doi.org/10.1016/j.neuroimage.2012.03.072>

Ziegler, G., Dahnke, R., Winkler, A. D., & Gaser, C. (2013). Partial least squares correlation of multivariate cognitive abilities and local brain structure in children and adolescents. *NeuroImage*, *82*, 284–294. <https://doi.org/10.1016/j.neuroimage.2013.05.088>

6. Appendix A

Supplementary figure 1: the association between each task and D2 scores at each voxel.



7. Appendix B

The Manuscript in preparation for the presentation and validation of Mahalanobis Distance in neuroimaging analysis. Referred to in-text as Tremblay et al., 2023.

MVComp: Multivariate Comparisons of White Matter Microstructure Accounting for Covariance Between Features

Stefanie Tremblay*, Zaki Alasmar*, Amir Pirhadi, Felix Carbonell, Yasser Iturria-Medina,
Claudine Gauthier, Christopher Steele

* Contributed equally to this work.

Table of Contents

1. Introduction	50
2. Method.....	58
2.1 General framework.....	58
2.1.1 Comparisons between subject(s) and a reference.....	59
2.1.2 Comparisons within a single subject	60
2.2 Data preparation	64
2.3 Computing the reference standard and covariance matrix	64
2.3.1 Comparisons between subjects and a reference	64
2.3.2 Within-subject comparisons	65
2.4 Computing D2.....	66
2.4.1 Comparisons between subjects and a reference	66
2.4.2 Within-subject comparisons	67
2.5 Statistical analysis	67
2.5.1 Comparisons between subjects and a reference	67
2.5.2 Within-subject comparisons	68
2.6 Determining feature importance.....	68
2.6.1 Comparisons between subjects and a reference	68
2.6.2 Within-subject comparisons	69
2.7 Experiments.....	69
2.7.1 Data Description.....	69
2.7.2 Preprocessing.....	70
2.7.3 Experiment 1: Comparisons between subjects and a reference.....	74
2.7.4 Experiment 2: Within-subject comparisons	75
3. Results	77
3.1 Experiment 1: Comparisons between subjects and a reference	77
3.2 Experiment 2: Within-subject comparisons	79
4. Discussion.....	82

1. Introduction

In the past decade, there has been exponential growth in the number of modelling approaches that link white matter (WM) microstructural properties and the MR signal (Novikov et al., 2018). Since none of the existing models (e.g., the diffusion tensor, neurite orientation dispersion and density imaging (NODDI), etc.) is a perfect representation of the underlying microstructure, choosing a model and contrast for analyses can be challenging and has the potential to lead to errors in interpretation (Novikov et al., 2018). Multi-modal imaging, and multivariate frameworks that combine several parameters derived from different models and modalities, have been suggested as a promising avenue to harness the complementarity of different neuroimaging-derived metrics to more accurately capture properties of the brain (Uddin et al., 2019; Tardif et al., 2016).

Multivariate frameworks have the potential to counteract issues arising from the physiologically-unspecific nature of commonly used neuroimaging metrics and to capture the complexity and heterogeneity of biological properties (Tardif et al., 2016; Dean et al., 2017; Taylor et al., 2020; Seidlitz et al., 2018; Guberman et al., 2022). Multiple mechanisms give rise to brain structure such as myelination and cell proliferation. These mechanisms support neuroplastic change (Azzarito et al., 2023) and behavioral performance (Seidlitz et al., 2018; Thiebaut de Schotten & Forkel, 2022), and are involved in neurological disorders (Iturria-Medina et al., 2017). Interpreting the results of univariate statistical analyses is thus challenging within this context. In addition to capturing a more nuanced picture of the expected mechanisms, multivariate statistical frameworks can offer greater statistical power than multiple univariate analyses as they reduce the amount of multiple comparisons correction required (Naylor et al., 2014; Avants et al., 2008; Owen et al., 2020). Lastly, and perhaps most importantly, multivariate

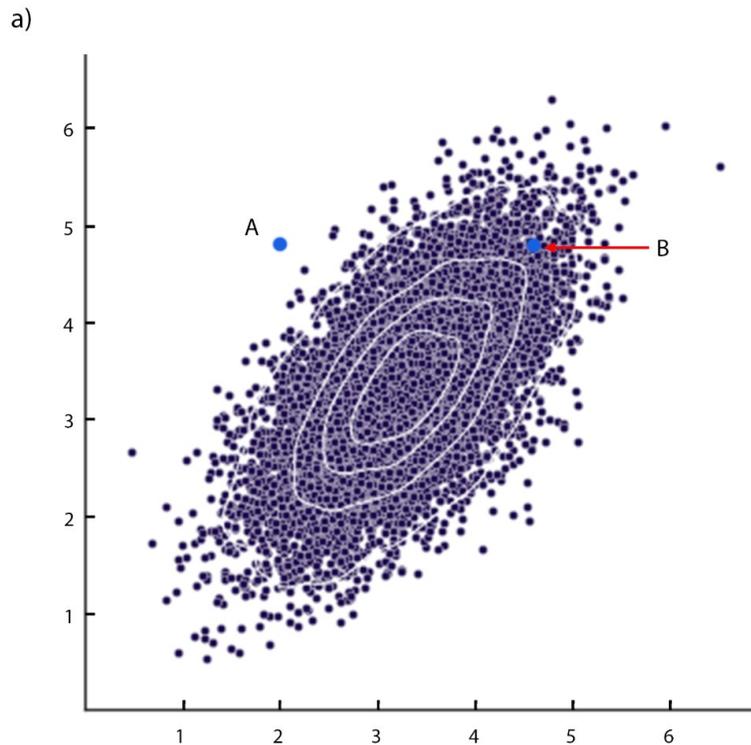
frameworks can be leveraged to move away from group comparisons and towards individual-level analyses, an essential step on the road to precision medicine (Chamberland et al., 2021; Marquand et al., 2016; Wolfers et al., 2018).

Multivariate approaches that combine structural MRI metrics have been used in a number of promising contexts. At the group level, partial least squares (PLS) analyses and their variants can be used to assess the covariance of multiple metrics. Other multivariate approaches that can be used in group analyses include principal component analysis (PCA), independent component analysis (ICA) and non-negative matrix factorization. At the individual level, inter-regional correlations across multiple metrics can be used to create individual-specific correlational maps that can be linked to behavior (Seidlitz et al., 2018). Individual network maps provide a more comprehensive structural mapping that captures both biological complexity and individual variability because they integrate multiple MRI features (e.g., Whitaker et al., 2016; Vandekar et al., 2016). However, in this study, the shared **covariance between metrics** was not accounted for. This has the potential to bias inferences made from such analyses, as there is significant covariance between many common imaging parameters (Uddin et al., 2019). **Various multivariate approaches** that can overcome this issue exist, including multivariate linear regression (Young et al., 2010; Naylor et al., 2014), machine-learning (e.g., Guberman et al., 2022; Carbonell et al., 2020), and Hotelling's T^2 test (Avants et al., 2008). However, many of these approaches (including multivariate linear regression and machine learning) are complicated to implement and computationally expensive (Gyebnár et al., 2019). The Hotelling's T^2 test, a multivariate extension of a two-sample t-test, is a simple yet powerful option for group comparisons (Hotelling, 1947; Avants et al., 2008), but provides little insight at the individual level (Guberman et al., 2022).

Here we propose using the **Mahalanobis distance** (D_2 ; Mahalanobis, 1936), which is closely related to Hotelling's T^2 , but provides an individual-level measure of deviation relative to a reference distribution. D_2 is defined as the multivariate distance between a point and a distribution in which covariance between features (i.e., imaging metrics) is accounted for. Initially developed by P. C. Mahalanobis in 1936 to quantify racial similarities based on anthropometric measurements of skulls (Mahalanobis, 1927), D_2 can be thought of as a multivariate z-score where the covariance between features is accounted for (Taylor et al., 2020). An example of the importance of taking covariance into account is shown in Figure 1. One can imagine calculating the Euclidean distance between a point and the centroid of a distribution (see Fig. 1a). Both points (A and B) have the same Euclidean distance from the centroid. However, we can intuitively see that point A (Fig. 1a) is more of an extreme value relative to the distribution than point B. When computing D_2 , the shape of the distribution is taken into account such that the values that are more likely (relative to the distribution) have lower distance values. In other words, D_2 tells us how improbable a certain combination of features is. For example, because height and weight are highly correlated, very tall individuals that have a very low weight would appear outside of the distribution (e.g., Fig. 1a, point A) and would have a high D_2 value. Relationships between variables are thus accounted for in the D_2 framework by dividing the Euclidean distance by the covariance matrix, which also scales the variables to have unit variance (Fig. 1b).

