Frequency-dependent diffusion tensor distribution imaging in the evaluation of ischemic stroke

Sara Gröhn[1*], Ángela Naranjo[2*], Omar Narvaez[1*], Maxime Yon[3], Buse Buz-Yalug[1], Santos Blanco[2], Daniel Topgaard[4], Esther Martinez-Lara[2], Mª Ángeles Peinado[2], Jussi Tohka[1], Alejandra Sierra[1†]

[1]A.I. Virtanen Institute for Molecular Sciences; University of Eastern Finland, Finland
[2]Department of Experimental Biology, University of Jaén, Spain
[3]University of Rennes, LTSI, Inserm-U1099, Rennes, France
[4]Department of Chemistry, University of Lund, Sweden

Abstract

Non-invasive MRI is widely used to assess and monitor ischemic stroke, yet conventional approaches often lack sensitivity to subtle microstructural changes and struggle to evaluate tissue viability across lesion, penumbra, and distal regions. In this study, frequency dependent diffusion tensor distribution imaging (ωDTD) was combined with clustering of diffusion tensor distributions $\boldsymbol{D}(\omega)$ *and multivariate regression modeling to characterize ischemic tissue alterations in a whole brain section. Ex vivo ωDTD and histology were performed in rats subjected to middle cerebral artery occlusion (MCAO) or sham surgery (P = 17) 24 hours after reperfusion. Lesions showed cell loss and an increased presence of smaller, likely glial, cells. A random forest (RF) model was used to explain and predict histological parameters from diffusion tensor imaging (DTI), manually bin resolved ωDTD features, and cluster resolved ωDTD parameters. Model performance was evaluated using leave one animal out cross validation (LOO CV). ωDTD features better represented cell number than DTI metrics (ωDTD* R^2 *= 0.73 vs. DTI* R^2 *= 0.49), with similar advantages for nuclear area and circularity (ωDTD* R^2 *= 0.64 and 0.61 vs. DTI* R^2 *= 0.40 and 0.35). The RF model further proved beneficial in capturing complex, nonlinear relationships between MRI parameters and tissue characteristics. Overall, these results indicate that ωDTD provides richer microstructural information than standard DTI, and that combining ωDTD with advanced machine learning methods enhances interpretation of ischemic tissue damage.*

Keywords: Ischemic stroke, MCAO, MRI, diffusion tensor distribution imaging, frequency-dependent diffusion tensor distribution imaging, random forest (RF), non-linear regression, non-parametric distributions, unsupervised clustering, histology

[*]Equal contribution

[†]**Correspondence:** Alejandra Sierra, Ph.D.
A.I. Virtanen Institute for Molecular Sciences,
University of Eastern Finland (Kuopio Campus)
P.O. Box 1627 (Street address: Neulaniementie 2),
FI-70211 Kuopio, Finland
Tel. +358 40 355 2219, Fax. +358 17 163 030
E-mail: alejandra.sierralopez@uef.fi

INTRODUCTION

Ischemic stroke is an significant health and economic burden on a global scale, with up to 7.6 million incident cases annually [1]. During the ischemic event, the brain undergoes a series of complex biochemical and structural changes such as shifts in tissue composition and neuronal damage. Ischemic stroke arises from obstructed blood flow in a section of the brain leading to a lack of oxygen and nutrients in the affected area. This results in cellular injury and eventually cell death if blood flow is not restored [2, 3]. While traditional magnetic resonance imaging (MRI) techniques used in clinic e.g. T2-weighted imaging, diffusion MRI (dMRI), diffusion tensor imaging (DTI) and perfusion weighted imaging (PWI), are effective in detecting large-scale structural damage in the ischemic lesion core, the viability of the tissue in the core, penumbra, and distal areas to the lesion often remains undefined based on these methods [4, 5, 6, 7, 8]. These methods extract rather limited tissue information during pathological processes and often fail to detect subtle changes crucial to understand tissue viability or treatment responses after reperfusion. Therefore, Ischemic stroke requires novel solutions to these limitations.

Diffusion tensor distribution imaging (DTD) is an emerging MRI method with the potential to provide a solution to the challenges posed by conventional MRI. The advanced acquisition protocols of DTD utilize tensor valued diffusion encoding. This allows subvoxel parameter estimation of cellular features such as "size", "shape" and orientation, enabling a more nuanced and detailed analysis of tissue microstructure [9, 10]. Additionally, frequency-dependent DTD (ωDTD) exploits frequency dependence of diffusion encoding to estimate the effect of restriction [11, 12]. Restriction describes how the diffusion of water is constrained by the barriers of different tissue microenvironments, such as cellular membranes, at a few micrometers scale depending on the diffusion encoding frequency range. It can therefore provide more detailed insights into tissue composition and integrity for the targeted feature scale [13, 10, 14]. The parameters diffusivity, micro-anisotropy, and orientation more familiar from the literature [15, 9], are calculated from the frequency-dependent diffusion tensor distribution $\mathbf{D}(\omega)$ within each voxel. These parameter distributions can then be manually segmented into three bins to discriminate the specific signal-fractions within the voxel, such as those specific of white matter (WM), grey matter (GM) and free water (FW) like tissues [16, 10]. This simple binning operation based on two dimensions only allows computing specific parameter maps taking advantage of the correlation between all parameters [12, 9]. However, even in a healthy brain these manually divided "bins" can not accurately represent the complexity of the tissue and especially in pathological conditions, the binning becomes all the more challenging to interpret. To utilize more than two dimensions in the binning operation, we recently introduced Gaussian mixture-model based clustering of the tensor-valued distributions [17]. This approach allows the unsupervised and automated division of the per-voxel parameter distributions into k cluster-resolved fractions instead of three manually divided bin-resolved fractions. The additional solution space subsets produced by the clustering could potentially provide increased specificity in defining the ischemia related changes with MRI.

In this study, we combined ωDTD, unsupervised clustering, and a multivariate statistical model to understand the relationship between MRI and histology following ischemic stroke. Our approach utilises the power of non-linear regression, specifically random forests (RF), to explain and predict ischemic damage related tissue changes at 24-hours after reperfusion. Additionally, we investigate how the prediction accuracy depends on the input MRI parameters. We evaluated the microstructural changes related to ischemia based on altered cell morphology in the whole brain slice and included four different MRI parameter sets ranging from conventional DTI parameters to ωDTD cluster-fraction parameters. This study brings novelty to the field in specifically concentrating on exploring the relation between ωDTD and histological changes in ischemic stroke.

