
Original Research Articles

MVComp toolbox: MultiVariate Comparisons of brain MRI features accounting for common information across measures

Stefanie A Tremblay, MSc^{1,2,3}^a, Zaki Alasmar, MSc^{2,4b}, Amir Pirhadi, MSc^{5,6}, Felix Carbonell, PhD⁷, Yasser Iturria-Medina, PhD^{8,9,10}, Claudine J Gauthier, PhD^{1,2,3}, Christopher J Steele, PhD^{2,4,11}

¹ EPIC Centre, Montreal Heart Institute, ² School of Health, Concordia University, ³ Physics, Concordia University, ⁴ Psychology, Concordia University, ⁵ ViTAA Medical Solutions, ⁶ Electrical Engineering, Concordia University, ⁷ Biospective Inc., ⁸ Ludmer Center for Neuroinformatics and Mental Health, ⁹ McConnell Brain Imaging Center, Montreal Neurological Institute and Hospital, ¹⁰ Neurology and Neurosurgery Department, McGill University, ¹¹ Neurology, Max Planck Institute for Human Cognitive and Brain Sciences

Keywords: Multivariate analysis, white matter, covariance, personalized assessment, toolbox, python, open-source

<https://doi.org/10.52294/001c.118427>

Aperture Neuro

Vol. 4, 2024

Introduction

Multivariate approaches have recently gained in popularity to address the physiological unspecificity of neuroimaging measures and to better characterize the complexity of biological processes underlying behavior. However, commonly used approaches are biased by the intrinsic associations between variables, or they are computationally expensive and may be more complicated to implement than standard univariate approaches. Here, we propose using the Mahalanobis distance (D2), an individual-level measure of deviation relative to a reference distribution that accounts for covariance between measures. To facilitate its use, we introduce an open-source python-based tool for computing D2 relative to a reference group or within a single individual: the MultiVariate Comparison (MVComp) toolbox (<https://github.com/neuralabc/mvcomp>). The toolbox allows different levels of analysis (i.e., group- or subject-level), resolutions (e.g., voxel-wise, ROI-wise) and dimensions considered (e.g., combining MRI measures or WM tracts). Several example cases are presented to showcase the wide range of possible applications of MVComp and to demonstrate the functionality of the toolbox.

Methods & Results

The D2 framework was applied to the assessment of white matter (WM) microstructure at 1) the group-level, where D2 can be computed between a subject and a reference group to yield an individualized measure of deviation. We observed that clustering applied to D2 in the corpus callosum yields parcellations that highly resemble known topography based on neuroanatomy, suggesting that D2 provides an integrative index that meaningfully reflects the underlying microstructure. 2) At the subject level, D2 was computed between voxels to obtain a measure of (dis)similarity. The loadings of each MRI measure (i.e., its relative contribution to D2) were then extracted in voxels of interest to showcase a useful option of the MVComp toolbox. These relative contributions can provide important insights into the physiological underpinnings of differences observed.

Discussion

Integrative multivariate models are crucial to expand our understanding of the complex brain-behavior relationships and the multiple factors underlying disease development and progression. Our toolbox facilitates the implementation of a useful multivariate method, making it more widely accessible.

^a Corresponding author: Stefanie A. Tremblay (stefanie.tremblay@mail.concordia.ca)

^b Stefanie A. Tremblay and Zaki Alasmar contributed equally to this work

1. INTRODUCTION

In the past decade, there has been exponential growth in the number of modeling approaches that link white matter (WM) microstructural properties and the MR signal.¹ Since none of the existing models (e.g., 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, potentially leading to errors in biological interpretation.^{1,2} 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 measures.³⁻⁵

Multivariate frameworks have the potential to counteract issues arising from the physiologically unspecific nature of commonly used neuroimaging measures and to capture the complexity and heterogeneity of biological properties.^{3, 6-9} Multiple mechanisms give rise to brain structure (e.g., myelination, cell proliferation), support neuroplastic change^{10,11} and behavioral performance,^{8,12} and are involved in neurological disorders.¹³ 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.¹⁴⁻¹⁶ 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.¹⁷⁻¹⁹

Multivariate approaches that combine structural MRI measures 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 between multiple measures.^{20,21} Other multivariate approaches that can be used in group analyses include principal component analysis (PCA), Sparse Group Lasso, independent component analysis (ICA) and non-negative matrix factorization.^{20,22-26} At the individual level, inter-regional correlations of multiple measures can be used to create individual-specific network maps based on morphometric similarity that can then be linked to behavior.⁸ Individualized network maps provide a more comprehensive structural mapping that captures both biological complexity and individual variability because they integrate multiple MRI features.^{27,28} However, in this study by Seidlitz et al., (2018), the shared covariance between MRI measures was not accounted for.⁸ This has the potential to bias inferences made from such analyses, as there is significant covariance between many commonly used imaging parameters.^{4,29} Various multivariate approaches that can overcome this issue exist, including multivariate linear regression,^{15, 30} machine-learning,^{7,20,22,24,31,32} and Hotelling's T^2 test.^{14,33} However, many of these approaches (including multivariate linear regression and machine learning) are computationally expensive and some necessitate making

subjective decisions.^{15,34-36} The Hotelling's T^2 test, a multivariate extension of a two-sample t-test, is a simple yet powerful option for group comparisons,^{14,33} but provides little insight at the individual level.⁷

Here we propose using the Mahalanobis distance (D_2)³⁷ for analyzing multimodal MRI measures. D_2 is closely related to Hotelling's T^2 but can also provide an individual-level measure of deviation relative to a reference distribution. It is defined as the multivariate distance between a point and a distribution in which covariance between features (i.e., imaging measures) is accounted for. Initially developed by P. C. Mahalanobis in 1936 to quantify racial similarities based on anthropometric measurements of skulls, D_2 can be thought of as a multivariate z-score where the covariance between features is accounted for.⁹

The D_2 approach has been used extensively in outlier detection, cluster analysis, and classification applications.³⁸⁻⁴⁰ D_2 has also been used in neuroimaging, mainly in the study of neurological disorders, to detect lesions,^{35, 41} or to evaluate the degree of abnormality in the brains of patients relative to controls,^{6,9,16,42} but also to study healthy WM development.⁴³ Despite promising implementations and its high versatility, D_2 has not yet been widely adopted. To facilitate its use, we present here an open-source python-based tool for computing D_2 relative to a reference group or within a single individual: the MultiVariate Comparison (MVComp) toolbox. We provide a step-by-step guide to computing D_2 using the MVComp tool (<https://github.com/neuralabc/mvcomp>) for two distinctive scenarios: a) comparisons between a subject and a reference group, and b) within-subject comparisons between voxels (Section 2). Lastly, the results of these example cases are presented (Section 3) and the general approach is discussed (Section 4).⁴⁴

2. METHODS

2.1. GENERAL FRAMEWORK

Since D_2 can be defined relative to virtually any reference of matching features, MVComp has been designed to support a wide range of applications. The first step is to define the set of multivariate data that will serve as the reference for computing D_2 . This choice depends on the hypothesis of interest, which will determine the *Level of Analysis* (Fig. 1). D_2 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 spatially correspondent regions. In each case, multiple different *Resolutions* of analysis are possible, including voxel-wise and region of interest- (ROI) based comparisons.

Lastly, the choice of which dimensions to combine, either MRI-derived measures 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,^{9,16} while combining features is useful in the presence of mechanistic heterogeneity (i.e., several concomitant underlying biological mechanisms) and when preserving regional specificity is