The Mahalanobis distance approach has been used extensively in outlier detection, cluster analysis, and classification applications (e.g., Kritzman & Li, 2010; Xiang et al., 2008; Ghorbani, 2019). D_2 has also previously been used in neuroimaging, mainly in the study of various disorders, to detect lesions (Gyebnár et al., 2019; Lindemer et al., 2015) or to evaluate the degree

of abnormality in the brains of patients relative to controls (Dean et al., 2017; Owen et al., 2021; Taylor et al., 2020), and to study healthy WM development (Kulikova et al., 2015). Despite promising implementations and its high versatility, D2 has not yet been widely adopted. To facilitate its use, we present here an open-source tool for computing D2 relative to a reference group or within a single individual: the MultiVariate Comparison (MVComp) toolbox. In this paper, we provide a step-by-step guide to computing D2 using the MVComp tool for two example cases: a) voxel-wise comparisons between a subject and a reference group and b) within-subject comparisons between voxels. Lastly, the results of these example cases are presented and the general approach is discussed.



b)

$$D^2 = (x - m)^T C^{-1} (x - m)$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ \dots \\ x_n \end{bmatrix} - \begin{bmatrix} m_1 \\ m_2 \\ m_3 \\ m_4 \\ \dots \\ m_n \end{bmatrix}$$

Inverse of covariance matrix

*where n = number of metrics

Figure. 1. Schematic representation of the Mahalanobis distance concept for two dimensions. (a) We can identify two points (A and B) at the same distance from the distribution centroid. This illustrates the need to account for the amount of covariance between features in multivariate analyses. If covariance is not accounted for, both points would have the same Euclidean distance from the centroid of the distribution, even though one point is a clear outlier (A), as it lies outside

the distribution, while the other is not (B). (b) The Mahalanobis distance (D_2) is computed by calculating the distance of a vector of data (e.g., the data of one subject; x_1, x_2, \dots, x_n) from the mean (e.g., the group average data; m_1, m_2, \dots, m_n) and multiplying by the inverse covariance matrix (C^{-1}).

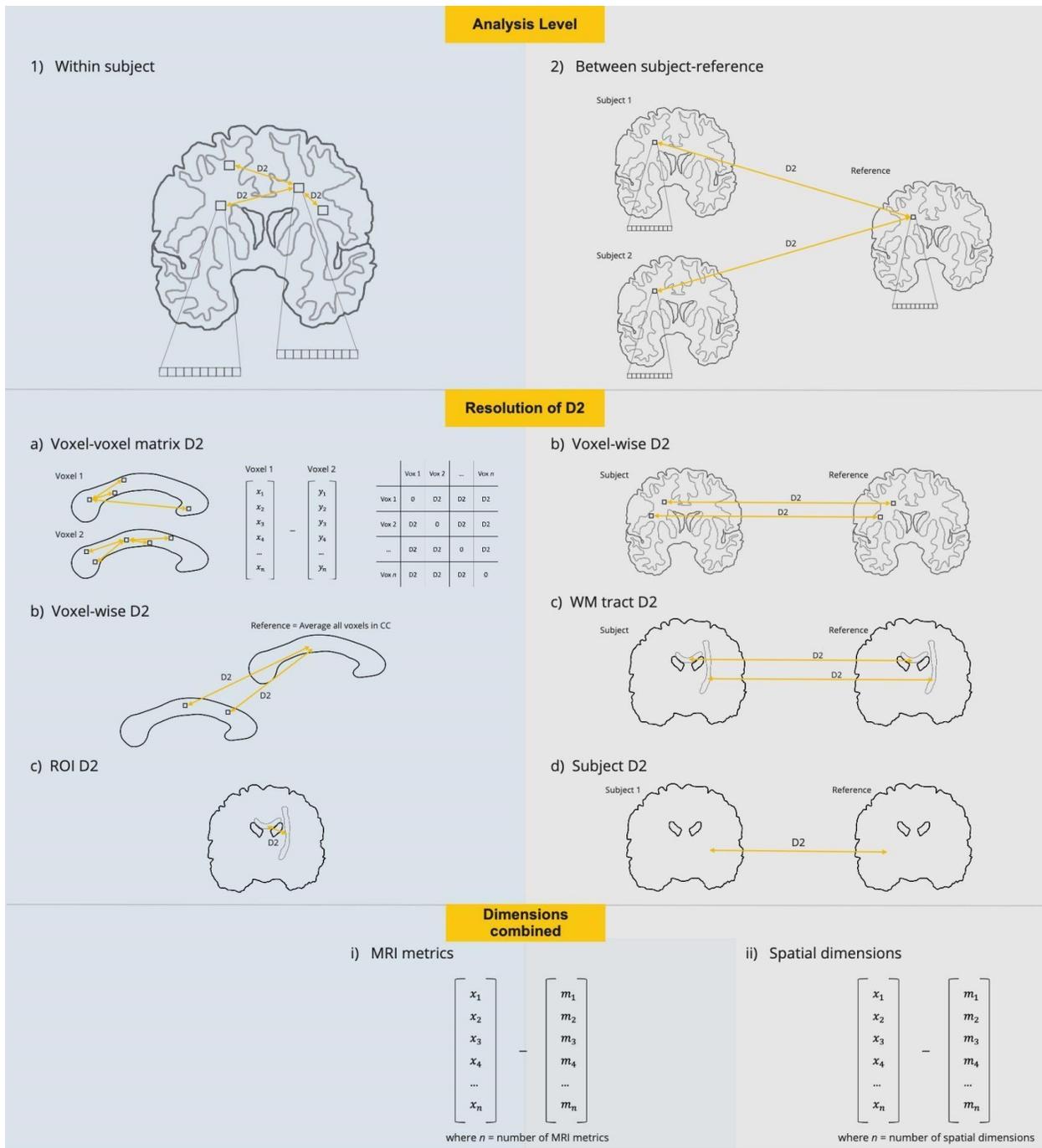


Figure. 2. Implementations of the D2 framework in neuroimaging studies. **Analysis level:** (1) Within-subject (left panel, in light blue): D2 can be computed between different voxels or brain regions (e.g., WM tracts) within a single subject. (2) Between subject-reference (right panel, light gray): D2 can be computed between a subject and a reference group (e.g., control group).

Resolution of D2: (a) Voxel-voxel matrix D2: For each voxel contained in a mask of analysis, we can compute the multivariate distance (D2) between that voxel and all other voxels, resulting in a D2 matrix of size n voxels \times n voxels (only applicable to within-subject analyses). (b) Voxel-wise D2: A D2 value can be computed at each voxel, resulting in a D2 value per voxel. (c) ROI D2: In this case, a D2 value is obtained for each WM tract, or other brain region (ROI) defined by the user. (d) Subject D2: A single D2 value can be obtained per subject, resulting in a measure of global brain deviation from the reference (only applicable to between subject-reference analyses).

Dimensions combined: (i) MRI metrics: the dimensions combined through D2 can be MRI metrics. In this case the length of the vector of data is the number of metrics. (ii) Spatial dimensions: If WM tracts, or other parcellated brain regions, are combined through D2, the length of the vector of data is equal to the number of WM tracts (only applicable to between subject-reference analyses).

2. Method

2.1 General framework

Since D2 can be defined relative to virtually any reference of matching features, MVComp has been designed to support a wide range of flexible analysis approaches. The first step is to define the set of multivariate data that will serve as the reference for computing distance. This choice depends on the hypothesis of interest, and is defined at the *Level of Analysis* as either the individual or group level (Fig 2). D2 can be computed between different brain regions within an individual (with the individual's data also serving as the reference) or between an individual and a group. In each case, multiple different *Resolutions* of analysis are possible, including voxel-wise and region of interest- (ROI) based comparisons.

To ensure that each subject's data will not bias their D2 values in single sample designs (i.e., where the entire sample is used as a reference) and to allow the evaluation of controls in two-sample designs, a leave-one-subject-out approach is also possible. In this way, the subject under evaluation is excluded from the group mean and covariance matrix prior to calculating D2.