MATERIALS AND METHODS

Animals and stroke model

Seventeen adult male Wistar rats (250-450 g; Charles River, Wilmington, MA, USA) were used in this study. The rats were housed individually in a climate-controlled room under 12/12-hour light/dark cycle on an ad libitum diet. A transient ischemic stroke was induced in the right hemisphere using the middle cerebral artery occlusion (MCAO) model [18]. The MCA was occluded with an intraluminal filament (Doccol Corporation, Redlands, CA, USA) inserted through the right common carotid artery for 45 minutes before reperfusion as described in [19, 20]. The animals were divided into three groups. The first MCAO group received empty nanoparticles (NP) (MCAO+∅NP, n = 6). The second MCAO group received NPs filled with neuroglobin (MCAO+Ngb-NP, n = 6), an endogenous oxygen-binding globin protein found to play a neuroprotective role against

many brain injuries, including ischemia [18, 21, 22]. The NPs were injected through the tail vein. The third group was sham-operated, and also received empty NPs (sham, n = 5) to maintain experimental consistency and eliminate confounding variables.

At 24-h post-reperfusion, the animals were deeply anesthetized for a transcardial perfusion (left ventricle) using 5% isoflurane in O_2/N_2 mixture. This time point was selected to capture acute stroke pathology. Previous studies have shown that MCA occlusion in rodents produces peak infarct volume (assessed by TTC staining) at 24 h post-reperfusion, which remains stable through day 7, while apoptosis peaks at 24–48 h [23, 24, 25]. The perfusion was performed first with 150 ml 0.01 M, pH 7.4 phosphate-buffered saline (PBS) for 5 minutes at 30 ml/min and continued with 200 ml 4% paraformaldehyde (PFA) and 10 u/ml of heparin for 7 minutes also at 30 ml/min. Then, finished with 300 ml 4% PFA for 20 minutes at 15 ml/min. After perfusion, the brains were carefully extracted from the skull and post-fixed in the same solution for 4 h, before being placed 4% PFA until imaging. The animal model was provided and supervised by the Animal Production and Experimental Centre and the Ethical Committee for Animal Experimentation (CEEA, approvals #23/05/2016/090 and #10/02/2021/011) at the University of Jaén, Spain. All procedures were performed in accordance with Spanish legislation (law 32/2007, amended by law 6/2012 and by the Royal Decree 53/2013) and European Union Directive guidelines 2010/63/EU.

Ex vivo ωDTD acquisition and data processing

Ex vivo MRI data was acquired in a vertical 11.7-T Bruker Avance-III HD spectrometer equipped with a MIC-5 probe (on-axis 3 T/m maximum gradient amplitude) and 10-mm volumetric coil. The brains were divided into two hemispheres and the right hemisphere containing the ischemic lesion was imaged. The hemispheres were then immersed in gadoteric acid (0.0005 mmol/ml, Clariscan, GE HeatlhCare) in PBS pH 7.4 for 24 h prior to imaging to recover longitudinal and transverse relaxations and improve SNR for ex vivo [26]. Before MRI, the hemispheres were placed inside a 10 mm polyethylene tube and immersed in a perfluorinated compound (Galden, Solvay, Italy) to effectively suppress the background signal.

MRI data was obtained with a multi-slice multi-echo (MSME) 2D sequence customized for tensor-valued diffusion encoding. The sequence included variable b-values (700 - 8000 s/mm2), normalized anisotropy bΔ (-0.5, 0 , 0.5 and 1), TR = 800 ms, TE = 30 ms, orientation (θ,ϕ) and centroid frequency $\omega_{\text{cent}}/2\pi$ at 10% and 90% percentile (34 - 115 Hz). The images were acquired at approximately -0.60 mm from bregma with an in-plane voxel resolution of 80 x 80 μm^2, slice thickness of 250 μm and a matrix size of 150 x 150 x 1. In addition, anatomical data of the same anterior-posterior location was acquired using the rapid acquisition with relaxation enhancement with variable repetition time TR (RARE VTR) sequence to reduce scan time and sample the longitudinal relaxation more effectively. The total acquisition time was 4 hours and 12 minutes. After scanning, the brains were stored in 2% PFA solution to await histology processing.

The images were reconstructed using Bruker's ParaVision 6.0.1 [27] into real and imaginary domains and separately preprocessed using 1) a python-based denoising algorithm [28] based on the exploitation of data redundancy in principal component analysis domain, and 2) a MRtrix3 based Gibbs-ringing removal algorithm [29] based on local sub-voxel shifts. After preprocessing the real and imaginary domains were used to obtain the magnitude image.

Estimation of nonparametric D(ω) distributions

The nonparametric frequency-dependant diffusion distributions $\mathbf{D}(\omega)$ were derived through approximating the **b** encoded signal S as shown in [30]

$$S(\mathbf{b}) = \sum_i w_i \exp(-\mathbf{b} : \mathbf{D}_i). \tag{1}$$

The ω-dependent self diffusion tensor $\mathbf{D}(\omega)$ is acquired by inverting equation (1) and approximating the $\mathbf{D}_i(\omega)$ as an axisymmetric tensor,

$$\mathbf{D}_i(\omega) = \mathbf{R}(\theta_i, \phi_i) \begin{pmatrix} D_{\perp,i}(\omega) & 0 & 0 \\ 0 & D_{\perp,i}(\omega) & 0 \\ 0 & 0 & D_{\parallel,i}(\omega) \end{pmatrix} \mathbf{R}^{-1}(\theta_i, \phi_i) \tag{2}$$

where $\mathbf{R}(\theta_i, \phi_i)$ is a rotation matrix and parallel and perpendicular eigenvalues $D_{\parallel,i}(\omega)$ and $D_{\perp,i}(\omega)$ are given by Lorentzian transformations [30] at the frequencies $\Gamma_{\parallel/\perp,i}$ according to

$$D_{\parallel,i/\perp,i}(\omega) = D_{0,i} - (D_{0,i} - D_{\parallel,i/\perp,i}) \frac{1}{1 + \frac{\omega^2}{\Gamma^2_{\parallel,i/\perp,i}}}. \tag{3}$$

Using the axisymmetric Lorentzian approximation, the components of the $\mathbf{D}(\omega)$ distribution are described by the weight w and parameter set [D_{A}, D_{R}, θ, ϕ, D_0, Γ_{A}, Γ_{R},]. The data inversion is then performed using an extension of the Monte Carlo algorithm with bootstrapping as described in [12, 30, 31] using the md-dmri MATLAB toolbox [32].