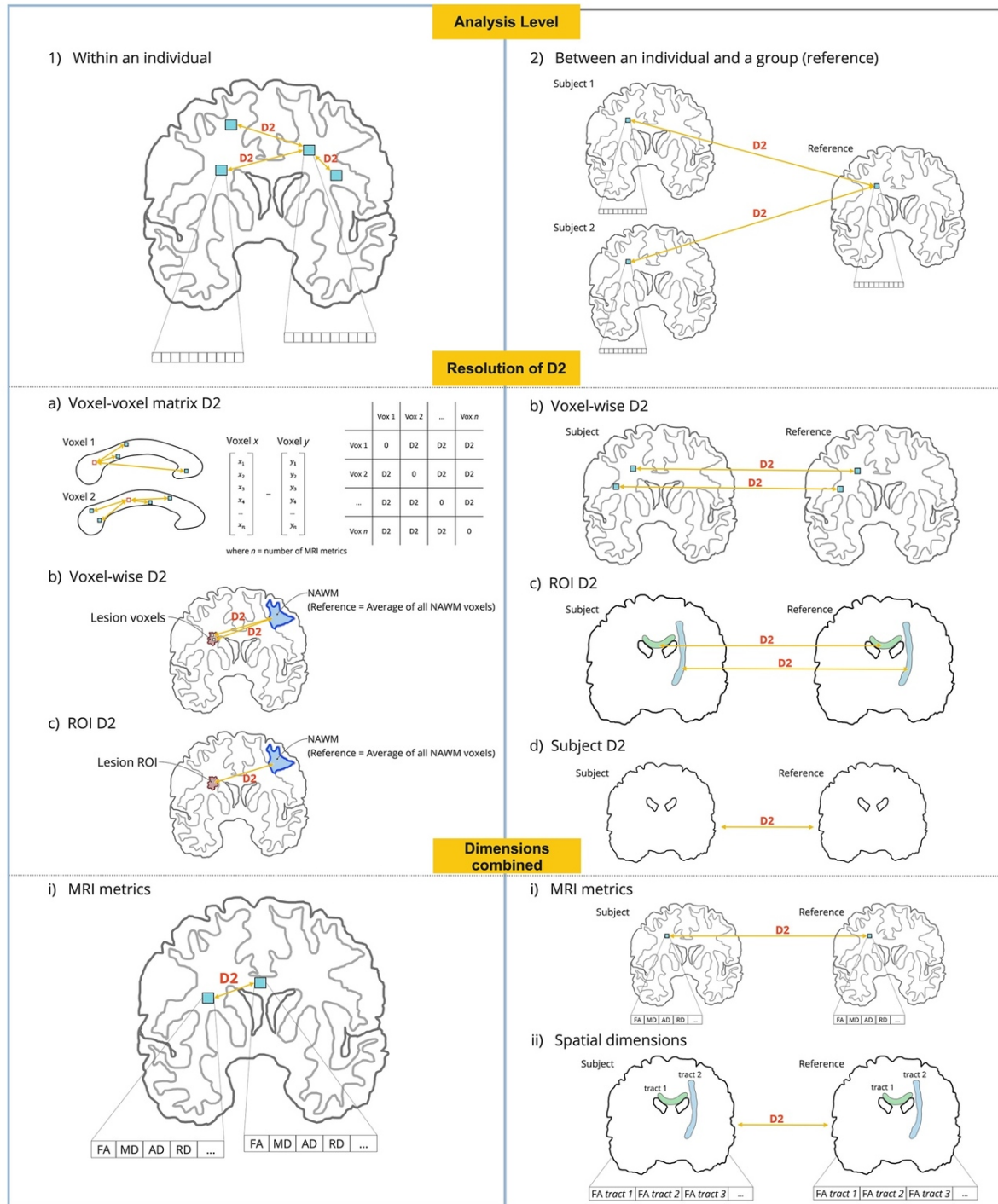


Figure 1. Implementations of the D2 framework in neuroimaging studies.

Analysis level: (1) Within an individual (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 an individual and a group (right panel, in 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: D2 can be computed between each voxel and all other voxels in a mask of analysis, resulting in a D2 matrix of size n voxels $\times$ n voxels (only applicable to analyses within an individual). (b) Voxel-wise D2: A D2 value can be computed at each 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 analyses between an individual and a group). **Dimensions combined:** (i) MRI measures: when the dimensions combined through D2 are MRI measures, the length of the vector of data is the number of measures. (ii) Spatial dimensions: when 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 analyses between an individual and a group; yields a single D2 value per subject).

desirable.^{35,41,42} Both brain regions and MRI measures can be combined, as has been done by Dean and colleagues (2017).⁶ See [Fig. 1](#) for a comprehensive view of the possible combinations of levels of analysis, resolutions and with different dimensions combined.

To illustrate the flexibility of the D2 approach, we present 4 examples:

2.1.1. COMPARISONS BETWEEN AN INDIVIDUAL AND A GROUP (REFERENCE)

Example 1: Computing a voxel-wise D2 map for each individual

Data: Diffusion MRI (dMRI) data in several subjects

Level of Analysis: Between an individual and a group ([Fig. 1](#) right panel)

Feature Resolution: Voxel-wise D2 (in all WM voxels) ([Fig. 1b](#))

Dimensions combined: dMRI-derived maps ([Fig. 1i](#))

In this example the reference would be defined as the voxel-wise group average for each dMRI-derived measure (m_1, m_2, m_n , where n is the number of measures) and D2 is computed by comparing the feature values in each voxel of an individual to the corresponding voxel in the reference (see [Fig. 2a-c](#)). The resulting 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 as the reference and D2 values computed between each patient and the reference would represent voxel-wise multivariate distance from controls.

Example 2 : Computing a single D2 score per individual

Data: dMRI data in several subjects

Level of Analysis: Between an individual and a group ([Fig. 1](#) right panel)

Feature Resolution: Subject D2 ([Fig. 1d](#))

Dimensions combined: WM tracts (spatial dimensions) ([Fig. 1ii](#))

A single MRI measure can also be used and combined across multiple ROIs (e.g., mean FA in pre-defined WM tracts). The reference is defined as the group mean of each tract (m_1, m_2, m_n , where n is the number of tracts) and a single D2 value is computed for each individual. In this case, D2 represents a measure of how different an individual's WM microstructure is relative to a reference, across multiple tracts. This workflow can also be used if the user wishes to combine both brain regions and MRI measures. The reference is then defined as the group mean of each MRI measure for each tract ($m_{11}, m_{12}, m_{1n}, m_{i1}, m_{i2}, m_{in}$, where n is the number of MRI measures, i is the number of tracts and the length of the vector is $n \times i$). Again, a single D2 score is obtained for each subject. These applications are not demonstrated in the present article but have been shown by others^{9,16} and can be implemented using MVComp.

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 (reference) and covariance matrix prior to calculating D2.

2.1.2. COMPARISONS WITHIN AN INDIVIDUAL

Example 3: Computing D2 between lesion voxels and normal appearing WM (NAWM)

Data: dMRI data in one subject

Level of Analysis: Within an individual ([Fig. 1](#) left panel)

Feature Resolution: Voxel-wise (in lesion voxels) ([Fig. 1b](#))

Feature Dimensions: dMRI-derived maps ([Fig. 1i](#))

Here, the level of analysis is within-subject, the dimensions combined are multiple dMRI-derived measures in each voxel, and the reference is the average of all voxels within a region of interest (ROI) for each dMRI measure. To investigate the distance between WM lesions and NAWM, the reference would be defined as the average of all NAWM voxels (m_1, m_2, m_n , where n is the number of measures) and D2 would be computed for each voxel classified as a lesion. Alternatively, the *resolution* could be ROI-wise if the user deems a single D2 value per lesion sufficient. This within-subject approach can also be used as a measure of similarity by computing D2 between all WM voxels and a reference ROI in a specific tract (e.g., voxels in the cortico-spinal tract, as in [Fig. 2d](#)). Voxels within the same WM tract as the reference ROI are likely to have lower D2 values (indicating higher similarity) than voxels in other tracts or in areas of crossing fibers ([Fig. 2e](#)).

Example 4: Computing D2 between each voxel and all other voxels in a mask

Data: dMRI data in one subject

Level of Analysis: Within an individual ([Fig. 1](#) left panel)

Feature Resolution: Voxel-voxel D2 matrix ([Fig. 1a](#))

Feature Dimensions: dMRI-derived maps ([Fig. 1i](#))

D2 can be calculated between every pair of voxels (voxel x – voxel y) within a mask of analysis to compute a voxel-voxel D2 matrix (see [Fig. 1a](#)). In this case, the reference for computing the covariance matrix would be the data in all voxels contained in the mask.

See Supplementary material for a summary of the workflow for the 4 examples.

2.2. DATA PREPARATION

In all cases, data for all subjects should be preprocessed and all MRI measures of interest computed and transformed to bring them into the same voxel space. If instead of voxel-wise comparisons the user is interested in performing ROI-based comparisons, summary measures should be calculated for each region of interest (e.g., mean FA in each WM