Lastly, the choice of which dimensions to combine, either MRI-derived metrics or brain regions (e.g., WM tracts), depends on what we want to capture. Combining brain regions within a multivariate measure allows to capture the degree of deviation from a reference even in the presence of high spatial heterogeneity (e.g., Owen et al., 2020; Taylor et al., 2020), while combining features is useful in the presence of mechanistic heterogeneity (i.e., several concomitant underlying biological mechanisms) and when preserving regional specificity is desirable (e.g., Lindemer et al., 2015; Gyebnár et al., 2019). Once the level of analysis, resolution, and dimensions combined are determined, the set of multivariate data that will serve as the reference for computing distance should become clear. The examples below, which

illustrate the flexibility of the D2 approach, should make this even clearer:

2.1.1 Comparisons between subject(s) and a reference

Example 1 : White matter (WM) microstructural differences

Data: Diffusion MRI (dMRI) in one group

Level of Analysis: Between subjects, relative to group

Feature Resolution: Voxel-wise (all WM voxels)

Feature Dimensions: dMRI-derived metric maps (metrics)

In this example the reference would be defined as the voxel-wise group average for each dMRI-derived metric (m_1, m_2, m_n , where n is the number of metrics) and since the *resolution* is voxel-wise, D2 is computed by comparing the feature values in each voxel of an individual to the reference (Fig. 3a-c). The resulting individual difference D2 maps can then be entered into second-level analyses to, for example, identify brain-behavior associations. If two groups are being analyzed (e.g., patients vs controls) the control group could be used to define the reference and D2 values would then represent voxel-wise multivariate distance from controls.

Example 2 : White matter (WM) microstructural differences

Data: Diffusion MRI (dMRI) in one group

Level of Analysis: Between subjects, relative to group

Feature Resolution: Mean FA in ROI

Feature Dimensions: WM tracts (*spatial locations*)

A single MRI metric can also be used and combined across multiple ROIs to serve as the feature dimensions (e.g., pre-defined WM tracts). If the reference is defined as the group mean of each tract (m_1, m_2, m_n , where n is the number of tracts), a single D2

value would be computed for each individual, representing how the combination of mean values within all tracts differ from the equivalent combination in the reference.

2.1.2 Comparisons within a single subject

Example 3: Lesions in the context of normal appearing white matter (NAWM)

Data: Diffusion MRI (dMRI) in one group

Level of Analysis: within an individual, relative to self

Feature Resolution: *voxel-wise within ROIs (all WM voxels, two ROIs)*

Feature Dimensions: *dMRI-derived metric maps (metrics)*

Here, the level of analysis is within-subject, the dimensions combined are multiple MRI-derived metrics in each voxel, and the reference is the average of all voxels within a region of interest (ROI) for each metric. To investigate the distance between WM lesions and NAWM, the reference would be the average of all NAWM voxels (m_1, m_2, m_n , where n is the number of metrics) and D2 would be computed for each voxel classified as a lesion. The *feature resolution* could also be changed to ROI, depending on whether the user wishes to obtain a D2 value for each voxel within lesion sites or a single D2 value per ROI. This within-subject approach can also be used to compute D2 between all WM voxels and a reference ROI (e.g., a voxel in cortico-spinal tract) (Fig. 3d). Voxels within the same WM tract as the reference ROI are likely to have lower D2 than voxels in other tracts or in areas of crossing fibers (Fig. 3e).

Example 4 :

Feature Resolution: *voxel-wise*

Feature Dimensions: *dMRI-derived metric maps (metrics)*

D2 can be calculated between every pair of voxels to compute a voxel-voxel D2

matrix by directly comparing across the feature dimensions. In this case, the reference for computing the covariance matrix would be the data in all voxels contained in the analysis mask.

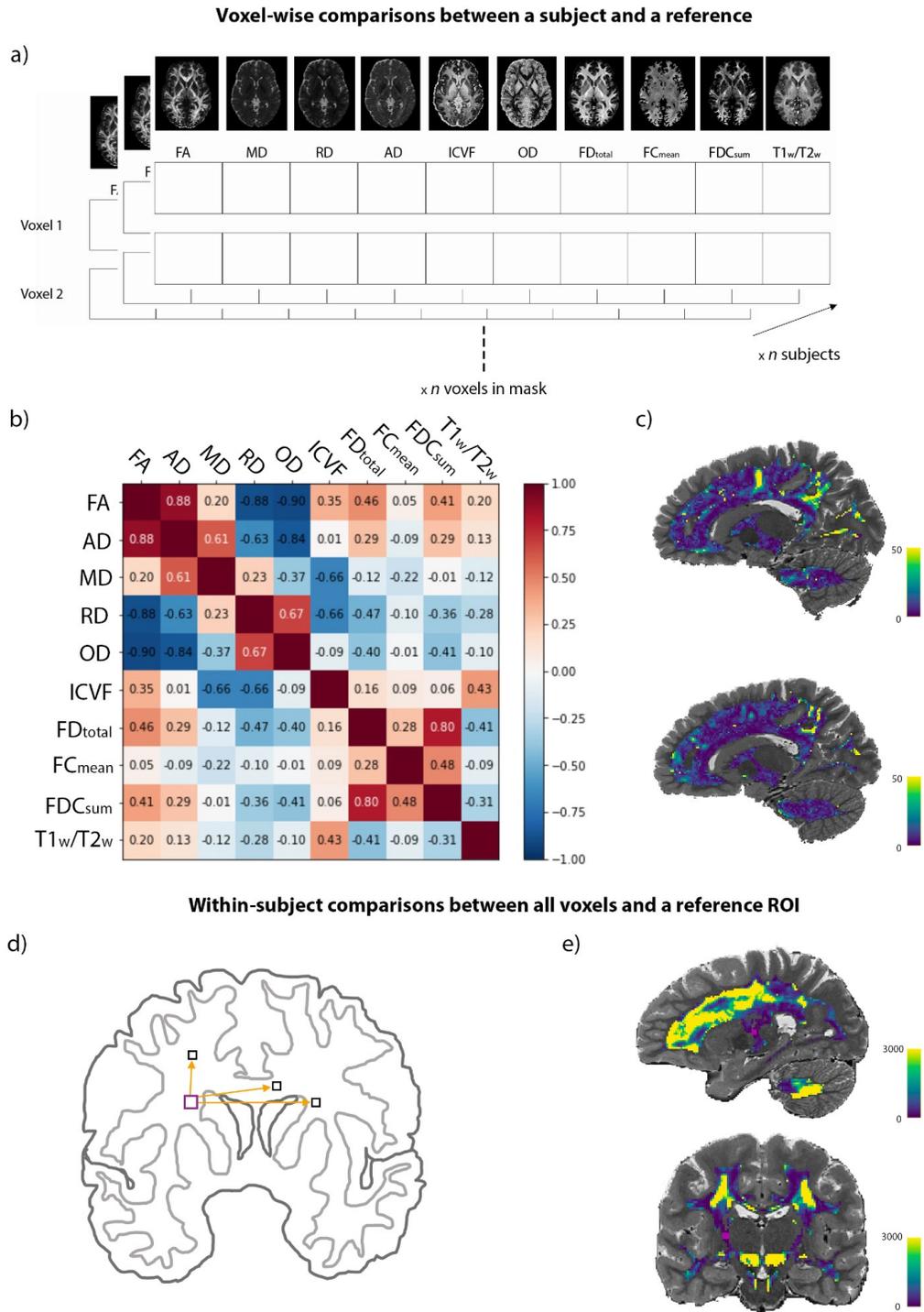


Fig. 3. D2 workflow. **Voxel-wise comparisons between a subject and a reference.** (a) The multivariate space is illustrated here. In this example, we have a vector of 10 MRI metrics at each WM voxel for each subject. (b) The covariance matrix is computed from the reference

feature matrix of shape n voxels \times n features. The plot shows the amount of correlation between features in the reference sample. (c) Voxel-wise D2 maps in two example subjects, where bright yellow represents areas of greater deviation from the reference population. Distinct patterns can be seen in the two subjects. **Within-subject comparisons between all voxels and a reference ROI.** (d) Schematic representation of the multivariate comparisons showing that D2 was computed between all WM voxels and a ROI of 24 voxels in the corticospinal tract (CST). (e) D2 map showing the multivariate distance between all voxels and the CST ROI (in pink).