The bootstrapping consisted of 100 repetitions using random sampling with replacement for 8 output components. The data inversion yielded an ensemble of 100 independent solutions of the $\mathbf{D}(\omega)$ per-voxel, each consisting of the 8 output components; weights and coordinates (w, [D_{A}, D_{R}, θ, ϕ, D_0, Γ_{A}, Γ_{R},]), within the pseudo-randomly sampled analysis space. The inversion limits were $D_{\mathrm{R/A/0}} \in [0.005, 5]\cdot 10^{-9}$ and $\Gamma_{\mathrm{R/A/0}} \in [0.000001, 1]\cdot 10^{5}$. Then the more familiar and meaningful parameters of isotropic diffusivity D_{iso}, normalized anisotropy D_{Δ} and normalized squared anisotropy D_{Δ}^2 were calculated through

$$D_{\mathrm{iso}}(\omega) = \frac{D_{\parallel,i}(\omega) + 2D_{\perp,i/\perp,i}(\omega)}{3}, \tag{4}$$

$$D_{\Delta}(\omega) = \frac{D_{\parallel,i}(\omega) - 2D_{\perp,i/\perp,i}(\omega)}{3D_{\mathrm{iso},i}(\omega)}, \tag{5}$$

and

$$D_{\Delta}^2(\omega) = \frac{[D_{\parallel,i}(\omega) - D_{\perp,i/\perp,i}(\omega)]^2}{[D_{\parallel,i}(\omega) + 2D_{\perp,i/\perp,i}(\omega)]^2}. \tag{6}$$

The effect of restriction on ω-dependent D_{iso} and D_{Δ}^2 has been explored in oscillating gradient encoding [33, 34] and other ωDTD [12, 35, 30, 36, 37] studies and can be quantified as

$$\Delta_{\omega/2\pi} D_{\mathrm{iso}} = \frac{D_{\mathrm{iso}}(\omega_{\mathrm{max}}) - D_{\mathrm{iso}}(\omega_{\mathrm{min}})}{\frac{\omega_{\mathrm{max}} - \omega_{\mathrm{min}}}{2\pi}}, \tag{7}$$

and

$$\Delta_{\omega/2\pi} D_{\Delta}^2 = \frac{D_{\Delta}^2(\omega_{\mathrm{max}}) - D_{\Delta}^2(\omega_{\mathrm{min}})}{\frac{\omega_{\mathrm{max}} - \omega_{\mathrm{min}}}{2\pi}}, \tag{8}$$

where ω is valued from ω_{min} to ω_{max} within the investigated frequency window (10% to 90%). The distributions can then be visualized as contour plots in a 2D D_{iso} - D_{Δ}^2 space as well as bin-resolved signal fractions $f_{\mathrm{bin}n}$, means $\mathrm{E}_{\mathrm{bin}n}[\cdot]$, variances $\mathrm{V}_{\mathrm{bin}n}[\cdot]$ and covariances $\mathrm{C}_{\mathrm{bin}n}[\cdot,\cdot]$ from which the per-voxel statistical descriptors values were obtained to display the quantitative parameter maps [9]. Here we have only considered the bin-resolved signal fractions $f_{\mathrm{bin}n}$, means $\mathrm{E}_{\mathrm{bin}n}[\cdot]$ calculated as

$$f_{\mathrm{bin}n} = \frac{1}{S_0} \sum_{l \in \mathrm{bin}n} w_l, n = 1,2,3 \tag{9}$$

and

$$\mathrm{E}_{\mathrm{bin}n}[z] = \frac{1}{S_0 f_{\mathrm{bin}n}} \sum_{l \in \mathrm{bin}n} w_l z_l, n = 1,2,3, \tag{10}$$

where z represents D_{iso}, $D_{\Delta}^2(\omega)$, $\Delta_{\omega/2\pi} D_{\mathrm{iso}}$ and $\Delta_{\omega/2\pi} D_{\Delta}^2$. Subscript l refers to the bootstrap samples and S_0 is the non-diffusion weighted image.

Additionally, basic DTI maps of mean diffusivity (MD) and fractional anisotropy (FA). These were derived from the diffusion tensor [38], calculated based on the orientation information in the ωDTD data.

Clustering of D(ω) distributions

The traditional manual binning is performed based on predefined limits of the isotropic D_{iso} and anisotropy D_{Δ}^2 space as follows; bin 1: $D_{\mathrm{iso}} \in$ [0,1] and $D_{\Delta}^2 \in$ [0.25,1], bin 2: $D_{\mathrm{iso}} \in$ [0,1] and $D_{\Delta}^2 \in$ [0,0.25] and bin 3: $D_{\mathrm{iso}} \in$ [1,2] and $D_{\Delta}^2 \in$ [0, 1] . The bin-resolved maps are produced based on this division. In turn, the clustering of the bootstrapped D_{iso}, D_{Δ} - for accessing information on prolate and oblate tensor shapes [39, 40] - and $\Delta_{\omega 2\pi} D_{\mathrm{iso}}$ per-voxel distributions was performed as described in [17] using Gaussian mixture model clustering. The clustering was performed with $k = 2, ..., 8$ to avoid predefining the number of clusters. BIC (Bayesian information criteria) was used to aid in the selection of optimal k (see supplementary figure S8). The spatial information from the clusters $clus_{\mathrm{n}}$ can be visualised with the per-voxel normalized frequency maps $F_{clus\mathrm{n}}, n = 1, 2, 3, ..., k$, where k is the number of clusters. Then, for the optimal k value, we produced the distributions and cluster-resolved signal fraction f_{clus} maps [17]. The cluster-resolved signal fractions can be obtained following a similar formulation as with the bin-resolved signal fractions. These cluster-resolved signal fractions f_{clus} can then be visualised using contours in the 2D D_{iso} - D_{Δ}^2 space and per-voxel means $\mathrm{E}_{\mathrm{clus}n}[\cdot]$ parametric maps calculated as

$$f_{\mathrm{clus}n} = \frac{1}{S_0} \sum_{l \in \mathrm{clus}n} w_l, n = 1,2,3,...,k \tag{11}$$

and

$$\mathrm{E}_{\mathrm{clus}n}[z] = \frac{1}{S_0 f_{\mathrm{clus}n}} \sum_{l \in \mathrm{clus}n} w_l z_l, n = 1,2,3,...,k \tag{12}$$

Histology

The hemispheres were first immersed in a cryoprotectant solution of glycerol 20% in 0.02 M potassium PBS (KPBS) pH 7.4 for 48 h before being frozen in dry-ice and stored at -80°C. Before sectioning, the brains were embedded in optimal temperature cutting medium. The sectioning was performed in the coronal plane at 30-μm slice thickness using a cryostat (Leica CM1950, Leica Microsistemas S.L.U., Spain). The sections were then stored in a cryoprotectant tissue-collecting solution (30% ethylene glycol, 25% glycerol in 0.05 M sodium phosphate buffer) at -20°C until staining. For the Nissl staining, the

sections were washed with 0.1 M PB (pH 7.4), then incubated in a chloroform and absolute ethanol (1:1) mixture. The sections were hydrated in descending concentrations of alcohol: absolute ethanol twice for 2 minutes; 96% for 2 minutes; 70% for 2 minutes; and 50% for 2 minutes. Then, the sections were distilled in water for 1 minute. Then, the sections were immersed in 0.125% thionine to assess both tissue damage severity and cytoarchitectonics until the appropriate coloring was achieved. The excess thionine was washed off with two washes using distilled water and then the sections were dehydrated increasing alcohol concentrations: 50% for 2 minutes; 70% for 2 minutes; 96% for 2 minutes and absolute ethanol twice for 2 minutes. Finally, 3 x 5 minutes washes were performed using xylol before mounting the sections in dibutylphthalate polystryrene xylene. The Nissl-stained sections were scanned at 20x magnification using a BX dotslide digital virtual microscope equipment, Olympus, at the Scientific Instrumentation Center of the University of Granada, Spain.