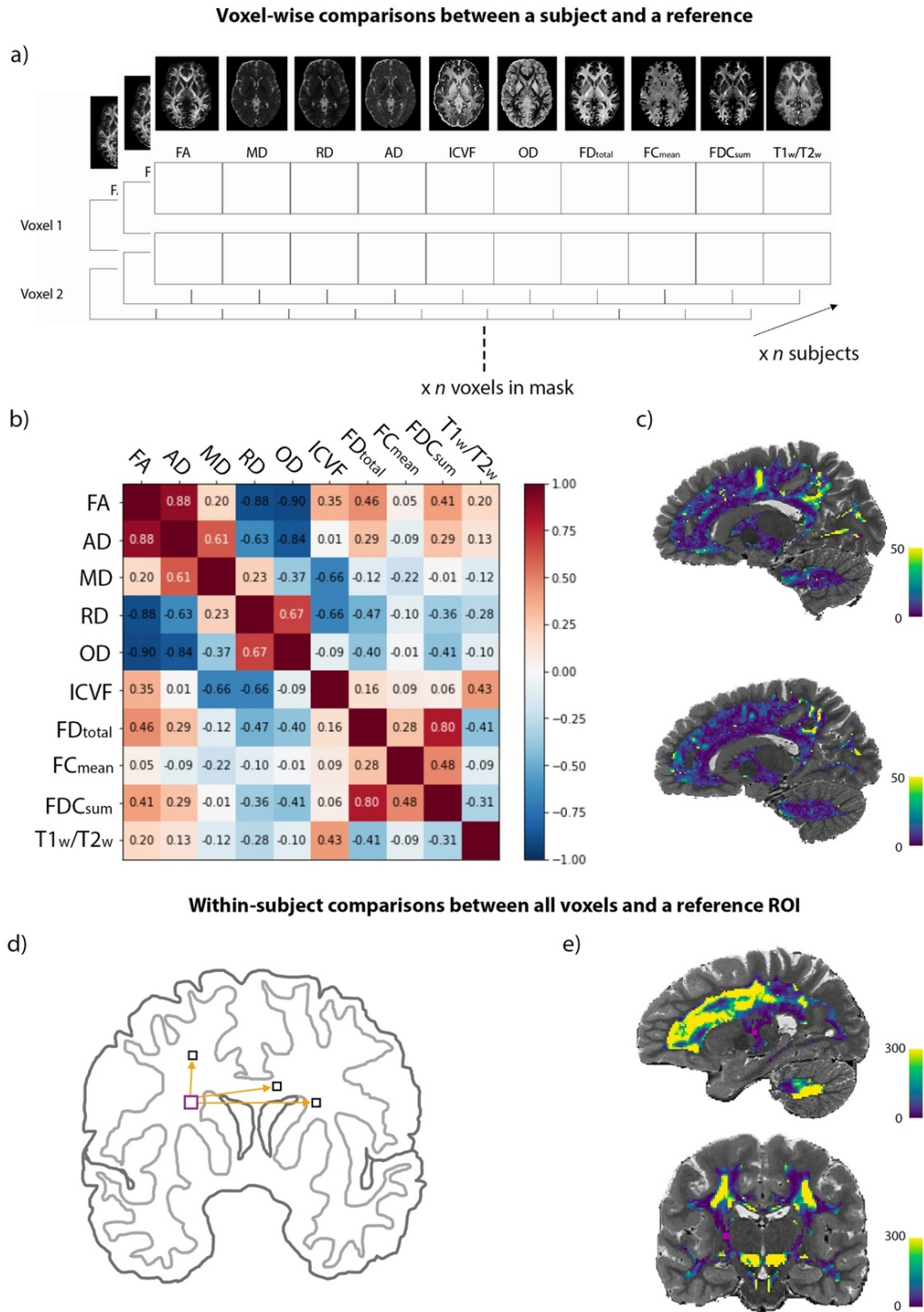


Figure 2. D2 workflow. Voxel-wise comparisons between a subject and a reference.

(a) The multivariate space is illustrated here. In this example, we have a vector of 10 dMRI measures at each WM voxel for each subject.* (b) The covariance matrix is computed from the reference feature matrix of shape n voxels in WM $\times$ n features. The plot shows the amount of correlation between features in the reference sample (i.e., the whole group). (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. Note that the leave-one-subject-out approach was used so that the data of the subject under evaluation was not included in the group mean (i.e., reference) and covariance matrix prior to D2 calculation. **Within-subject comparisons between all WM 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 WM voxels and the CST ROI (in pink).

*Data used for these examples will be presented in section 2.7.

tract of interest) for each subject. Masks should also be generated to restrict analyses to chosen regions (e.g., WM) and these should also be transformed into the same space. Masks can be binary or thresholded at a later step within MVComp.

2.3. COMPUTING THE REFERENCE MEAN AND COVARIANCE MATRIX

In the case of analyses between subject(s) and a reference (Fig. 1 right panel), the reference mean and covariance matrix are derived either from multiple features (Fig. 1i) or multiple ROIs (Fig. 1ii) 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 (Fig. 1 left panel), multiple features can be compared between voxels (e.g., Fig. 1 a-b) or between ROIs (e.g., Fig. 1c).

2.3.1. COMPARISONS BETWEEN AN INDIVIDUAL AND A GROUP (REFERENCE)

COMBINING MRI MEASURES

For this application, the group average of each measure must be computed from the reference group (`mvcomp.compute_average` can be used to perform this task). The `mvcomp.feature_list` function can then be used to create a list of feature names and a list of the full paths of the average maps that were created with the `compute_average` function. The `feature_gen` function extracts the feature matrix from a set of input images. Run on the reference group mean images with a provided mask, it returns the feature matrix (`m_f_mat` of shape n voxels in the mask $\times$ n features), a mask vector (`mat_mask` of shape n voxels) and a nibabel object of the mask (`mask_img`). 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 `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 `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.

A leave-one-out approach (where the individual to be compared to the reference 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`). If the leave-one-out approach is used, it is not necessary to compute the group average nor to use the `feature_gen` and `norm_covar_inv` functions since the average and covariance matrix will be computed within the `model_comp` function from a group that excludes the subject for which D2 is being computed.

COMBINING SPATIAL DIMENSIONS

The reference mean values (e.g., reference group mean FA in each WM tract) and covariance matrix are computed within the `spatial_mvcomp` function described in detail below. See^{9,16} for example applications of this implementation.

2.3.2. COMPARISONS WITHIN AN INDIVIDUAL

VOXEL-WISE D2 RESOLUTION

In the case of comparisons within a single subject, one of the possible applications is to compute D2 between specific ROIs. If the reference ROI is a set of NAWM voxels, 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 measure) 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 `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 `correlation_fig` function can again be used to visualize relationships between measures.

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 $\times$ n features.

2.4. COMPUTING D2

Once the mean of the reference and the covariance matrix have been computed and the data for comparisons has been prepared, the D2 computation can be performed. Depending on the *resolution* of D2, this computation may be repeated several times (i.e., between every pair of voxels or once for each voxel or each ROI; Fig. 1a-c), or it may only be done once if the user is interested in obtaining a single individualized score of deviation from a group (Fig. 1d). The MVComp tool contains functions to easily compute D2 for each of these applications, according to this equation:

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

where x is the vector of data for one observation (e.g., one subject), m is the vector of averages of all observations for each independent variable (e.g., MRI measures), and C^{-1} is the inverse of the covariance matrix.

2.4.1. COMPARISONS BETWEEN AN INDIVIDUAL AND A GROUP (REFERENCE)

COMBINING MRI MEASURES

The `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) (Fig. 1 right panel; b). The user should specify the directories and

suffix of the subjects' features and of the 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 user wishes to use the leave-one-out approach, the `exclude_comp_from_mean_cov` option should be set to True. If this option is set to True, the mean (reference) and `pinv_s` are computed for each subject comparison, excluding the subject being compared before computing its D2. Therefore, it is not necessary to specify the directory of the reference (`model_dir`) in this application. The `model_comp` function yields a matrix containing the D2 data of all subjects (of size number of voxels x number of subjects). To obtain a D2 map (in nifti format) for each subject, the `dist_plot` function can then be used. The function also outputs a mean D2 map of all subjects and a histogram of all D2 values.

COMBINING SPATIAL DIMENSIONS

The `spatial_mvcomp` function is used to compute a D2 score between each subject and the reference computed from all subjects (Fig. 1ii). A matrix containing the data (e.g., mean FA in each WM tract) of all subjects (n subjects x n tracts) should be provided to the function. The `spatial_mvcomp` function returns a vector with a single D2 value per subject, reflecting the subject's individualized score of deviation from the group. 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. COMPARISONS WITHIN AN INDIVIDUAL

VOXEL-WISE D2 RESOLUTION

The `mah_dist_mat_2_roi` function is used to compute voxel-wise D2 between all voxels within a mask and a specific ROI (Fig. 1 left panel; b). Here, in addition to the feature matrix containing the data for the voxels to be evaluated (n voxels in the mask x 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 measure) and the inverse of the covariance matrix (`pinv_s`) calculated previously.

VOXEL-VOXEL MATRIX D2 RESOLUTION

The `voxel2voxel_dist` function is used to compute D2 between each voxel and all other voxels in a mask (Fig. 1 left panel; a). This yields a symmetric 2-D matrix of size n voxels x n voxels containing D2 values between each pair of voxels.

2.5. STATISTICAL ANALYSIS

Once D2 values are computed, second-level statistical analyses can be used to assess group differences and longitudinal trajectories, to explore relationships between D2 and behavior. Machine learning techniques can also be used to reduce dimensionality and produce network maps based on (dis)similarity.

2.5.1. COMPARISONS BETWEEN AN INDIVIDUAL AND A GROUP (REFERENCE)

For group comparisons, a two-samples t-test can be performed on D2 values (e.g., D2 values in patients vs D2 in controls), which would be equivalent to performing a Hotelling's T^2 test on raw measures (i.e., without computing D2). 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.⁶ However, such group analyses are likely to average out interindividual variability and may be problematic when heterogeneity is high.⁷ Wilk's criterion is another approach that can be used to define abnormality based on a calculated critical value that accounts for normative sample size, number of features, and multiple comparisons.^{35,42,45}

2.5.2. COMPARISONS WITHIN AN INDIVIDUAL

In within-subject analyses, clustering approaches can be applied to the voxel-voxel D2 matrix to partition brain voxels into networks or other parcellations.

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.^{41,43} 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.^{6,9,16}

2.6. DETERMINING FEATURE IMPORTANCE

D2 summarizes the amount of deviation from a reference, based on several measures or brain regions, into a single score. This yields a useful metric to easily quantify *abnormalities*, whether due to pathology or to exceptional abilities such as musical skills. However, when summarizing several features into a single score, we lose specificity. To help address this limitation, it is possible to extract the contribution of each feature to the multivariate distances (D2) using functions of the MVComp tool to recover biological or spatial specificity.

2.6.1. COMPARISONS BETWEEN AN INDIVIDUAL AND A GROUP (REFERENCE)

COMBINING MRI MEASURES

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 `model_comp` function can be used. This allows the extraction of each measure's weight in D2. If `return_raw` is set to True, the function returns a 3D array of size (number of voxels) x (number of measures) x (number of subjects) that contains the voxel-wise distances for each feature and each subject. A flattened mask of the region of interest (e.g., a region of high D2) can then be applied to

select voxels from the 3D array. The distances can be summarized across voxels and/or subjects to obtain a % contribution to D2 for each MRI measure 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. COMPARISONS WITHIN AN INDIVIDUAL

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 measures).