2.2 Data preparation

In all cases, data for all subjects should be preprocessed and all MRI metrics of interest computed and transformed to bring them into the same voxel space. Masks should also be generated to restrict analyses to chosen regions (e.g., white matter) and be transformed into the same space. Masks can be binary or thresholded at a later step. Moreover, in group analysis it is necessary to compute the average of each metric from the reference group (*mvcomp.compute_average* is included as a convenience function to perform this task). A leave-one-out approach (where the individual to be compared to the standard is left out of the average) is preferred in cases where the individual subject is also a member of the reference group. This functionality is directly available in the model comparison function (*model_comp*).

Combining spatial dimensions (Feature dimension: Regions of interest)

Summary metrics should be calculated from each region of interest (e.g., mean FA in each WM tract of interest) for each subject. Creating a group average is not necessary for this application since the group average of the reference is computed within a subsequent function (*spatial_mvcomp*).

2.3 Computing the reference standard and covariance matrix

In the case of group analyses, the reference standard mean and covariance matrix are derived from either multiple features or multiple regions of interest in another group (e.g., control group). The comparison can also be between each individual and the mean of all other individuals if only a single group is available. In the case of analyses within an individual, multiple features across all voxels or a specific region of interest can be used.

2.3.1 Comparisons between subjects and a reference

Combining MRI metrics

The *mvcomp.feature_gen* extracts the feature matrix from a set of input images. Run on the reference group mean images, it returns the feature matrix (*m_f_mat* of shape n voxels \times n features) and a mask vector (*mat_mask* of shape n voxels). The mask array contains zeros at voxels where values are *nan* or *inf* for at least one of the reference average maps in addition to the voxels below the threshold set for the mask. The *mvcomp.norm_covar_inv* function is then used to compute the covariance matrix (*s*) and its pseudoinverse (*pinv_s*) from the reference feature and mask matrices (*m_f_mat* and *mat_mask*). The *mvcomp.correlation_fig* function can be used to generate a correlation matrix from the covariance matrix (*s*), which is informative to verify if expected relationships between features are present. If the leave-one-out approach is used, *mvcomp.feature_gen* is not necessary as the covariance matrix will be computed within the *model_comp* function from a group average that excludes the subject for which D2 is being computed.

Combining spatial dimensions

The comparison mean values and covariance matrix are computed within the *spatial_mvcomp* function described in detail below.

2.3.2 Within-subject comparisons

Voxel-wise D2 resolution

If the reference ROI is a set of voxels in the CST, the covariance matrix will be computed based on all voxels within that ROI in that subject. The path of the images (i.e., one image per metric) can be provided to the *feature_gen* function, along with the ROI mask, to create the reference feature matrix (*m_f_mat*) and mask vector (*mat_mask*). The *mvcomp.norm_covar_inv* function is then used to compute the covariance matrix (*s*) and its pseudoinverse (*pinv_s*) from the feature and mask matrices. The *mvcomp.correlation_fig* function can again be used to

visualize relationships between metrics.

Voxel-voxel matrix D2 resolution

For this approach, the covariance matrix is computed from a feature matrix that includes all voxels in the mask of analysis. For instance, if we are interested in computing D2 between each voxel and all other voxels in the whole WM, the covariance matrix is based on all WM voxels. Therefore the matrix provided to the *norm_covar_inv* function will be of shape n voxels in the mask x n features.

2.4 Computing D2

2.4.1 Comparisons between subjects and a reference

Combining MRI metrics

The *mvcomp.model_comp* function allows the calculation of voxel-wise D2 between each subject contained in the provided `subject_ids` list and the reference (group average). The user should specify the directories and suffix of the subjects' features and reference images (`feature_in_dir`, `model_dir`, `suffix_name_comp` and `suffix_name_model`), the mask of analysis (`mask_f`) and a threshold if the mask is not binary (`mask_threshold`). If subjects or features are to be excluded at this point, they can be specified with the `exclude_subject_ids` and the `feat_sub` options, respectively. If the leave-one-out approach is to be used, the `exclude_comp_from_mean_cov` option should be set to `True`, in which case the `model_dir` is not necessary. If this option is set to `True`, the mean (reference) and `pinv_s` are computed on each subject comparison, excluding the subject being compared before computing its D2.

Combining spatial dimensions

The *spatial_mvcomp* function is used to compute D2 between each subject and the reference computed from all subjects. A matrix containing the data (e.g., mean FA in each WM

tract) of all subjects (n subjects \times n tracts) should be provided to the function. The *spatial_mvcomp* function returns a vector with a single D2 value per subject. As in *model_comp*, setting the *exclude_comp_from_mean_cov* to True leaves out the current subject when computing the mean and covariance.

2.4.2 Within-subject comparisons

Voxel-wise D2 resolution

The *mah_dist_mat_2_roi* function is used to compute voxel-wise D2 between all voxels and a specific ROI. Here, in addition to the feature matrix containing the data for the voxels to be evaluated (n voxels \times n features), the user will need to provide a vector of data for the reference ROI (i.e., mean across voxels in the ROI for each metric) and the inverse of the covariance matrix (*pinv_s*).

Voxel-voxel matrix D2 resolution

The *mah_dist_mat_connectivity* function is used to compute D2 between each voxel and all other voxels in a mask. This yields a symmetric 2-D matrix of size n voxels \times n voxels containing D2 values between each pair of voxels.

2.5 Statistical analysis

2.5.1 Comparisons between subjects and a reference

A two-samples t-test on D2 values would then allow group comparisons, in a similar manner as a Hotelling's T^2 test, or alternatively a statistical method such as the Bhattacharyya coefficient can be used to estimate the degree of overlap between the distribution of each group (where less overlap indicates a higher probability that the groups differ) as in (Dean et al., 2017). However, such group analyses are likely to average out interindividual variability and be problematic when heterogeneity is high (Guberman et al., 2022).

2.5.2 Within-subject comparisons

Clustering approaches can be applied to the voxel-voxel matrix D2 to parcellate brain voxels into networks/subdivisions. Changes in D2, either from the group or subject-level, can also be assessed through longitudinal analyses, to investigate WM damage progression or brain maturation for instance (e.g., Lindemer et al., 2015; Kulikova et al., 2015). D2, or changes in D2, can also be related to behavioral outcomes (e.g., cognitive score, performance on a skill test, or symptom severity) in the same way one would with univariate measures of fractional anisotropy for instance (Owen et al., 2020; Dean et al., 2017; Taylor et al., 2020).

2.6 Determining feature importance

2.6.1 Comparisons between subjects and a reference

Combining MRI metrics

If the user is interested in understanding the physiological mechanisms underlying microstructural deviations in a region of interest (e.g., voxels where D2 is high), the *return_raw* option of the *mvcomp.model_comp* function can be used. This allows the extraction of each features' weight in D2. If *return_raw* is set to True, the function returns a 3D array of size (number of voxels) x (number of features) x (number of subjects) that contains the voxel-wise distances for each feature and each subject. A mask of the region of interest (e.g., a region of high D2) can then be applied to the 3D array and the distances can be summarized across voxels and/or subjects to obtain a percent contribution to D2 for each feature within that region.

Combining spatial dimensions

The *return_raw* option is also available in the *spatial_mvcomp* function. If set to True, a 2D array of size (number of subjects) x (number of tracts) containing the distances between every subject's tract and the mean tract values is returned. These *raw* distances provide

information regarding the contribution of each WM tract to D2, which gives insights on the localization of greatest deviation for each subject.

2.6.2 Within-subject comparisons

Voxel-wise D2 resolution

The *return_raw* option of the *mah_dist_mat_2_roi* function can be used to extract features' contributions. In this case, the distances between features in all voxels being compared and feature values in the ROI are returned. The output will be of shape (number of voxels) x (number of features).

Voxel-voxel matrix D2 resolution

The *return_raw* option of the *mah_dist_mat_connectivity* function can be used to extract features' contributions. A matrix of shape (number of voxels) x (number of voxels) x (number of features) is returned, with the contribution of each feature to the voxel-voxel D2 positioned on the last axis.