Non-linear regression and analysis

To assess the relationship between MRI and histology, we performed a non-linear regression analysis using the RF algorithm [41]. A total of $P = 11$ animals were included. From the original sample size of 17, 1 brain per group was allocated for serial-block electron microscopy (SBEM) analysis and therefore had an MRI acquisition not comparable with the MRI sequence described in this paper. Additionally, histology was performed for only one sham animal. For the regression analysis, we extracted quantitative parameter maps from the histology and selected the photomicrographs (0.3231 x 0.3231 μm^2) that gave the closest match to the MRI slices (approx -0.60 mm from bregma) for analysis. The photomicrographs were then divided to 80 x 80 μm^2 tiles before applying the QuPath watershed segmentation technique-based cell detection [42]. From the resulting cell detection measurements, we reconstructed three quantitative parameter maps depicting the number of cells per tile as well as the average nucleus area and circularity per tile. These histological parameter maps were then masked and downsampled using nearest neighbour interpolation. Lastly, the histological parameter maps were registered with the DTI FA map of the corresponding brain using a rigid similarity transform optimized by MATLAB's imregtform function. The registration minimized the negative mutual information [43] via the One-Plus-One Evolutionary (OPOE) algorithm [44, 45] suited for multimodal images. The maps were then prepared for the regression analysis by flagging possible negative (errors) and zero-valued voxels within the brain tissue as missing values. The pipeline is outlined in **Figure 1**.

The MRI data was then divided into four sets to evaluate, which one best predicted the histological parameters. The sets were: (1) conventional DTI-MRI consisting of MD and FA maps (2 maps); (2) ωDTD per-voxel mean and rates of change per frequency maps (4 maps: per-voxel mean $\mathrm{D_{iso}}$, D^2_Δ , $\Delta_{\omega 2\pi}\mathrm{E}[D_{\mathrm{iso}}]$ and, $\Delta_{\omega 2\pi}\mathrm{E}[D^2_\Delta]$; (3) ωDTD bin-resolved mean and rates of change per frequency maps (12 maps: bin-resolved $\mathrm{D_{iso}}$, D^2_Δ , $\Delta_{\omega 2\pi}\mathrm{E}[D_{\mathrm{iso}}]$ and, $\Delta_{\omega 2\pi}\mathrm{E}[D^2_\Delta]$); and (4) ωDTD cluster-resolved mean and rates of change per frequency maps (28 maps: cluster-resolved $\mathrm{D_{iso}}$, D^2_Δ , $\Delta_{\omega 2\pi}\mathrm{E}[D_{\mathrm{iso}}]$ and, $\Delta_{\omega 2\pi}\mathrm{E}[D^2_\Delta]$). The MRI data was prepared for the regression analysis as follows: (1) Individual maps were masked and processed for outliers by using the third quartile to remove such as, the bright halo effect around DTI MD maps. (2) The MRI datasets were scaled using z-score normalization. This preprocessing was done to help optimize the data for the RF regression task [46].

The RF analysis was performed using the TreeBagger function by MATLAB with 1000 trees. The predictor data matrix was built as

$$X = \begin{array}{c} \begin{array}{ccccc} \text{MRI parameter 1} & \text{MRI parameter 2} & \text{MRI parameter 3} & & \text{MRI parameter m} \end{array} \\ \left[\begin{array}{ccccc} x_{11}^{j=1} & x_{12}^{j=1} & x_{13}^{j=1} & \cdots & x_{1m}^{j=1} \\ x_{21}^{j=1} & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ x_{h1}^{j=1} & \ddots & \ddots & \ddots & \vdots \\ x_{(h+1)1}^{j=2} & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ x_{N1}^{j=11} & \cdots & \cdots & \cdots & x_{Nm}^{j=11} \end{array}\right] \begin{array}{l} \left.\begin{array}{l} \\ \\ \\ \\ \end{array}\right\} \text{Animal 1} \\ \left.\begin{array}{l} \\ \\ \end{array}\right\} \vdots \\ \left.\begin{array}{l} \\ \end{array}\right\} \text{Animal 11} \end{array} \end{array} \tag{13}$$

where m is the number of MRI parameters per set and the MRI maps from all animals j were concatenated row-wise, with the number of datapoints all together being N. The target Y (Nx1) was built following the same logic as X

$$Y = \left[\begin{array}{c} y_{11}^{j=1} \\ y_{21}^{j=1} \\ \vdots \\ y_{h1}^{j=1} \\ y_{(h+1)1}^{j=2} \\ \vdots \\ y_{N1}^{j=11} \end{array}\right] \begin{array}{l} \left.\begin{array}{l} \\ \\ \\ \\ \end{array}\right\} \text{Animal 1} \\ \left.\begin{array}{l} \\ \\ \end{array}\right\} \vdots \\ \left.\begin{array}{l} \\ \end{array}\right\} \text{Animal 11} \end{array} \tag{14}$$

A separate model was trained for each individual histological and MRI set combination. Based on the trained RF models, we predicted the histological parameter maps.

The performance of the RF models was assessed using leave-one-animal-out cross-validation (LOO