2.7. EXPERIMENTS

2.7.1. DATA DESCRIPTION

We computed 10 microstructural features for 1001 subjects from the Human Connectome Project S1200 data release⁴⁶ 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.^{46,47} 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), us-

ing 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.⁴⁸ 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 MRtrix3.⁴⁹ The tensor was computed on the bias field-corrected DWI data (using `dwi2tensor`) and DTI measures were then calculated (FA, MD, AD and RD) using `tensor2metric`.⁵⁰⁻⁵²

MULTI-TISSUE MULTI-SHELL CONSTRAINED SPHERICAL DECONVOLUTION

The multi-tissue Constrained Spherical Deconvolution (CSD) was performed following the [fixel-based analysis \(FBA\) workflow](#).⁴⁸ The T1-w images were segmented using the `5ttgen` FSL function of MRtrix3, which uses the FAST algorithm.⁵³⁻⁵⁷ 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).⁵⁸ The bias-uncorrected DWI data was used because bias field correction is performed at a later step in the FBA pipeline.⁵⁹ 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.⁵⁸ 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 normalization, which normalizes signal amplitudes to make subjects comparable, were then performed on the ODFs, using the `mt_nor-malise` function in MRtrix3.^{60,61}

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 and the “nodif” brain masks of a subset of 200 participants using the `population_template` function of MRtrix3 (with regularization parameters: `nl_update_smooth=1.0` and `nl_disp_smooth=0.75`), resulting in a group template for each of the three tissue types.⁴⁸

Subject-to-template warps were computed using `mrregister` in MRtrix3 with the same regularization 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`.⁶² T1w and T2w images were kept in native resolution (0.7mm) and the ratio of T1w/T2w was calculated to produce a myelin map.⁶³ 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 are then used to generate a WM mask for analyses.

COMPUTING FIXEL MEASURES

The WM FOD template was segmented to generate a fixel mask using the `fod2fixel` function.^{64,65} This mask determines the fiber bundle elements (i.e., fixels), within each voxel of the template mask, that will be considered for subsequent analyses. Fixel segmentation was then performed from the WM FODs of each subject using the `fod2fixel` function, which also yields the apparent fibre density (FD) metric. The `fixelreorient` and `fixelcorrespondence` functions were then used to ensure subjects' fixels map onto the fixel mask.⁴⁸

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 MEASURES INTO VOXEL SPACE

In order to integrate all measures 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. 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).⁶⁶ 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.

Lastly, the voxel-wise sum of FDC, reflecting the total information-carrying capacity at each voxel, was computed using the `fixel2voxel` sum option.

NODDI MEASURES

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).^{67,68} 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. 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 bvects, 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 measures 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 AN INDIVIDUAL AND A GROUP (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 dMRI measures. 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 from 10 microstructural features was computed in the CC for each subject, yielding a D2 matrix of 1001 subjects X 2845 voxels. The D2 values represent voxel-wise microstructural differences 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).

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 profiles.⁶⁹⁻⁷¹ We therefore hypothesized that these segments could be extracted via clustering, an unsupervised machine learning technique, of D2 values in the CC. We performed k-means clustering on the D2 matrix, setting the number of clusters to 9 based on literature on CC topography.⁶⁹⁻⁷¹ Prior to clustering, we applied z-score and power transformation on the D2 matrix to achieve gaussian distributions of the standardized scores. 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. Final visualization was done using BrainNet Viewer (<http://www.nitrc.org/projects/bnv/>).

2.7.4. EXPERIMENT 2: COMPARISONS WITHIN AN INDIVIDUAL

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 were used to compute the covariance matrix and this same covariance matrix was 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 (Fig. 4a-b). We standardized the matrix to z-scores and applied Principal component analysis (PCA) to reduce the matrix dimensionality (Fig. 4c). Of note, eigendecomposition on the D2 matrix would be a valid alternative approach to reduce matrix dimensionality. In fact, because data was centered prior to applying PCA, the first eigenvector should be equivalent to the first principal component from PCA. We then extracted the contributions of each metric to D2 within the voxels with the largest and the lowest scores on the first principal component (Fig. 4d-f).

3. RESULTS

3.1. EXPERIMENT 1: COMPARISONS BETWEEN AN INDIVIDUAL AND A GROUP (REFERENCE)

For this experiment, D2 was computed voxel-wise in the CC between each subject and a reference consisting in all other

subjects (Fig. 3a-b). K-means clustering was applied to the D2 matrix of size (subjects) X (voxels). We observed that the 9 clusters were distributed along the anterior-posterior axis, in accordance with past evidence on CC microstructure and connectivity.⁶⁹⁻⁷¹ Fig. 3c shows the clusters identified via k-means and Fig. 3d shows the topography expected according to literature. The genu of the CC was clustered into 3 segments, while the midbody displays 2 segments. The splenium was divided into 4 segments (with one segment positioned on the isthmus). Segmentations from individual MRI measures are also presented in Supplementary material.

3.2. EXPERIMENT 2: COMPARISONS WITHIN AN INDIVIDUAL

For the within-subject experiment, D2 was computed between all voxel pairs in the CC of a single individual, yielding a voxel X voxel D2 matrix (Fig. 4a-b). PCA was applied to the D2 matrix. Fig. 4c shows the first 10 principal components (PCs). We then extracted the contributions (i.e., loadings) of each metric to D2 within the voxels with the largest and the lowest scores on the first principal component. The first PC explained 95% of the variance in the voxel X voxel dense D2 matrix. The highest and the lowest PC1 scores were in the genu and in the midbody of the CC, respectively (Fig. 4d). 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 (Fig. 4f). On the other hand, in the voxel with the lowest score on PC1, all microstructural features had nearly

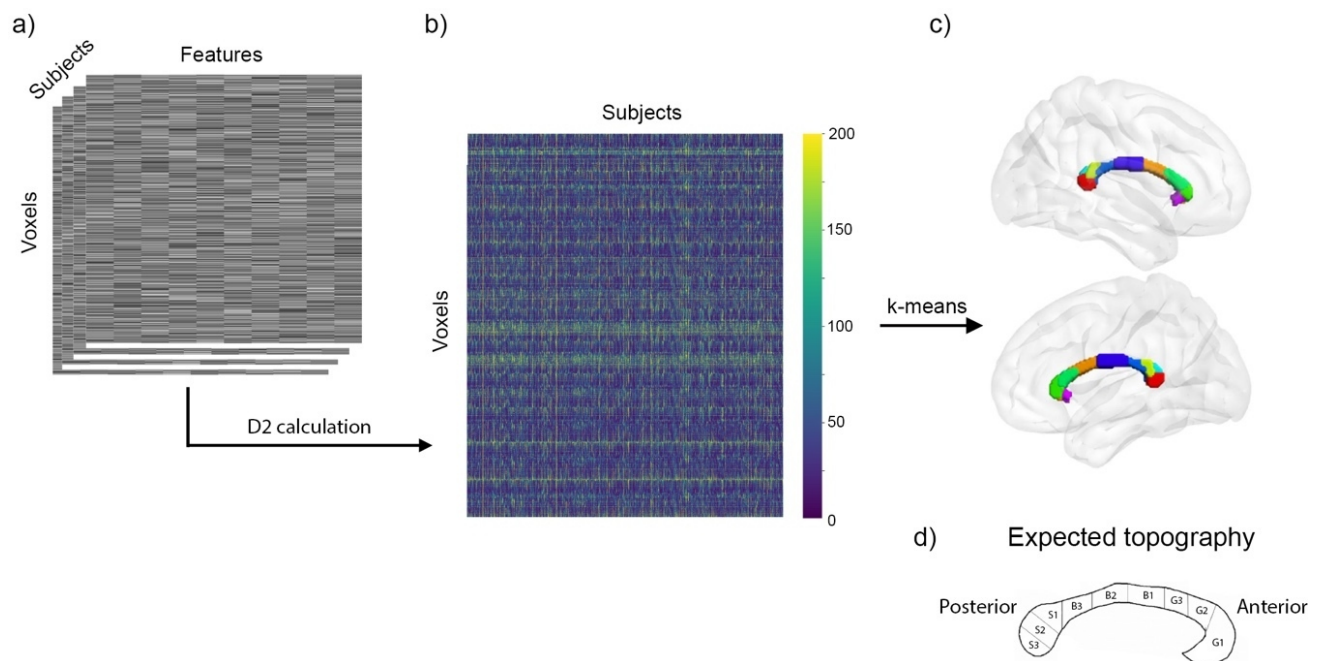


Figure 3. Voxel-wise comparisons between each subject and the reference.