2.7 Experiments

2.7.1 Data Description

We computed 10 microstructural features for 1001 subjects from the Human Connectome Project S1200 data release (Van Essen et al., 2013) for these experiments. DWI, T1- and T2-weighted data were acquired using a custom-made Siemens Connectom Skyra 3 Tesla scanner with a 32-channel head coil. The DWI data (TE/TR=89.5/5520 ms, FOV=210×180 mm) were multi-shell with b-values of 1000, 2000 and 3000 s/mm² and a 1.25 mm isotropic resolution, 90 uniformly distributed directions, and 6 b=0 volumes. T1-w data was acquired with a 3D-MPRAGE sequence and T2w images with a 3D T2-SPACE sequence, both with a 0.7mm isotropic resolution (T1w: 0.7 mm iso, TI/TE/TR=1000/2.14/2400 ms, FOV=224×224 mm;

T2w: 0.7 mm iso, TE/TR=565/3200 ms, FOV=224×224 mm). Anatomical scans were acquired during the first session, and DWI data were acquired during the fourth session. More details on the acquisitions can be found at: <https://www.humanconnectome.org/hcp-protocols-ya-3t-imaging>. The imaging data of 1065 young healthy adults, those who had undergone T1w, T2w and diffusion-weighted imaging, were preprocessed. The data of 64 participants were excluded due to poor cerebellar coverage.

2.7.2 Preprocessing

Diffusion Tensor Imaging

The minimally preprocessed HCP data was used (Van Essen et al., 2013; Glasser et al., 2013). The minimal preprocessing pipeline for DWI data includes intensity normalization of the b_0 images, eddy current and susceptibility-induced distortions correction, using DWI volumes of opposite phase-encoding directions, motion correction and gradient nonlinearity correction. DWI data were registered to native structural space (T1w image), using a rigid transform computed from the mean b_0 image, and diffusion gradient vectors (bvecs) were rotated accordingly.

Most subsequent processing steps were performed using the MRtrix3 toolbox (Tournier et al., 2019). The minimally preprocessed DWI data was converted to the mif format, with the bvals and bvecs files embedded, after which a bias field correction was performed using the ANTs algorithm (N4) of the `dwibiascorrect` function of MRtrix3 (Tustison et al., 2010). The tensor was computed on the bias field-corrected DWI data (using `dwi2tensor`) and DTI metrics were then calculated (FA, MD, AD and RD) using `tensor2metric` (Basser & LeBihan, 1994; Veraart et al., 2013; Basser, Mattiello & LeBihan, 1994).

Multi-tissue Multi-shell Constrained Spherical Deconvolution

The multi-tissue Constrained Spherical Deconvolution (CSD) was performed following

the fixel-based analysis (FBA) workflow

(https://mrtrix.readthedocs.io/en/latest/fixel_based_analysis/mt_fibre_density_cross-section.html). The T1-w images were segmented using the *5ttgen* FSL function of MRtrix3, which uses the FAST algorithm (Smith et al., 2012; Smith, 2002; Zhang et al., 2001; Patenaude et al., 2011; Smith et al., 2004). Response functions for each tissue type were then computed from the minimally preprocessed DWI data (without bias field correction) and the five-tissue-type (5tt) image using the *dwi2response* function (*msmt_5tt* algorithm) (Jeurissen et al., 2014). The bias-uncorrected DWI data was used because bias field correction is performed at a later step in the FBA pipeline (Raffelt et al., 2017). The WM, GM and CSF response functions were then averaged across all participants, resulting in a single response function for each of the three tissue types. Multi-shell multi-tissue CSD was then performed based on the response functions to obtain an estimation of orientation distribution functions (ODFs) for each tissue type (Jeurissen et al., 2014). This step is performed using the *dwi2fod msmt_csd* function of MRtrix3 within a brain mask (i.e., *nodif_brain_mask.nii.gz*). Bias field correction and global intensity normalisation, which normalises signal amplitudes to make subjects comparable, were then performed on the ODFs, using the *mtnormalise* function in MRtrix3 (Raffelt et al., ISMRM, 2017; Dhollander et al., ISMRM, 2021).

Registration

In order to optimize the alignment of WM as well as gray matter, multi-contrast registration was performed. Population templates were generated from the WM, GM and CSF FODs of a subset of 200 participants using the *population_template* function of MRtrix3 (with regularisation parameters: *nl_update_smooth*= 1.0 and *nl_disp_smooth*= 0.75; including the “nodif” brain masks), resulting in a group template for each of the three tissue types (Tournier et

al., 2019).

Subject-to-template warps were computed using *mrregister* in MRtrix3 with the same regularisation parameters and warps were then applied to the brain masks, WM FODs, DTI metrics (i.e., FA, MD, AD and RD), T1w, and T2w images using *mrtransform* (Raffelt et al., 2011). T1w and T2w images were kept in native resolution (0.7mm) and the ratio of T1w/T2w was calculated to produce a myelin map (Glasser & Van Essen, 2011). WM FODs were transformed but not reoriented at this step, which aligns the voxels of the images but not the *fixels* (“*fibre bundle elements*”). A template mask was computed as the intersection of all warped brain masks (*mrmath min* function). This template mask includes only the voxels that contain data in all subjects. The WM volumes of the five-tissue-type (5tt) 4-D images were also warped to the group template space since these will be used to generate a WM mask for analyses.

Computing fixel metrics

The WM FOD template was segmented to generate a *fixel* mask using the *fod2fixel* function (Smith et al., 2013; Raffelt et al., 2012). This mask determines the fiber bundle elements (i.e., *fixels*), within each voxel of the template mask, that will be considered for subsequent analyses. The *fixel* mask is stored in a specific file format which embeds information such as the number and direction of each *fixel* in every voxel. *Fixel* segmentation was then performed from the WM FODs of each subject using the *fod2fixel* function. The apparent fibre density (FD) of each *fixel* was also computed in this step. Since reorientation was not performed in the registration step above, the FODs of each subject are in the FOD template space voxel-wise, but the *fixels* are not aligned yet at this point. The *fixelreorient* function is thus used to align the *fixels* of each subject’s FOD to those of the template, based on the subject-to-template warps

applied previously. Next, each subject's *fixels* were explicitly mapped to the corresponding *fixels* in the fixel mask using the *fixelcorrespondence* function. The FD data of each subject is thus expressed with respect to a single set of *fixels* common to all subjects (i.e., with a single set of *fixels*' directions). The fibre bundle cross-section (FC) metric was then computed from the warps generated during registration (using the *warp2metric* function) as FC is a measure of how much a fiber bundle has to be expanded/contracted for it to fit the fiber bundles of the *fixel* template. Lastly, a combined metric, fibre density and cross-section (FDC), representing a fibre bundle's total capacity to carry information, was computed as the product of FD and FC.

Transforming fixel metrics into voxel space

In order to integrate all metrics into the same multi-modal model, *fixel* metric maps were transformed into voxel-wise maps. As a voxel aggregate of fiber density, we chose to use the $l=0$ term of the WM FOD spherical harmonic expansion (i.e., 1st volume of the WM FOD, which is equal to the sum of FOD lobe integrals) to obtain a measure of the total fibre density (FD_{total}) per voxel since this was shown to result in more reproducible estimates than when deriving this measure from fiber specific FD (i.e., by summing the FD *fixel* metric) (Calamante et al., 2015). The FOD $l=0$ term was scaled by the spherical harmonic basis factor (by multiplying the intensity value at each voxel by the square root of 4π).

For the fiber cross-section voxel aggregate measure, we opted for computing the mean of FC, weighed by FD, using the mean option of the *fixel2voxel* function. We thus obtained the typical expansion/contraction necessary to align fiber bundles in a voxel to the *fixels* in the template. Since this is weighted by FD, bundles with higher density have a greater influence on voxel-wise FC value than lower density bundles.

Lastly, the voxel-wise sum of FDC, reflecting the total information-carrying capacity at

each voxel, was computed using the *fixel2voxel sum* option.

NODDI metrics

Bias field corrected DWI data was fitted to the neurite orientation dispersion and density imaging (NODDI) model using the python implementation of Accelerated Microstructure Imaging via Convex Optimization (AMICO) (Daducci et al., 2015; Zhang et al., 2012). First, small variations in b values were removed by assigning the closest target bval (0, 1000, 2000 or 3000) to each value of the bvals file (Github link of Chris function). This is to prevent the fitting algorithm from interpreting every slightly different bval as a different diffusion shell. A diffusion gradient scheme file is then created from the bvecs, and the new bvals file. The response functions are computed for all compartments and fitting is then performed on the unbiased DWI volumes, within the non-diffusion weighted brain mask (*nodif_brain_mask.nii.gz*). The resulting parameters obtained are: the intracellular volume fraction (ICVF, also referred to as neurite density), the isotropic volume fraction (ISOVF), and the orientation dispersion index (OD). In this study, we will use ICVF and OD.