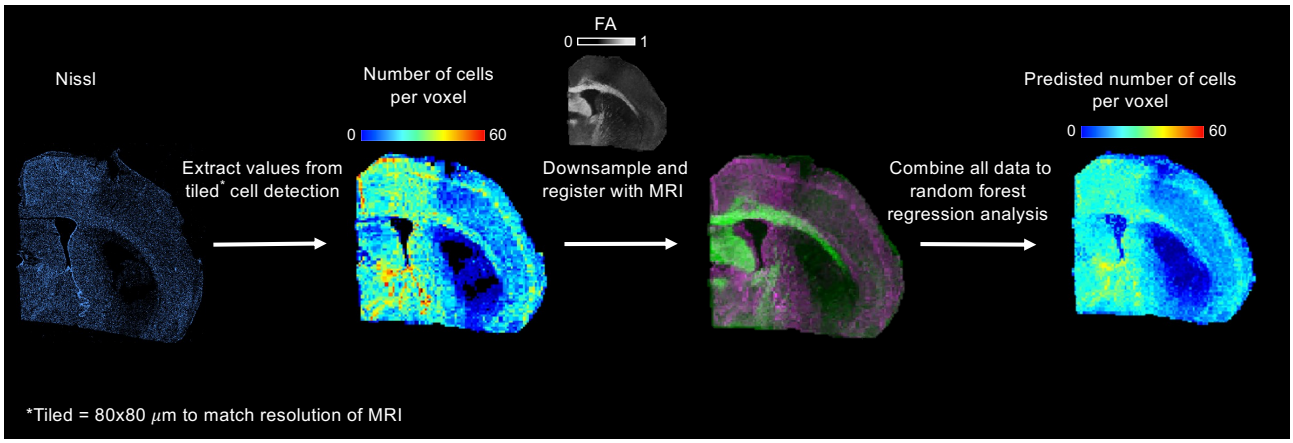

Figure 1: ***Overview of the non-linear regression analysis pipeline.*** *For each brain, the Nissl-stained photomicrographs were tiled to 80 x 80 μm2 sections to match the resolution of the MRI data. Cell detection was performed on these images to extract quantitative information. Tiled histological parameter maps were then reconstructed, downsampled and individually registered with the corresponding DTI FA map. For the non-linear regression analysis, MRI parameters were grouped as follows: (1) conventional MRI consisting of T2-weighted, MD and FA maps; (2) ωDTD per-voxel mean and rates of change per frequency maps; (3) ωDTD bin-resolved mean and rates of change per frequency maps; and (4) ωDTD cluster-resolved mean and rates of change per frequency maps. For each MRI parameter set, all brains were combined into a single dataset, which was used to predict the histological parameters using a with 1000 trees.*

CV) as in [47] using the mean (E[·]) and variance (V[·]) of the different metrics:

$$E[Q^2] = \frac{1}{11}\sum_{j=1}^{11}\left(1 - \frac{\sum_{i=1}^{n(j)}(y_i^{(j)} - \hat{y}_i^{(j)})^2}{\sum_{i=1}^{n(j)}(y_i^{(j)} - \bar{y}_i^{(j)})^2}\right) \tag{15}$$

$$E[\mathrm{MAE}] = \frac{1}{11}\sum_{j=1}^{11}\left(\frac{\sum_{i=1}^{n(j)}|(y_i^{(j)} - \hat{y}_i^{(j)})|}{n(j)}\right) \tag{16}$$

$$E[\mathrm{MSE}] = \frac{1}{11}\sum_{j=1}^{11}\left(\frac{\sum_{i=1}^{n(j)}(y_i^{(j)} - \hat{y}_i^{(j)})^2}{n(j)}\right) \tag{17}$$

$$\mathrm{V}[z] = \frac{1}{11}\sum_{i=1}^{11}(z^{(i)} - \mathrm{E}[z])^2, \tag{18}$$

where $y_j^{(i)}$ is the true histological parameter of animal j at the voxel i, $n(j)$ is the number of datapoints for animal j, $\hat{y}_i^{(j)}$ is the predicted histological parameter and $\bar{y}^{(j)}$ is the average of the true histological values. The Q^2 metric sheds light on the model's ability to generalize as it provides an internal measure for consistency between CV predictions and original data [48], while MSE and MAE tell of the accuracy and generalizability of the model as well as potential overfitting [46, 49]. The R^2 [50] was used as a performance measure to estimate the goodness-of-fit for the model. The CV analysis was performed using an in-house MATLAB code, available at https://github.com/jussitohka/mrihistology.

To compare the predictive performance of the different MRI feature sets, we used a paired permutation test (100,000 permutations) on the subject-wise performance metrics (R, Q^2, MSE, and MAE). For each histological parameter and performance metric, we treated every MRI method–histology combination as a separate model and then tested all pairwise differences between these models (e.g. DTI–cellularity vs ωDTD per-voxel mean and rates of change with frequency maps–cellularity). For a given metric and histological parameter, the test statistic was the mean within-subject difference between two methods (e.g. R for DTI minus R for ωDTD per-voxel mean and rates of change with frequency maps). This approach avoids distributional assumptions such as normality of differences and is well suited to the relatively small sample sizes and potentially non-Gaussian performance metrics common in imaging-based prediction studies.

Shapley additive explanations (SHAP) were calculated with the shapley and fit functions in MATLAB and used to obtain local feature attributions for a subset of 40 voxels per animal across cross-validation folds. This subset size was chosen to balance representativeness with the substantial computational cost of SHAP estimation, and voxels were strategically selected to capture both lesioned and adjacent tissue. Local Shapley values were then aggregated across query points, folds, and animals to derive global feature importance, following the framework proposed for tree ensembles by [51], where many local explanations are combined to characterize global model behavior.

RESULTS

Qualitative ischemic changes in conventional- and bin-resolved ωDTD parameter maps

Figure 2A and **D** show a conventional T2-weighted anatomic map along with DTI parameter maps mean diffusivity (MD) and fractional anisotropy (FA) for representative sham-operated and MCAO rats. The anatomic T2-weighted map of the MCAO brain (**Figure 2D**) shows a slightly increased signal intensity within the lesion, which is consistent with increased presence of water in the region. Both the MD and FA maps, in turn, show a decrease in signal values in the lesion. The MD map provides a visually strong contrast for the lesion with clearly delineable borders.

The ωDTD parameters display multiple contrasts and highlight different regions of the lesion depending on the parameter map. **Figure 2B** and **E** show the rates of change with frequency maps $\Delta_{\omega 2\pi}\mathrm{E}[D_{\mathrm{iso}}]$ and $\Delta_{\omega 2\pi}\mathrm{E}[D_{\Delta}^2]$) derived from the $\mathbf{D}(\omega)$ distributions. The classic figure for displaying the ωDTD parameter maps, as in e.g. [12, 30], can be seen in supplementary figure S9) for a representative sham brain. The $\Delta_{\omega 2\pi}\mathrm{E}[D_{\mathrm{iso}}]$ map shows excellent lesion contrast across all the MCAO brains. The bin-resolved fractions, means and rates of change per frequency maps in **Figure 2C** and **F** depict the traditional manual division of the $\mathbf{D}(\omega)$ distributions into "bins". We present these classical bin-resolved maps as a qualitative point of reference for the clustering-based approach.

The MCAO frequency-dependent $\Delta_{\omega 2\pi}\mathrm{E}[D_{\mathrm{iso}}]$ map shows visually a increased signal in the somatosensory cortex as well as the striatum. The $\Delta_{\omega 2\pi}\mathrm{E}[D_{\Delta}^2]$) map also shows an increase in values in those same areas. The $\mathrm{E}[D_{\Delta}^2]$ map of f_{bin1} shows increased values in the external capsule indicating more uniform orientation. In turn, the $\mathrm{E}[D_{\mathrm{iso}}]$ of f_{bin1} , shows signal decrease in the external capsule denoting slower diffusion. Both maps also demonstrate changes in the somatosensory cortex when compared qualitatively to the representative sham-operated rat in **Figure 2B**. The $\Delta_{\omega 2\pi}\mathrm{E}[D_{\mathrm{iso}}]$ map of f_{bin1} shows an apparent increase in the values in the external capsule, striatum and somatosensory cortex (blue arrows). The $\mathrm{E}[D_{\mathrm{iso}}]$ map of f_{bin2} delineates visually very clear lesion borders similar to the MD map of **Figure 2D** and along with the $\mathrm{E}[D_{\Delta}^2]$ map of f_{bin2} shows hypointensity in the somatosensory cortex as well as the striatum. The $\Delta_{\omega 2\pi}\mathrm{E}[D_{\mathrm{iso}}]$ and $\Delta_{\omega 2\pi}\mathrm{E}[D_{\Delta}^2]$) maps of f_{bin2} show increased signal values in those same regions.