(a) Voxel-wise D2 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 corpus callosum (CC). (b) This results in a D2 matrix of size subject (723 after exclusion of outliers) X voxel (2845) containing the multivariate distance between a subject's data and the reference at each CC voxel. (c) Applying k-means clustering to the D2 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.⁶⁹⁻⁷¹

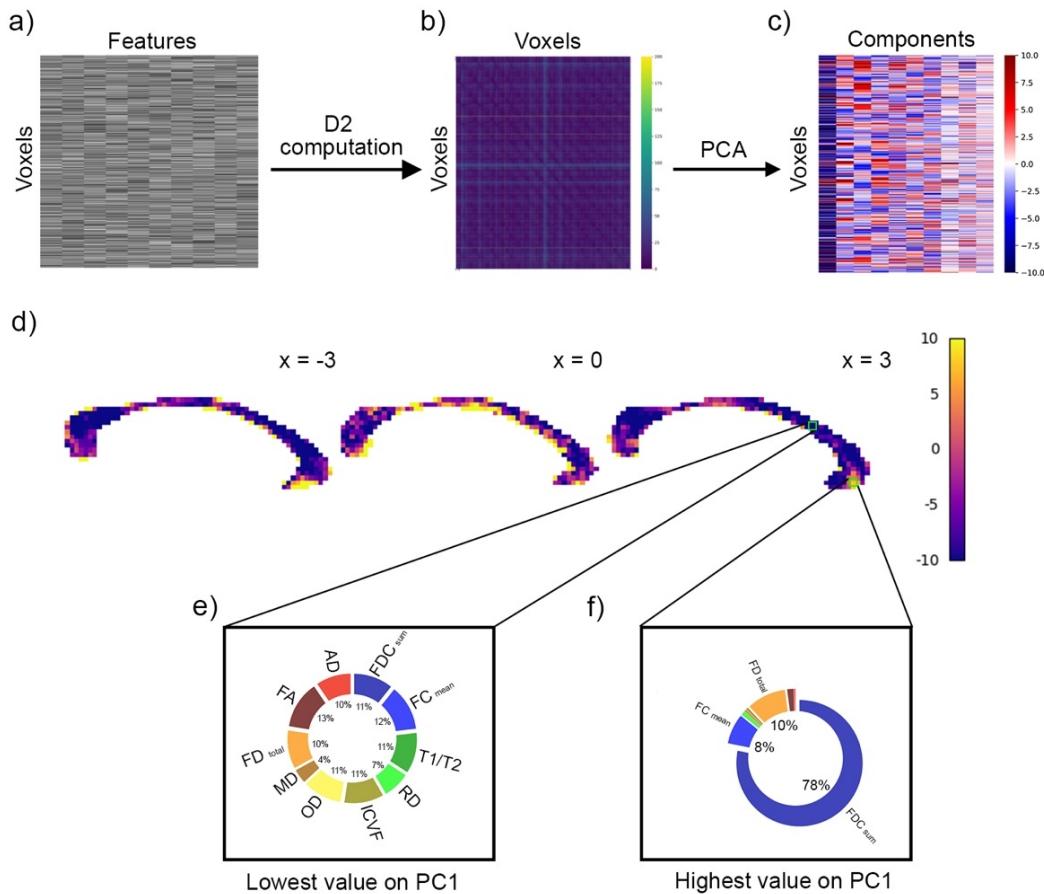


Figure 4. Within-subject voxel-voxel comparisons.

D2 was computed between all voxel pairs from the (a) (features) $\times$ (voxels in the CC) matrix of a subject. (b) A voxel $\times$ voxel D2 matrix was generated. (c) PCA was then applied to the D2 matrix. The PCA matrix shows the first 10 principal components. (d) Voxels with the highest and lowest score on PC1 are shown. PC1 scores were scaled between -10 and 10 to facilitate visualization. (e) In the voxel with the lowest value on PC1, located in the midbody of the CC, all measures had approximately equal contribution to D2. (f) SumFDC contributed most to D2 in the voxel with the highest PC1 score, located in the genu of the CC.

equal contributions to D2, indicating minimal variability in this voxel (Fig. 4e).

4. DISCUSSION

In the present study, we introduced the MVComp tool,⁴⁴ a set of python-based functions that can be used to compute the Mahalanobis distance (D2) for a wide range of neuroimaging applications. At the group-level, MVComp allows the calculation of a score that quantifies how different the brain structure of an individual is from a reference group. The MVComp tool provides a versatile framework that can be used to answer various research questions, from quantifying the degree of abnormality relative to a control group in individuals with a pathology, to exploring interindividual variability in healthy cohorts. At the subject level, D2 can be used to assess differences between regions of interest or to compute a measure of similarity that can then be used for subsequent analyses (e.g., graph theory/network analyses). Lastly, D2 can combine multiple MRI measures in the same spatial locations, or it can combine a single metric across several brain regions.

Our approach allows the integration of several variables while accounting for the relationships between these variables. Several biological properties influence the same neuroimaging metric and multiple neuroimaging measures indirectly reflect a similar underlying physiological property. This overlap means that accounting for covariance between measures is essential. It also means that using a single neuroimaging measure, or measures stemming from a single model, offers limited potential for interpretation and is biased by the set of assumptions of the chosen model (e.g., some models assume fixed compartment diffusivities while others attempt to estimate them).¹ Similarly, integrating the assessment of multiple brain regions may map better onto behavior (e.g., cognition or disease severity) than assessing each region separately. Here, again the relationships between variables should be accounted for as observations are not completely independent from each other (i.e., in the same individual, there is likely a great amount of covariance between FA in different voxels or in different WM tracts). 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,^{8,18,33,72} and do not extend from one level to another.

Frameworks using PCA can be implemented at both levels of analysis and recent studies have used PCA to combine dMRI measures, forming a reduced set of biologically interpretable variables.^{7,26} While this method is interesting as it allows for dimensionality reduction and visualization of patterns between variables, it differs from D2 in several aspects. For instance, in PCA, the component loadings are usually calculated once on the whole sample and yield a set of new variables that will then be used for all subjects and in all brain regions. On the other hand, the features' loadings in D2 can be different for each subject and at each voxel. In this sense, D2 would provide a more individualized measure of deviation or dissimilarity relative to a reference, whereas PCA can be seen more as a dimensionality reduction method. Another advantage of the D2 framework is that it is highly versatile and relatively simple to implement. Moreover, the open-source MVComp toolbox we propose makes its implementation accessible for flexibly assessing a wide variety of research questions (see Fig. 1).

One of the novelties of this work is that it provides the option to extract the contributions of all features within the D2 measure, similar to extracting loadings in PCA.²⁶ This addresses one of the main limitations of some multivariate frameworks, allowing researchers to develop more mechanistic interpretations. 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 measures) were not extracted, which has been a significant limitation.⁶ Characterizing the extent by which each feature contributes to D2 can provide important insights into the physiological underpinnings of the differences observed and/or their localization. To our knowledge, MVComp is the only available toolbox for computing D2 on imaging data. In this paper, we detailed the usage of MVComp through 4 example cases (see Supplementary material) covering a wide range of applications and presented the results of 2 experiments.

4.1. D2 REFLECTS THE UNDERLYING MICROSTRUCTURE OF WM

To provide specific examples of how MVComp can be used, the D2 framework was applied to the assessment of WM microstructure. We found the approach to be particularly suitable for the study of WM because of the number of modeling methods available for dMRI data. However, it is important to note that other types of tissues and imaging techniques can also be used within the MVComp framework. By applying K-Means clustering to D2 in the corpus callosum, we observed a clear segmentation along the anterior-posterior axis (Fig. 3), consistent with known topography from ex-vivo anatomical studies and tractography-based connectivity.⁶⁹⁻⁷¹ This high correspondence between clustered D2 and previously described CC topography suggests that the microstructural score obtained by combining several WM neuroimaging measures through D2 provides a useful index of microstructure.

At the individual level, D2 can capture the amount of (dis)similarity between voxels and, through the extraction of features' contributions (i.e., loadings), the specific mi-

crostructural properties underlying regional differences can be inferred. For example, in our within-subject experiment (Fig. 4) we found high spatial heterogeneity in the relative contributions of different features to D2. The voxel with the highest loading on the first latent component (PC1) was primarily dominated by one metric (sumFDC) while the voxel with the lowest loading was characterized by similar weightings across all features. In the voxel with the highest PC1 score, sumFDC (combined metric of fiber cross-section and density, indicative of the amount of information-carrying capacity) contributed most to D2, meaning sumFDC had higher variability across CC voxels than other measures. This is consistent with the known microstructural properties of the CC, which shows regional variations in densities of fibers of different sizes along the CC.^{69,71} Further, given that the CC is composed of tightly packed fiber tracts, MD would likely be very low in all those CC voxels (i.e., low variability), which would explain its low contribution. Overall, this supports the relevance of D2 in assessing variability in WM microstructure properties and showcases the use of the features contribution option (i.e., `return_raw`) included in MVComp.

4.2. D2 IN THE STUDY OF PATHOLOGIES

Given the complexity of underlying pathological changes in various brain conditions, multiparametric approaches are a promising avenue to capture the combination of multiple changes in brain properties.^{6,7,9,13,16,42} 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 (e.g., mean FA in a single WM tract).¹⁶ Another study reported better performance using D2 encompassing FA in several WM tracts, vs using FA in a single tract, in discriminating between controls and individuals with TBI.⁹ The multivariate D2 measure allowed for the discrimination of even mild TBI cases from controls and correlated significantly with cognitive scores. Similarly, using D2 combining both spatial (i.e., WM regions) and feature (i.e., different DTI measures) dimensions led to improved detection between autistic and typically developing individuals compared to univariate approaches or to D2 computed by combining brain regions only.⁶ Associations between D2 and autism symptom severity were also reported in this study, providing additional evidence that D2 can serve as a behaviorally relevant measure of WM abnormality.