Generating masks for analyses

The maps of each of the 10 metrics of interest (FA, AD, RD, MD, T1w/T2w, FDtotal, FCmean, FDCsum, ICVF and OD) were then averaged across all subjects. These average maps served as the reference. A WM mask was created by computing the group average of the corresponding volume of the T1 5tt image (volume 2). A threshold of 0.99 was applied within the MVComp toolbox's functions.

2.7.3 Experiment 1: Comparisons between subjects and a reference

Here, we present an example case of using D2 in a large sample from the HCP dataset to quantify voxel-wise microstructural differences in WM according to several MRI metrics. Since

the dataset used in this study contains the data of healthy young adults, a relatively homogeneous population, the entire sample was set as the reference and the leave-one-out approach was used to exclude the subject under evaluation. The analysis was restricted to the corpus callosum (CC). Voxel-wise D2 values were computed in the CC for each subject from 10 microstructural features, yielding a D2 matrix of 1001 subjects X 2845 voxels. The D2 values represent voxel-wise microstructural distances in an individual's CC relative to the group average, while accounting for the covariance between features. Large D2 scores in a voxel indicate greater deviation from the group average, whereas scores closer to 0 indicate lower distance (i.e., more typical microstructure). Due to the large number of datapoints and potential effects of partial voluming, we observed several outliers in D2 maps of several subjects . We therefore excluded participants with at least 50 voxels that were deemed as outliers (i.e. exceeded a threshold of 5 standard deviations from the voxel mean D2). This yielded a final sample of 723 participants.

Past literature on CC neuroanatomy shows several segments that are distributed along the anterior to posterior axis, where each segment is defined by common microstructural properties and/or connectivity profile (Aboitiz et al., 1992; Chao et al., 2009; Hofer & Frahm, 2006). We therefore hypothesized that D2 values in the CC could extract these segments via unsupervised machine learning. We performed K-means clustering on the D2 matrix of size subject X voxel, setting the number of clusters to 9 based on literature on CC topography (Aboitiz et al., 1992; Chao et al., 2009; Hofer & Frahm, 2006). Prior to clustering, we applied z-score and power transformation on the D2 matrix to achieve gaussian distributions of standardized scores. Final visualization was done using BrainNet Viewer.

2.7.4 Experiment 2: Within-subject comparisons

The within-subject approach allows the computation of voxel-voxel D2 in a single

individual from multiple microstructural features. Here, D2 was calculated between each voxel and every other voxel in a subject's CC, while accounting for the covariance between the 10 microstructural features. All voxels within the CC of that subject are used to compute the covariance matrix and this same covariance matrix is used in the D2 calculation of every voxel. The resulting D2 matrix is a 2845 voxel X 2845 voxel dense matrix representing the distance between each voxel and every other voxel in the CC. We standardized the matrix to z-scores, and then applied Principal component analysis (PCA). We extracted the first principal component explaining the highest variance, and the contributions of each metric to D2 were extracted within the voxels with the largest and the lowest scores on the first principal component.

3. Results

3.1 Experiment 1: Comparisons between subjects and a reference

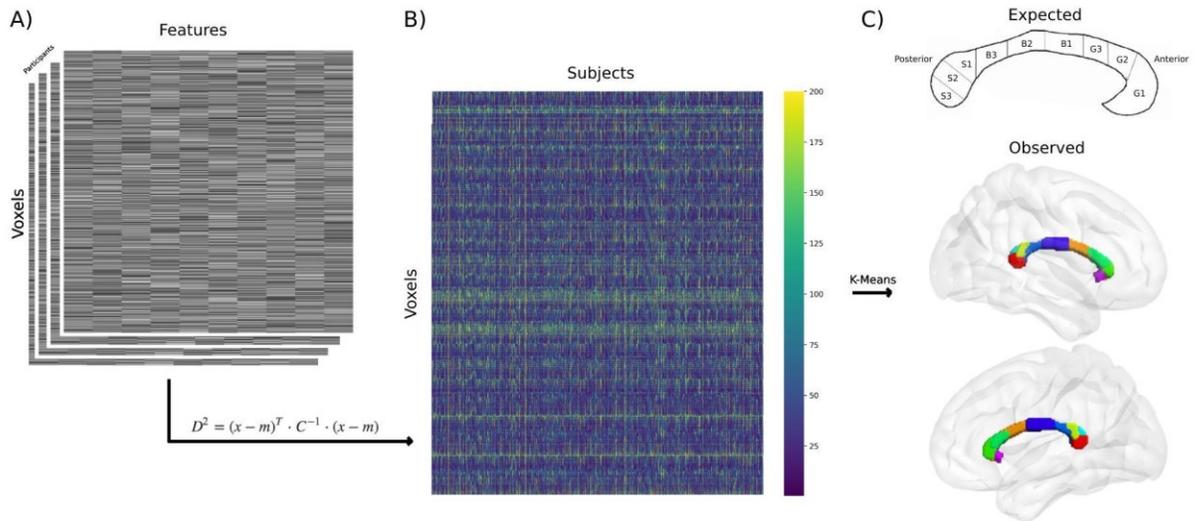


Fig. 4. Voxel-wise comparisons between each subject and the reference. (a) Voxel-wise D^2 is calculated between the reference (group average of the whole sample, except the subject under evaluation) and each subject's data (feature (10) X voxel (2845) matrix), in voxels of the the corpus callosum (CC). (b) This results in a D^2 matrix of size subject (1001) X voxel (2845) containing the multivariate distance between a subject's data and the reference at each CC voxel for each of the 1001 subjects. (c) Applying k-means clustering to the D^2 matrix, voxels of the CC were partitioned into 9 clusters distributed along the anterior-posterior axis, in close accordance with known topography of the CC as seen in (d). (d) Schematic representation of CC topography based on literature (Aboitiz et al., 1992; Chao et al., 2009; Hofer & Frahm, 2006).

Voxel-wise D2 was computed between each subject (feature X voxel matrix) and the reference (feature X voxel matrix of the reference group average). This computation was repeated for each voxel within the CC. K-means clustering was then applied to the resulting subject X voxel D2 matrix. We observed 9 clusters distributed along the anterior-posterior axis in accordance with past evidence on CC microstructure and connectivity (Aboitiz et al., 1992; Chao et al., 2009; Hofer & Frahm, 2006). Fig. 4b shows the clusters identified via k-means. The Genu of the CC is clustered into 3 segments, while the midbody displays 2 segments. The splenium was also clustered into 4 segments (with one segment positioned on the isthmus).