Clustering of D(ω) distributions

The parametric distributions of D_{iso}, and D_{Δ}^2 are presented in Fig. 3 for representative sham-operated and MCAO brains derived from the unsupervised clustering of the $(\mathrm{D})_{\omega}$, distributions. The optimum k based on the BIC (see supplementary figure S8) and [17] was 7. Using k = 7, the clustering consistently produced distributions presenting identifiable characteristics for each cluster-fraction (f_{clust}) as can be seen in **Figure 3** . The first cluster-fraction is a WM-like fraction (f_{clust1}), with D_{iso}, and D_{Δ} distributions depicting low diffusivities, along with high anisotropy, diffusion characteristics typical in WM-like tissue, such as the corpus callosum and external capsule, but also including WM in cortical layers. The second cluster-faction is a D_{Δ} = -0.5 fraction (f_{clust2}), followed by fraction of oblate shapes sensitive to frequency (f_{clust3}). These first three $\mathrm{E}[D_{\mathrm{iso}}]$ cluster-fraction distributions for the MCAO also displayed a two peaked distribution in comparison to the single peaked appearance of the distributions in the sham-operated brains (black arrows in **Figure 3**).

The fourth and sixth cluster-fractions (f_{clust4} and (f_{clust6}) were characterised by high diffusivities and a widespread in the D_{Δ} distributions along with a dependency on frequency, indicating restriction. The fifth cluster-fraction is a FW-like fraction (f_{clust5}) defined by with high diffusivity and low anisotropy values. And finally, a last spurious cluster-fraction (f_{clust7}) also appeared consistently for all brains and is omitted from all analyses. For plotting the frequency dependency of these parameters, we used centroid frequencies 34.4, 74.7 and 115 $\omega/2\pi$ Hz.

Figure 4 shows the cluster-resolved mean parametric maps for representative sham-operated and MCAO brains derived from the unsupervised clustering of the $(\mathrm{D})_{\omega}$, distributions using k = 7. All the cluster-fractions, apart from the FW-like fraction (f_{clust5}), showed distinct lesion contrast. While the cluster-resolved maps of $\mathrm{E}[D_{\mathrm{iso}}]$ showed a decrease in values in the lesion for the first three fractions, the lesion borders vary between the maps. The f_{clust1} depicts the changes in WM-like tissues, and while the different parameter maps show minor changes in the external capsule, the corpus callosum remains unchanged when compared to the corresponding sham-operated maps (**Figure 4A**). These f_{clust1} maps also show changes in the WM-like structures in the cortex, e.g. cortical axons. The $\Delta_{\omega 2\pi}\mathrm{E}[D_{\Delta}^2]$ map of f_{clust1} shows prominent contrast for the lesion in of the somatosensory cortex.

The $\mathrm{E}[D_{\mathrm{iso}}]$ map of f_{clust2} depicts a diffuse and irregular lesion boundary, especially between the motor and somatosensory cortex. In comparison, the lesion boundaries in the $\mathrm{E}[D_{\mathrm{iso}}]$ map of f_{clust1}