Other interesting implementations have used D2 to detect and characterize lesions. Gyebrár et al. (2019) combined DTI eigenvalues into a voxel-wise D2 measure between epilepsy patients and controls to detect cortical malformations in patients.³⁵ Voxels were identified as belonging to a lesion if their D2 value exceeded a critical value calculated using Wilks' criterion,⁴⁵ 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 WM hyperintensities and those in normal appearing WM (NAWM).⁴¹ D2 in WM hyperin-

tensities progressed at a quicker rate in individuals who converted from mild cognitive impairment to Alzheimer's disease (AD) compared to those who did not convert. Interestingly, the rate of change of WM hyperintensities volume (i.e., lesion load), a metric more commonly used,^{73,74} 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.⁴¹

4.3. LIMITATIONS

There are some limitations of D2 computation as presented in MVComp. First, D2 itself is a squared measure, thus the directionality of the difference is non-specific. As it is currently implemented, it is not possible to determine whether a given subject's features are higher or lower than the average, although this information can be easily extracted by comparing the subject's voxel values or ROI means to the mean of the group average on a per-metric basis. Future studies could potentially address this limitation indirectly by integrating with studies that model ground-truth biophysical properties to better interpret differences and/or splitting groups based on expected direction of change. Then, the directions of deviations from the average could be hypothesized a priori.

D2 is a sensitive multivariate distance measure that has since found applications in various fields, such as classification, cluster analysis, and outlier detection. Our implementation makes use of the sensitivity of D2 to detect multivariate deviations in WM microstructure. This high sensitivity also means the method can be affected by registration inaccuracies and partial voluming (PV). Therefore, special attention must be paid to ensure optimal alignment across subjects and modalities (e.g., using directional information from dMRI to align WM tracts). Strict tissue type masking (e.g., using a high threshold on probabilistic segmentation images) can also be used to limit the amount of PV. However, this may result in a large number of excluded voxels, especially for low resolution images. Alternatively, the PV effect can be quantified and accounted for.^{35,75} The latter option would be preferable if the D2 framework was used to detect tumors and estimate their volume, for instance.

Another limitation of D2 as presented in MVComp is that its use is restricted to continuous variables. However, more recent formulations of D2 allow for nominal and ordinal variables to be incorporated in the model, in addition to continuous variables.^{76,77} Future developments of MVComp could thus allow generalization of D2 to include mixed data types (e.g. WM, sex, or other grouping variable).

5. CONCLUSION

We introduce a new open-source tool for the computation of the Mahalanobis distance (D2), the MVComp (MultiVariate Comparisons) toolbox.⁴⁴ D2 is a multivariate distance measure relative to a reference that inherently accounts

for covariance between features. MVComp can be used in a wide range of neuroimaging implementations, at both the group and subject levels. In line with the current shift towards precision medicine, MVComp can be used to obtain personalized assessments of brain structure and function, which is essential in the study of brain conditions with high heterogeneity.

DATA AND CODE AVAILABILITY

The data is openly available from the Human Connectome Project (<https://www.humanconnectome.org/study/hcp-young-adult/document/1200-subjects-data-release>) and the code of the MVComp toolbox is available at <https://github.com/neuralabc/mvcomp>.⁴⁴

AUTHOR CONTRIBUTIONS

Stefanie A Tremblay: Writing - Original Draft, Conceptualization, Data Curation, Methodology, Software, Validation, Visualization

Zaki Alasmar: Methodology, Software, Formal analysis, Validation, Conceptualization, Data Curation, Writing - Original Draft, Visualization

Amir Pirhadi: Methodology, Software, Validation, Data Curation

Felix Carbonell: Methodology, Writing - Review & Editing

Yasser Iturria-Medina: Methodology, Writing - Review & Editing, Conceptualization, Funding acquisition

Claudine J Gauthier: Supervision, Conceptualization, Writing - Review & Editing

Christopher J Steele: Supervision, Conceptualization, Methodology, Writing - Review & Editing, Software, Funding acquisition

FUNDING SOURCES

This study was supported by the Canadian Institutes of Health Research (FRN: 175862, to Stefanie A. Tremblay), the Canadian Natural Sciences and Engineering Research Council (RGPIN-2015-04665, to Claudine J. Gauthier), the Michal and Renata Hornstein Chair in Cardiovascular Imaging (to Claudine J. Gauthier). Christopher J. Steele has received funding from the Natural Sciences and Engineering Research Council of Canada (DGECR-2020-00146), the Canada Foundation for Innovation (CFI-JELF Project number 43722) and Heart and Stroke Foundation of Canada (National New Investigator), and the Canadian Institutes of Health Research (HNC 170723). Christopher J. Steele and Yasser Iturria-Medina were also supported by the Quebec Bioimaging Network.

Data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

DECLARATION OF COMPETING INTERESTS

Submitted: March 01, 2024 CDT, Accepted: May 23, 2024 CDT

The authors have no competing interests to declare.



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-4.0). View this license's legal deed at <http://creativecommons.org/licenses/by/4.0> and legal code at <http://creativecommons.org/licenses/by/4.0/legalcode> for more information.

REFERENCES

1. Novikov DS, Kiselev VG, Jespersen SN. On modeling. *Magn Reson Med*. 2018;79(6):3172-3193. [doi:10.1002/mrm.27101](https://doi.org/10.1002/mrm.27101)
2. Alexander DC, Dyrby TB, Nilsson M, Zhang H. Imaging brain microstructure with diffusion MRI: practicality and applications. *NMR Biomed*. 2019;32(4):e3841. [doi:10.1002/nbm.3841](https://doi.org/10.1002/nbm.3841)
3. Tardif CL, Gauthier CJ, Steele CJ, et al. Advanced MRI techniques to improve our understanding of experience-induced neuroplasticity. *Neuroimage*. 2016;131:55-72. [doi:10.1016/j.neuroimage.2015.08.047](https://doi.org/10.1016/j.neuroimage.2015.08.047)
4. Uddin MN, Figley TD, Solar KG, Shatil AS, Figley CR. Comparisons between multi-component myelin water fraction, T1w/T2w ratio, and diffusion tensor imaging measures in healthy human brain structures. *Sci Rep*. 2019;9(1):2500. [doi:10.1038/s41598-019-39199-x](https://doi.org/10.1038/s41598-019-39199-x)
5. Bells S, Cercignani M, Deoni S, et al. Tractometry—comprehensive multi-modal quantitative assessment of white matter along specific tracts. In: *Proc. ISMRM*. Vol 678. ; 2011.
6. Dean DC, Lange N, Travers BG, et al. Multivariate characterization of white matter heterogeneity in autism spectrum disorder. *NeuroImage Clin*. 2017;14:54-66. [doi:10.1016/j.nicl.2017.01.002](https://doi.org/10.1016/j.nicl.2017.01.002)
7. Guberman GI, Stojanovski S, Nishat E, et al. Multi-tract multi-symptom relationships in pediatric concussion. Shackman A, Büchel C, Rokem A, eds. *eLife*. 2022;11:e70450. [doi:10.7554/eLife.70450](https://doi.org/10.7554/eLife.70450)
8. Seidlitz J, Váša F, Shinn M, et al. Morphometric Similarity Networks Detect Microscale Cortical Organization and Predict Inter-Individual Cognitive Variation. *Neuron*. 2018;97(1):231-247.e7. [doi:10.1016/j.neuron.2017.11.039](https://doi.org/10.1016/j.neuron.2017.11.039)
9. Taylor PN, Silva NM da, Blamire A, Wang Y, Forsyth R. Early deviation from normal structural connectivity: A novel intrinsic severity score for mild TBI. *Neurology*. 2020;94(10):e1021-e1026. [doi:10.1212/WNL.0000000000008902](https://doi.org/10.1212/WNL.0000000000008902)
10. Azzarito M, Emmenegger TM, Ziegler G, et al. Coherent, time-shifted patterns of microstructural plasticity during motor-skill learning. *NeuroImage*. 2023;274:120128. [doi:10.1016/j.neuroimage.2023.120128](https://doi.org/10.1016/j.neuroimage.2023.120128)
11. Taubert M, Villringer A, Ragert P. Learning-Related Gray and White Matter Changes in Humans: An Update. *The Neuroscientist*. 2012;18(4):320-325. [doi:10.1177/1073858411419048](https://doi.org/10.1177/1073858411419048)
12. Thiebaut de Schotten M, Forkel SJ. The emergent properties of the connected brain. *Science*. 2022;378(6619):505-510. [doi:10.1126/science.abq2591](https://doi.org/10.1126/science.abq2591)
13. Iturria-Medina Y, Carbonell FM, Sotero RC, Chouinard-Decorte F, Evans AC. Multifactorial causal model of brain (dis)organization and therapeutic intervention: Application to Alzheimer's disease. *NeuroImage*. 2017;152:60-77. [doi:10.1016/j.neuroimage.2017.02.058](https://doi.org/10.1016/j.neuroimage.2017.02.058)
14. Avants B, Duda JT, Kim J, et al. Multivariate Analysis of Structural and Diffusion Imaging in Traumatic Brain Injury. *Acad Radiol*. 2008;15(11):1360-1375. [doi:10.1016/j.acra.2008.07.007](https://doi.org/10.1016/j.acra.2008.07.007)
15. Naylor MG, Cardenas VA, Tosun D, Schuff N, Weiner M, Schwartzman A. Voxelwise multivariate analysis of multimodality magnetic resonance imaging. *Hum Brain Mapp*. 2014;35(3):831-846. [doi:10.1002/hbm.22217](https://doi.org/10.1002/hbm.22217)
16. Owen TW, de Tisi J, Vos SB, et al. Multivariate white matter alterations are associated with epilepsy duration. *Eur J Neurosci*. 2021;53(8):2788-2803. [doi:10.1111/ejn.15055](https://doi.org/10.1111/ejn.15055)
17. Chamberland M, Genc S, Tax CMW, et al. Detecting microstructural deviations in individuals with deep diffusion MRI tractometry. *Nat Comput Sci*. 2021;1(9):598-606. [doi:10.1038/s43588-021-00126-8](https://doi.org/10.1038/s43588-021-00126-8)
18. Marquand AF, Rezek I, Buitelaar J, Beckmann CF. Understanding Heterogeneity in Clinical Cohorts Using Normative Models: Beyond Case-Control Studies. *Biol Psychiatry*. 2016;80(7):552-561. [doi:10.1016/j.biopsych.2015.12.023](https://doi.org/10.1016/j.biopsych.2015.12.023)
19. Wolfers T, Doan NT, Kaufmann T, et al. Mapping the Heterogeneous Phenotype of Schizophrenia and Bipolar Disorder Using Normative Models. *JAMA Psychiatry*. 2018;75(11):1146-1155. [doi:10.1001/jamapsychiatry.2018.2467](https://doi.org/10.1001/jamapsychiatry.2018.2467)
20. Khedher L, Ramírez J, Górriz JM, Brahim A, Segovia F. Early diagnosis of Alzheimer's disease based on partial least squares, principal component analysis and support vector machine using segmented MRI images. *Neurocomputing*. 2015;151:139-150. [doi:10.1016/j.neucom.2014.09.072](https://doi.org/10.1016/j.neucom.2014.09.072)