3.2 Experiment 2: Within-subject comparisons

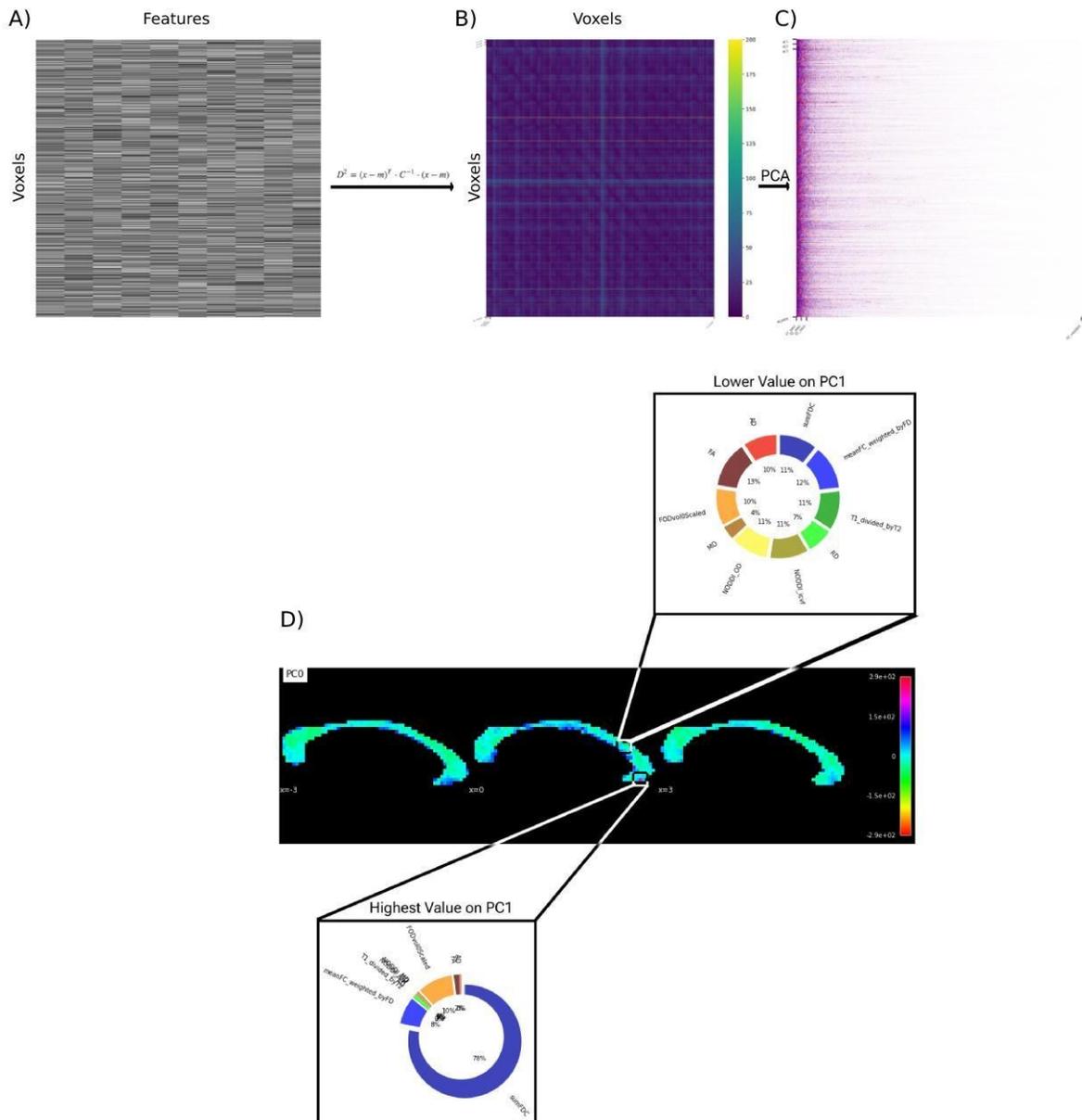


Fig. 5. Within-subject voxel-voxel comparisons. (a) D2 is computed from the features x voxels matrix of a subject and results in a (b) voxel x voxel D2 matrix. (c) PCA was applied to the D2 matrix. (d) The metrics's contributions to D2 are extracted in the voxels with the highest and lowest score on PC1. SumFDC contributed most to D2 in the voxel with the highest value, located in the genu of the CC. In the voxel with the lowest value on PC1, located in the midbody

of the CC, all metrics had approximately equal contribution to D2.

D2 was computed between each and every pair of voxels in a subject's CC as a case example of the within-subject level of analysis. We applied PCA on the resulting voxel X voxel D2 matrix and extracted metrics' contributions to D2 within the voxels with the largest and the lowest scores on the first principal component. Fig. 5 shows the steps involved in the within-subject analysis. The first principal component explained 95% of the variance in the voxel X voxel dense D2 matrix. The largest and the lowest scores were in the genu and in the midbody of the CC, respectively (Fig. 5d). We then extracted metrics' contributions in these voxels. In the voxel with the largest value on PC1, the fibre density and cross-section metric (sumFDC) contributed most to D2, while mean diffusivity (MD) contributed the least. On the other hand, in the voxel with the lowest score on PC1, all microstructural features had nearly equal contributions to D2, indicating minimal variability in this voxel.

4. Discussion

White matter (WM) tracts have long been viewed as passive passageways, serving as routes for information transmission between neurons, to which were attributed most, if not all, the credit in higher-order functions. Since the 1965 seminal paper by Geschwind in which WM lesions were associated to neurobehavioral syndromes (Geschwind, 1965), the importance of WM to normal cognitive and affective functions has started to be acknowledged and it is now widely recognized by the scientific community (Filley, 1998; Fields, 2008a, 2008b). Early studies on disconnection of white matter and their role in neurobehavioral syndromes (Geschwind, 1965) to more recent work on subtle alterations to WM microstructure and their role in cognition and pathologies (Fields, 2008). Moreover, work in the last decade has highlighted the plastic potential of WM and its importance in learning, not only during development, but throughout life (Scholz et al., 2009; Fields 2015; Sampaio-Baptista & Johansen-Berg, 2017). Together, these studies highlight the importance of quantitatively assessing WM to evaluate alterations induced by pathology, training or inter-individual variability in healthy populations and to investigate the relationships between these deviations in WM microstructure and behavior.

Quantifying even the most subtle abnormalities in WM health could have a profound impact on fields such as psychiatry and aging. For several psychiatric conditions, interventions are likely to be more effective if initiated at the earliest possible stage, before irreversible damage occurs and when there is likely more potential for neuroplasticity (Bonzano et al., 2014; Sachdev et al., 2013; Iturria-Medina et al., 2016, 2017). Since we do not have a cure for several neurological disorders, including Alzheimer's disease and other dementias, prevention is currently the best line of treatment. Precise quantification of WM integrity could also have

important impacts in normal aging, where preventing neurological disorders could be achieved through early detection of WM alterations in at-risk populations (Poirier et al., 2021; Anazodo et al., 2019).

Multiparametric approaches appear promising in studying pathological alterations in WM microstructure in a number of conditions. For instance, D2 incorporating fractional anisotropy (FA) in multiple WM tracts in epileptic patients was found to show stronger associations with epilepsy duration than any univariate measure (i.e., mean FA in a single WM tract) (Owen et al., 2020). Another study reported better performance using D2 encompassing FA in several WM tracts than using univariate measures in discriminating between controls and individuals with TBI (Taylor et al., 2020). Their multivariate measure allowed for the discrimination of even mild TBI cases from controls and correlated significantly with cognitive scores. Similarly, D2 combining both spatial (i.e., WM regions) and feature (i.e., different DTI metrics) dimensions showed greater delineation between autistic and typically developing individuals compared to univariate approaches or to D2 formed by combining brain regions only (Dean et al., 2017). Associations between D2 and autism symptom severity were also reported in this study, suggesting D2 served as a meaningful, behaviorally relevant, measure of WM abnormality.

Other interesting implementations have used D2 to detect and characterize lesions. Gyebnár et al. (2019) combined DTI eigenvalues into a voxel-wise D2 distance measure between epilepsy patients and controls to detect cortical malformations in patients. Voxels were identified as lesion voxels if their D2 value exceeded a critical value calculated using Wilks' criterion (Wilks, 1963), a criterion used for multivariate statistical outlier detection. In another implementation, D2 was employed to characterize the heterogeneity within WM lesions by computing the multivariate distance (combining T1-w, T2-w and PD-w signal intensities)

between voxels in regions of WM signal abnormalities (i.e., WM hyperintensities; WMSA) and those in regions of normal appearing WM (NAWM) (Lindemer et al., 2015). D2 within regions of WMSA, defined as WMSA quality by the authors, was shown to progress at a quicker rate in individuals who converted from mild cognitive impairment to Alzheimer's disease (AD) compared to those who did not convert. Moreover, the time point of greatest rate of change coincided with the time of conversion to AD. Interestingly, the rate of change of WMSA volume (i.e., lesion load), a metric more commonly used (Schmidt et al., 2005; Bilello et al., 2015), did not differentiate converters from non-converters cross-sectionally and longitudinally, suggesting that a characterization of WM lesion heterogeneity through a multivariate framework was more informative than the volume of WM lesions (Lindemer et al., 2015).