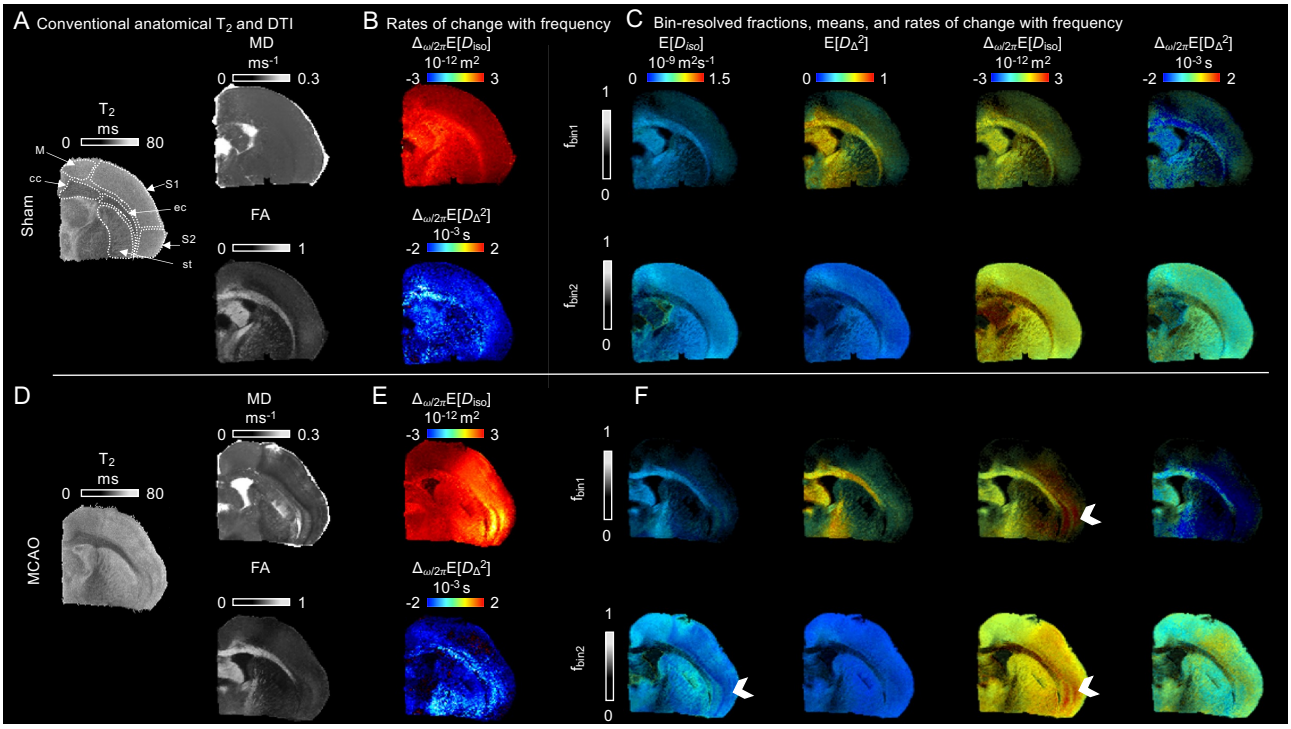

Figure 2: ***Conventional anatomical T2 and DTI metrics map compared to ωDTD parameter maps for representative sham and MCAO brains.*** ***(A)*** *and* ***(D)*** *Conventional T2-weighted anatomic maps together with classic DTI parameter maps of MD and FA.* ***(B)*** *and* ***(E)*** *ωDTD maps showing per-voxel rates of change with frequency, along with* ***(C)*** *and* ***(F)*** *bin-resolved parametric maps for manually defined bin limits in the same representative brains (****equations*** (4),(6), (7)***, and*** (8)*). The brightness of the bin-resolved maps reflects signal fraction, while the color indicates parameter values. The ωDTD maps were computed using the ωDTD distribution (****equation*** (2)*) at* $\omega/2\pi$ *= 34.4 Hz and the rates of change with frequency maps were calculated with* $\omega/2\pi$ *= 34.4-115 Hz. The labels are M = motor cortex, S1 = primary somatosensory cortex, S2 = Secondary somatosensory cortex, cc = corpus callosum, ec = external capsule, and st = striatum. The arrows indicate the lesion in the somatosensory cortex.*

and f_{clust3} are distinct. The f_{clust4} and f_{clust6} demonstrate sensitivity to the lesion specifically, with f_{clust6} showing primarily the lesion and the visibility of the surrounding brain structures is minimal (white arrows in **Figure 4B**). The $\mathrm{E}[D_{\text{iso}}]$ of f_{clust4} in turn shows hyperintensity in both primary and secondary somatosensory cortex, while f_{clust6} predominantly has signal only the secondary somatosensory cortex, the epicentre for the most intense damage. However, when comparing the lesion boundaries within the f_{clust4} and f_{clust6}, there is a correspondence across the parameters. In turn, the $\Delta_{\omega 2\pi}\mathrm{E}[D_{\Delta}^2]$ of f_{clust4} displays decreased signal values inside the lesion region. Additionally, $\Delta_{\omega 2\pi}\mathrm{E}[D_{\text{iso}}]$ shows high values in f_{clust4} and f_{clust6} indicating high values of restriction.

Relationship between histological and MRI parameters

A photomicrograph of a Nissl-stained section from a representative MCAO brain is shown in **Figure 5A**. The lesion was primarily located in the somatosensory cortex and was characterized by extensive cell loss and a presence of smaller cells that are most likely glial cells. The boundary of the lesion in the motor and primary somatosensory cortex (S1) demonstrated a clear transition between adjacent and damaged tissue, highlighting the regionality of the ischemic damage. The lesion boundary in the primary somatosensory cortex is reflected in the MRI parameter maps. The striatum (st) exhibited a significant loss of structures. The corpus callosum (cc) and the motor cortex (M) (see **Figure 5A**) were intact, exhibiting normal cellular density.

The results of the RF regression fitted for the histological parameters in their original scale are summarized in **Table 1**. The R values for the CV LOO were positive for all the histological parameters and MRI parameter set combinations, indicating that the models captured a relationship between MRI and histology. **Table 1** also shows an increase in the performance metrics for all histological parameters associated with the change of the MRI parameter set from conventional DTI to ωDTD. The variation of these metrics was minimal inside the ωDTD parameter sets. Based on the R^2 and Q^2 values, the model fits (training error) were more accurate than the out-of-sample predictions suggesting overfitting. The best model of cluster-resolved ωDTD for number of cells explained 73% of the variance in the data. The improved performance of the model associated with the changing of MRI parameter sets from conventional to ωDTD can be seen in supplementary figure S10.

The permutation test results comparing DTI against the three ωDTD methods are summarized in Table 2. Across all metrics and histological parameters, ωDTD methods showed significant superior predictive performance over conventional DTI in 22 out of 36 comparisons ($p<0.05$), with particularly strong effects for number of cells prediction where all three ωDTD variants significantly outperformed DTI ($p<0.05$ across R, Q^2, MSE, and MAE). For area of nucleus, 9 out of 12 comparisons significantly favored ωDTD, with per-voxel and bin-resolved methods showing consistent significance across metrics. Circularity of nucleus showed the weakest MRI-histology relationships overall, with only 3 significant comparisons. P-values showing significantly superior DTI were rare, while p-values showing superior ωDTD dominated across all metrics. The rest of the permutation testing results can be found in Table S3.

The normalised global SHAP mean absolute values for the top ten features of the ωDTD cluster-resolved mean and rates of change per frequency maps can be seen in figure 6. The $\mathrm{E}[D_{\text{iso}}]$ of f_{clust1} emerged as the dominant feature, followed by $\mathrm{E}[D_{\text{iso}}]$ of f_{clust4} and $\mathrm{E}[D_{\Delta}^2]$ of f_{clust1}. The different clusters of $\mathrm{E}[D_{\text{iso}}]$ occupied six of the top ten positions in global feature importance. And while $\mathrm{E}D_{\text{iso}}$ of f_{clust1} ranked highest overall, $\mathrm{E}D_{\text{iso}}$ of f_{clust3} and f_{clust4} demonstrated the greatest consistency within the top ten global features.

The predicted histological parameter maps based on the ωDTD cluster-resolved MRI parameter set is shown in **Figure 7**. There was no significant underestimation of values when comparing the original histological parameter maps (**Figure 7A**) and to predicted ones (**Figure 7B**) as shown by the MAE values in **Table 1**. The RF based predictions also retained information on the different anatomical structures, separating between WM and GM as well as the ventricles, which are distinguishable, though they are not valued as null as in the original maps. The model predictions also demonstrated the capacity to separate between the lesion and adjacent tissue. Each histological parameter map of the MCAO brain presented in **Figure 7B** shows subtle differences in lesion boundaries, differences that are replicated in the predicted histological parameter maps (**Figure 7D**). The average difference (in percentages) between the prediction and original show smaller differences in the circularity of nucleus parameter, when compared to the area of nucleus and cellularity parameters.

Overall, the number of cells histological parameter in the MCAO brains show there were less cells in the lesion in comparison to the lesion adjacent tissue. The area and circularity of nucleus parameter maps of the MCAO brains also show decreased nucleus area and increased circularity in the lesion. The distinct boundaries of the lesion within the histological parameters were accurately captured by the predictions.