21. Nestor PG, O'Donnell BF, McCarley RW, et al. A new statistical method for testing hypotheses of neuropsychological/MRI relationships in schizophrenia: partial least squares analysis. *Schizophr Res.* 2002;53(1):57-66. [doi:10.1016/S0920-9964\(00\)00171-7](https://doi.org/10.1016/S0920-9964(00)00171-7)
22. Calhoun VD, Adali T, Pearlson GD, Pekar JJ. A method for making group inferences from functional MRI data using independent component analysis. *Hum Brain Mapp.* 2001;14(3):140-151. [doi:10.1002/hbm.1048](https://doi.org/10.1002/hbm.1048)
23. Plitman E, Ochi R, Patel R, et al. Using Non-Negative Matrix Factorization to Examine Treatment Resistance and Response in Patients With Schizophrenia: A Multimodal Imaging Study. *Biol Psychiatry.* 2020;87(9):S350. [doi:10.1016/j.biopsych.2020.02.899](https://doi.org/10.1016/j.biopsych.2020.02.899)
24. Yang W, Lui RLM, Gao JH, et al. Independent Component Analysis-Based Classification of Alzheimer's Disease MRI Data. *J Alzheimers Dis.* 2011;24(4):775-783. [doi:10.3233/JAD-2011-101371](https://doi.org/10.3233/JAD-2011-101371)
25. Richie-Halford A, Yeatman JD, Simon N, Rokem A. Multidimensional analysis and detection of informative features in human brain white matter. *PLOS Comput Biol.* 2021;17(6):e1009136. [doi:10.1371/journal.pcbi.1009136](https://doi.org/10.1371/journal.pcbi.1009136)
26. Chamberland M, Raven EP, Genc S, et al. Dimensionality reduction of diffusion MRI measures for improved tractometry of the human brain. *NeuroImage.* 2019;200:89-100. [doi:10.1016/j.neuroimage.2019.06.020](https://doi.org/10.1016/j.neuroimage.2019.06.020)
27. Vandekar SN, Shinohara RT, Raznahan A, et al. Subject-level measurement of local cortical coupling. *NeuroImage.* 2016;133:88-97. [doi:10.1016/j.neuroimage.2016.03.002](https://doi.org/10.1016/j.neuroimage.2016.03.002)
28. Whitaker KJ, Vértés PE, Romero-Garcia R, et al. Adolescence is associated with genomically patterned consolidation of the hubs of the human brain connectome. *Proc Natl Acad Sci.* 2016;113(32):9105-9110. [doi:10.1073/pnas.1601745113](https://doi.org/10.1073/pnas.1601745113)
29. Carter F, Anwender A, Goucha T, et al. Assessing Quantitative MRI Techniques using Multimodal Comparisons. Published online June 21, 2022:2022.02.10.479780. [doi:10.1101/2022.02.10.479780](https://doi.org/10.1101/2022.02.10.479780)
30. Young K, Govind V, Sharma K, Studholme C, Maudsley AA, Schuff N. Multivariate statistical mapping of spectroscopic imaging data. *Magn Reson Med.* 2010;63(1):20-24. [doi:10.1002/mrm.22190](https://doi.org/10.1002/mrm.22190)
31. Carbonell F, Zijdenbos AP, Bedell BJ. Spatially Distributed Amyloid- β Reduces Glucose Metabolism in Mild Cognitive Impairment. *J Alzheimers Dis.* 2020;73(2):543-557. [doi:10.3233/JAD-190560](https://doi.org/10.3233/JAD-190560)
32. Chen C, Cao X, Tian L. Partial least squares regression performs well in MRI-based individualized estimations. *Front Neurosci.* 2019;13:1282. [doi:10.3389/fnins.2019.01282](https://doi.org/10.3389/fnins.2019.01282)
33. Hotelling H. Multivariate quality control. *Tech Stat Anal.* Published online 1947.
34. Alexopoulos EC. Introduction to Multivariate Regression Analysis. *Hippokratia.* 2010;14(Suppl 1):23-28.
35. Gyebnár G, Klimaj Z, Entz L, et al. Personalized microstructural evaluation using a Mahalanobis-distance based outlier detection strategy on epilepsy patients' DTI data – Theory, simulations and example cases. *PLOS ONE.* 2019;14(9):e0222720. [doi:10.1371/journal.pone.0222720](https://doi.org/10.1371/journal.pone.0222720)
36. Hayasaka S, Du AT, Duarte A, et al. A non-parametric approach for co-analysis of multi-modal brain imaging data: application to Alzheimer's disease. *NeuroImage.* 2006;30(3):768-779. [doi:10.1016/j.neuroimage.2005.10.052](https://doi.org/10.1016/j.neuroimage.2005.10.052)
37. Mahalanobis PC. On the generalized distance in statistics. In: National Institute of Science of India; 1936.
38. Ghorbani H. MAHALANOBIS DISTANCE AND ITS APPLICATION FOR DETECTING MULTIVARIATE OUTLIERS. *Facta Univ Ser Math Inform.* 2019;0:583-595. [doi:10.22190/FUMI1903583G](https://doi.org/10.22190/FUMI1903583G)
39. Kritzman M, Li Y. Skulls, Financial Turbulence, and Risk Management. *Financ Anal J.* 2010;66(5):30-41. [doi:10.2469/faj.v66.n5.3](https://doi.org/10.2469/faj.v66.n5.3)
40. Xiang S, Nie F, Zhang C. Learning a Mahalanobis distance metric for data clustering and classification. *Pattern Recognit.* 2008;41(12):3600-3612. [doi:10.1016/j.patcog.2008.05.018](https://doi.org/10.1016/j.patcog.2008.05.018)
41. Lindemer ER, Salat DH, Smith EE, Nguyen K, Fischl B, Greve DN. White matter signal abnormality quality differentiates mild cognitive impairment that converts to Alzheimer's disease from nonconverters. *Neurobiol Aging.* 2015;36(9):2447-2457. [doi:10.1016/j.neurobiolaging.2015.05.011](https://doi.org/10.1016/j.neurobiolaging.2015.05.011)
42. Guerrero-Gonzalez JM, Yeske B, Kirk GR, Bell MJ, Ferrazzano PA, Alexander AL. Mahalanobis distance tractometry (MaD-Tract) – a framework for personalized white matter anomaly detection applied to TBI. *NeuroImage.* 2022;260:119475. [doi:10.1016/j.neuroimage.2022.119475](https://doi.org/10.1016/j.neuroimage.2022.119475)