In previous work using the D2 approach, the loadings (or weights) of the elements combined in the multivariate measure (i.e., either WM tracts or MRI metrics) were not extracted and this was reported as a limitation of D2 in Dean et al. (2017). Characterizing the extent by which each element contributes to D2 can provide important insights into the physiological underpinnings of the alterations observed and/or their localization. Therefore, the MultiVariate Comparison (MVComp) tool we propose allows the extraction of features contribution to D2. Previous methods of examining neuroanatomical and microstructural differences using neuroimaging metrics have largely focused on univariate and some limited multivariate frameworks. Several of these frameworks do not offer the possibility to correct for the shared covariance between metrics, which arise from overlapping microstructural phenomena as well as from mathematical similarities between some modelling approaches. Non-quantitative nature of neuroimaging metrics

Furthermore, the widely used acquisition protocols and their parameters are non-

quantitative in nature, and hence do not tie intimately to microstructure. This stems from the mutual influence of several biological properties to the same neuroimaging metric. This also extends in the opposite direction, where multiple neuroimaging metrics indirectly reflect a similar underlying physiological property. Therefore, using a single neuroimaging metric, or metrics stemming from a single model, offers limited potential for interpretation and is biased by the set of assumptions of the chosen model (Novikov et al., 2018). For instance, some models assume fixed compartment diffusivities (e.g., NODDI) while others attempt to estimate them (white matter tract integrity (WMTI)) (Novikov et al., 2018). While some multivariate frameworks have been implemented in the neuroimaging field, several of them are either applicable at the group level or at the subject level, and do not extend from one level to another. The D2 framework on the other hand is highly versatile and fairly easy to implement. Moreover, the open-source MVComp toolbox we propose makes the implementation of D2 in various research questions (see Fig.2) even more accessible.

Here, we present MVComp, a multivariate comparisons framework that integrates multiple neuroimaging metrics while accounting for their shared covariance, applicable at either the group- or subject-levels. When applied at the group level, the clustering of D2 maps in the corpus callosum using the unsupervised K-Means clustering technique revealed high accordance between D2 and known WM microstructure properties. When applied at the subject level, principal component analysis on the dense voxel-voxel distance matrix showed extracted voxels with large and low variability in the underlying neuroimaging metrics, as could be seen in the highest and lowest values on the first principal component.

Accounting for covariance between the neuroimaging metrics is crucial when these metrics share common underlying microstructural and biophysical characteristics. Therefore, D2

was able to resolve the microstructure of the CC.

There are several advantages of using a multivariate distance via MVComp. First, The covariance between neuroimaging metrics is largely overlooked, and at its core, MVcomp accounts for the covariance while computing distances. This is done by computing the Mahalanobis distance (D2) at the level of analysis (group voxelwise, group ROI level, and within subject). The researcher hence decides on the features used in computing D2, with the possibility of integrating several diffusion models and their derived metrics. Therefore, MVComp provides a versatile framework in assessing microstructural differences in WM. Moreover, the toolbox is an easy-to-use python package with the options to compute the sample average, compute the pseudo covariance between metrics, and calculating D2 using well-documented functions. The MVComp tool also allows the extraction of raw distances, which gives information regarding the relative contribution of each metric to the multivariate distance (D2).

We show that combining multiple neuroimaging metrics while accounting for their shared covariance exhibit ties to known microstructure. By applying K-Means clustering to the corpus callosum, we observed a clear and homogeneous segmentation along the anterior-posterior axis similar to known anatomy. This high correspondence between clustered D2 and previously described CC topography indicates that the microstructural score obtained by combining several WM neuroimaging metrics while accounting for the shared covariance between metrics would provide a pertinent index encompassing several important microstructural properties. At the individual level D2 can capture the amount of (dis)similarity between voxels and through the extraction of features' contributions we can infer the physiological mechanisms underlying this extent of (dis)similarity between voxels (it can capture the metrics/features with the highest variability/variance). SumFDC was the metric that

contributed most to D2 in the voxel with highest score on PC1: indicates that the amount of *information-carrying capacity* (combined fiber cross-section and density) is the feature that varies the most in the CC. MD was the metric that contributed least to D2 in the voxel with highest score on PC1: MD would likely be very low in all those CC voxels (tightly packed white matter tracts). These conclusions could further aid in assessing brain-behaviour associations, by maximising variability while reducing redundancy (shared information).

The link between D2 and white matter extends beyond healthy microstructure. Previous studies have shown the necessity of integrating neuroimaging metrics in disease via a multivariate way, and some proposed using D2 (Owen et al., 2021; Taylor et al., 2020). Our framework expands on these findings, and provides several levels of analysis that can be applied in cases of neurodegeneration and neural injury. The aim in these cases would be to evaluate the degree of microstructural differences between a control group (serving as the model) and the comparison group (e.g. AD). Since D2 is a measure of multivariate distance from a model, higher scores would indicate larger extents of damage at the voxel or tract level. It is possible to implement this comparison at a single subject level, where regions of maximum degeneration or extensive lesion would exhibit very large D2 scores that those with healthy tissue. As such, we would expect that D2 scores would increase along a continuum from health to extensive tissue damage, and this continuum could exist from healthy ageing, to mild cognitive impairment, to AD or inversely correlated with the distance to damage location at the within-subject level.

The MVcomp toolbox is easily extendable, and double multivariate steps could be implemented. It is possible to apply D2 computation at the voxel level, integrating multiple different measures of microstructure to extract D2 scores, and then apply spatial multivariate comparison to extract D2 at the tract level. In this case, each subject in a between-subject designs

would have one D2 score, which could be used in brain-behaviour assessments or classification based on previously defined conditions. It is worth noting that in this case, spatial information is lost, and the location of highest D2 scores is masked when applying the spatial analysis.

Given the interactivity between the brain and behaviour, where changes on one could alter the other (such is the case in experience-induced plasticity), we would expect high correlation between D2 and behavioural measures. For instance, it is possible to use MVcomp to extract D2 values in motor tracts, and examine the association between microstructural differences in a given sample and motor functioning. If high motor performance is due to microstructural differences, then a positive correlation between microstructure and motor tasks is to be observed. It is also possible to investigate the relationship between behavioural impairment and the extent of neural damage in targeted regions in a spatially specific manner while integrating multiple measures of microstructure, and then extract the most influential measure for further analysis.

There are some limitations of D2 computation as presented in MVComp. First, the Mahalanobis Distance itself is a squared measure, thus the directionality of the difference (feature from the average) is non-specific. As it is currently implemented, it is not possible to determine whether a given subject's features are higher or lower from the average. Future studies could address this limitation by using modeling techniques to define the average based on normative modeling techniques, and splitting groups based on expected direction of change. Then, the directions of deviations from the average could be hypothesized a priori and an additional statistical term could be included in formal statistical analyses.

References

Owen, T. W., de Tisi, J., Vos, S. B., Winston, G. P., Duncan, J. S., Wang, Y., & Taylor, P. N. (2021). Multivariate white matter alterations are associated with epilepsy duration. *The European Journal of Neuroscience*, 53(8), 2788–2803.

<https://doi.org/10.1111/ejn.15055>

Taylor, P. N., Silva, N. M. da, Blamire, A., Wang, Y., & Forsyth, R. (2020). Early deviation from normal structural connectivity: A novel intrinsic severity score for mild TBI. *Neurology*, 94(10), e1021–e1026. <https://doi.org/10.1212/WNL.00000000000008902>

Raffelt, D.; Dhollander, T.; Tournier, J.-D.; Tabbara, R.; Smith, R. E.; Pierre, E. & Connelly, A. Bias Field Correction and Intensity Normalisation for Quantitative Analysis of Apparent Fibre Density. In Proc. ISMRM, 2017, 26, 3541

Dhollander, T.; Tabbara, R.; Rosnarho-Tornstrand, J.; Tournier, J.-D.; Raffelt, D. & Connelly, A. Multi-tissue log-domain intensity and inhomogeneity normalisation for quantitative apparent fibre density. In Proc. ISMRM, 2021, 29, 2472

B. Jeurissen, J.-D. Tournier, T. Dhollander, A. Connelly, and J. Sijbers. Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion MRI data. *NeuroImage*, 103 (2014), pp. 411–426.

Raffelt, D.; Tournier, J.-D.; Fripp, J; Crozier, S.; Connelly, A. & Salvado, O. Symmetric diffeomorphic registration of fibre orientation distributions. *NeuroImage*, 2011, 56(3), 1171-1180

Smith, R. E.; Tournier, J.-D.; Calamante, F. & Connelly, A. SIFT: Spherical-deconvolution informed filtering of tractograms. *NeuroImage*, 2013, 67, 298-312
(Appendix 2)

Raffelt, D.; Tournier, J.-D.; Rose, S.; Ridgway, G.R.; Henderson, R.; Crozier, S.;
Salvado, O.; Connelly, A. Apparent Fibre Density: a novel measure for the analysis of
diffusion-weighted magnetic resonance images. *Neuroimage*, 2012, 15;59(4), 3976-94