43. Kulikova S, Hertz-Pannier L, Dehaene-Lambertz G, Buzmakov A, Poupon C, Dubois J. Multi-parametric evaluation of the white matter maturation. *Brain Struct Funct*. 2015;220(6):3657-3672. [doi:10.1007/s00429-014-0881-y](https://doi.org/10.1007/s00429-014-0881-y)
44. Tremblay SA, Alasmar Z, Pirhadi A, et al. mvcomp: Multivariate Comparisons using Whole-brain and ROI-derived Metrics from MRI. Published online 2024. [doi:10.5281/zenodo.10713027](https://doi.org/10.5281/zenodo.10713027)
45. Wilks SS. Multivariate Statistical Outliers. *Sankhyā Indian J Stat Ser 1961-2002*. 1963;25(4):407-426.
46. Van Essen DC, Smith SM, Barch DM, et al. The WU-Minn Human Connectome Project: an overview. *NeuroImage*. 2013;80:62-79. [doi:10.1016/j.neuroimage.2013.05.041](https://doi.org/10.1016/j.neuroimage.2013.05.041)
47. Glasser MF, Sotiropoulos SN, Wilson JA, et al. The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage*. 2013;80:105-124. [doi:10.1016/j.neuroimage.2013.04.127](https://doi.org/10.1016/j.neuroimage.2013.04.127)
48. Tournier JD, Smith R, Raffelt D, et al. MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation. *NeuroImage*. Published online 2019:116137. [doi:10.1016/j.neuroimage.2019.116137](https://doi.org/10.1016/j.neuroimage.2019.116137)
49. Tustison NJ, Avants BB, Cook PA, et al. N4ITK: Improved N3 Bias Correction. *IEEE Trans Med Imaging*. 2010;29(6):1310-1320. [doi:10.1109/TMI.2010.2046908](https://doi.org/10.1109/TMI.2010.2046908)
50. Basser PJ, Mattiello J, LeBihan D. Estimation of the effective self-diffusion tensor from the NMR spin echo. *J Magn Reson B*. 1994;103(3):247-254. [doi:10.1006/jmrb.1994.1037](https://doi.org/10.1006/jmrb.1994.1037)
51. Basser PJ, Mattiello J, LeBihan D. MR diffusion tensor spectroscopy and imaging. *Biophys J*. 1994;66(1):259-267. [doi:10.1016/S0006-3495\(94\)80775-1](https://doi.org/10.1016/S0006-3495(94)80775-1)
52. Veraart J, Sijbers J, Sunaert S, Leemans A, Jeurissen B. Weighted linear least squares estimation of diffusion MRI parameters: strengths, limitations, and pitfalls. *NeuroImage*. 2013;81:335-346. [doi:10.1016/j.neuroimage.2013.05.028](https://doi.org/10.1016/j.neuroimage.2013.05.028)
53. Patenaude B, Smith SM, Kennedy DN, Jenkinson M. A Bayesian model of shape and appearance for subcortical brain segmentation. *NeuroImage*. 2011;56(3):907-922. [doi:10.1016/j.neuroimage.2011.02.046](https://doi.org/10.1016/j.neuroimage.2011.02.046)
54. Smith RE, Tournier JD, Calamante F, Connelly A. Anatomically-constrained tractography: improved diffusion MRI streamlines tractography through effective use of anatomical information. *NeuroImage*. 2012;62(3):1924-1938. [doi:10.1016/j.neuroimage.2012.06.005](https://doi.org/10.1016/j.neuroimage.2012.06.005)
55. Smith SM. Fast robust automated brain extraction. *Hum Brain Mapp*. 2002;17(3):143-155. [doi:10.1002/hbm.10062](https://doi.org/10.1002/hbm.10062)
56. Smith SM, Jenkinson M, Woolrich MW, et al. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*. 2004;23:S208-S219. [doi:10.1016/j.neuroimage.2004.07.051](https://doi.org/10.1016/j.neuroimage.2004.07.051)
57. Zhang Y, Brady M, Smith S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans Med Imaging*. 2001;20(1):45-57. [doi:10.1109/42.906424](https://doi.org/10.1109/42.906424)
58. Jeurissen B, Tournier JD, Dhollander T, Connelly A, Sijbers J. Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion MRI data. *NeuroImage*. 2014;103:411-426. [doi:10.1016/j.neuroimage.2014.07.061](https://doi.org/10.1016/j.neuroimage.2014.07.061)
59. Raffelt D, Tournier JD, Smith RE, et al. Investigating white matter fibre density and morphology using fixel-based analysis. *NeuroImage*. 2017;144:58-73. [doi:10.1016/j.neuroimage.2016.09.029](https://doi.org/10.1016/j.neuroimage.2016.09.029)
60. Dhollander T, Tabbara R, Rosnarho-Tornstrand J, Tournier JD, Raffelt D, Connelly A. Multi-tissue log-domain intensity and inhomogeneity normalisation for quantitative apparent fibre density. *Proc ISMRM*. Published online 2021.
61. Raffelt D, Dhollander T, Tournier JD, et al. Bias Field Correction and Intensity Normalisation for Quantitative Analysis of Apparent Fibre Density. *Proc ISMRM*. Published online 2017.
62. Raffelt D, Tournier JD, Fripp J, Crozier S, Connelly A, Salvado O. Symmetric diffeomorphic registration of fibre orientation distributions. *NeuroImage*. 2011;56(3):1171-1180. [doi:10.1016/j.neuroimage.2011.02.014](https://doi.org/10.1016/j.neuroimage.2011.02.014)
63. Glasser MF, Essen DCV. Mapping Human Cortical Areas In Vivo Based on Myelin Content as Revealed by T1- and T2-Weighted MRI. *J Neurosci*. 2011;31(32):11597-11616. [doi:10.1523/JNEUROSCI.2180-11.2011](https://doi.org/10.1523/JNEUROSCI.2180-11.2011)

64. Raffelt D, Tournier JD, Rose S, et al. Apparent Fibre Density: A novel measure for the analysis of diffusion-weighted magnetic resonance images. *NeuroImage*. 2012;59(4):3976-3994. [doi:10.1016/j.neuroimage.2011.10.045](https://doi.org/10.1016/j.neuroimage.2011.10.045)
65. Smith RE, Tournier JD, Calamante F, Connelly A. SIFT: Spherical-deconvolution informed filtering of tractograms. *NeuroImage*. 2013;67:298-312. [doi:10.1016/j.neuroimage.2012.11.049](https://doi.org/10.1016/j.neuroimage.2012.11.049)
66. Calamante F, Smith RE, Tournier JD, Raffelt D, Connelly A. Quantification of voxel-wise total fibre density: Investigating the problems associated with track-count mapping. *NeuroImage*. 2015;117:284-293. [doi:10.1016/j.neuroimage.2015.05.070](https://doi.org/10.1016/j.neuroimage.2015.05.070)
67. Daducci A, Canales-Rodríguez EJ, Zhang H, Dyrby TB, Alexander DC, Thiran JP. Accelerated Microstructure Imaging via Convex Optimization (AMICO) from diffusion MRI data. *NeuroImage*. 2015;105:32-44. [doi:10.1016/j.neuroimage.2014.10.026](https://doi.org/10.1016/j.neuroimage.2014.10.026)
68. Zhang H, Schneider T, Wheeler-Kingshott CA, Alexander DC. NODDI: Practical in vivo neurite orientation dispersion and density imaging of the human brain. *NeuroImage*. 2012;61(4):1000-1016. [doi:10.1016/j.neuroimage.2012.03.072](https://doi.org/10.1016/j.neuroimage.2012.03.072)
69. Aboitiz F, Scheibel AB, Fisher RS, Zaidel E. Fiber composition of the human corpus callosum. *Brain Res*. 1992;598(1):143-153. [doi:10.1016/0006-8993\(92\)90178-C](https://doi.org/10.1016/0006-8993(92)90178-C)
70. Chao Y, Cho K, Yeh C, Chou K, Chen J, Lin C. Probabilistic topography of human corpus callosum using cytoarchitectural parcellation and high angular resolution diffusion imaging tractography. *Hum Brain Mapp*. 2009;30(10):3172-3187. [doi:10.1002/hbm.20739](https://doi.org/10.1002/hbm.20739)
71. Hofer S, Frahm J. Topography of the human corpus callosum revisited--comprehensive fiber tractography using diffusion tensor magnetic resonance imaging. *NeuroImage*. 2006;32(3):989-994. [doi:10.1016/j.neuroimage.2006.05.044](https://doi.org/10.1016/j.neuroimage.2006.05.044)
72. Alexander-Bloch A, Giedd JN, Bullmore E. Imaging structural co-variance between human brain regions. *Nat Rev Neurosci*. 2013;14(5):322-336. [doi:10.1038/nrn3465](https://doi.org/10.1038/nrn3465)
73. Bilello M, Doshi J, Nabavizadeh SA, et al. Correlating Cognitive Decline with White Matter Lesion and Brain Atrophy Magnetic Resonance Imaging Measurements in Alzheimer's Disease. *J Alzheimers Dis*. 2015;48(4):987-994. [doi:10.3233/JAD-150400](https://doi.org/10.3233/JAD-150400)
74. Schmidt R, Ropele S, Enzinger C, et al. White matter lesion progression, brain atrophy, and cognitive decline: The Austrian stroke prevention study. *Ann Neurol*. 2005;58(4):610-616. [doi:10.1002/ana.20630](https://doi.org/10.1002/ana.20630)
75. González Ballester MÁ, Zisserman AP, Brady M. Estimation of the partial volume effect in MRI. *Med Image Anal*. 2002;6(4):389-405. [doi:10.1016/S1361-8415\(02\)00061-0](https://doi.org/10.1016/S1361-8415(02)00061-0)
76. Barhen A, Daudin JJ. Generalization of the Mahalanobis Distance in the Mixed Case. *J Multivar Anal*. 1995;53(2):332-342. [doi:10.1006/jmva.1995.1040](https://doi.org/10.1006/jmva.1995.1040)
77. de Leon AR, Carrière KC. A generalized Mahalanobis distance for mixed data. *J Multivar Anal*. 2005;92(1):174-185. [doi:10.1016/j.jmva.2003.08.006](https://doi.org/10.1016/j.jmva.2003.08.006)

SUPPLEMENTARY MATERIALS

Supplementary Materials

Download: <https://apertureneuro.org/article/118427-mvcomp-toolbox-multivariate-comparisons-of-brain-mri-features-accounting-for-common-information-across-measures/attachment/229113.docx>