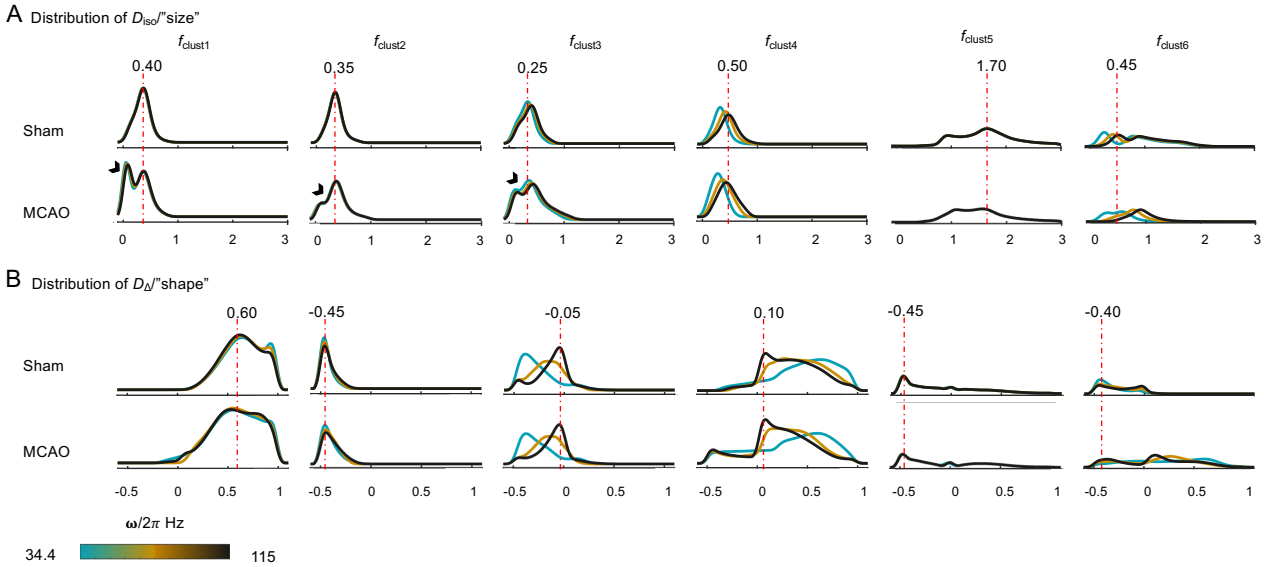

Figure 3: ***Distributions of D_{iso} and D_{Δ} for the cluster-fractions for representative sham and MCAO brains.*** *The dotted red lines indicate the more prevalent value of D_{iso} and D_{Δ} for the sham-operated brain. Different frequencies are represented by colors; 34.4 $\omega/2\pi$ Hz in blue , 74.7 $\omega/2\pi$ Hz in brown and 115 $\omega/2\pi$ Hz in black. The arrows point to the dual peak appearance of the D_{iso} distribution of the MCAO brain.*

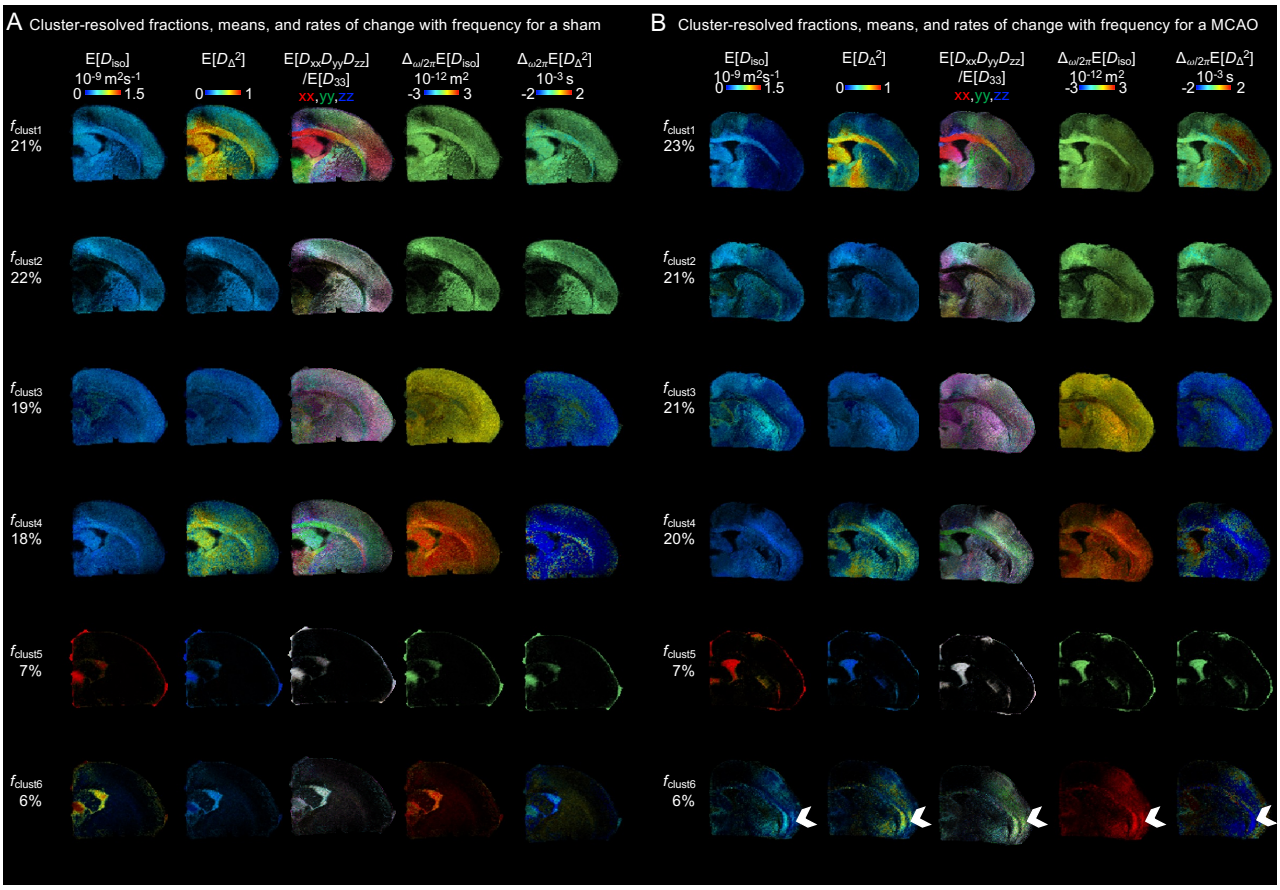

Figure 4: ***Cluster-resolved ωDTD parameter maps derived from $\mathbf{D}(\omega)$ distributions for representative sham and MCAO brains.*** ***(A)*** *and* ***(B)*** *Cluster-resolved ωDTD maps showing the cluster-fraction, with percentages indicating the per cluster normalized frequency in representative sham and MCAO brains. The color of the cluster-resolved maps indicates parameter values. The ωDTD maps were computed using the $\mathbf{D}(\omega)$ distribution (eq. 2) at $\omega/2\pi$ = 34.4 Hz and the rates of change with frequency maps were calculated with $\omega/2\pi$ = 34.4-115 Hz. The arrows indicate the lesion in the somatosensory cortex.*

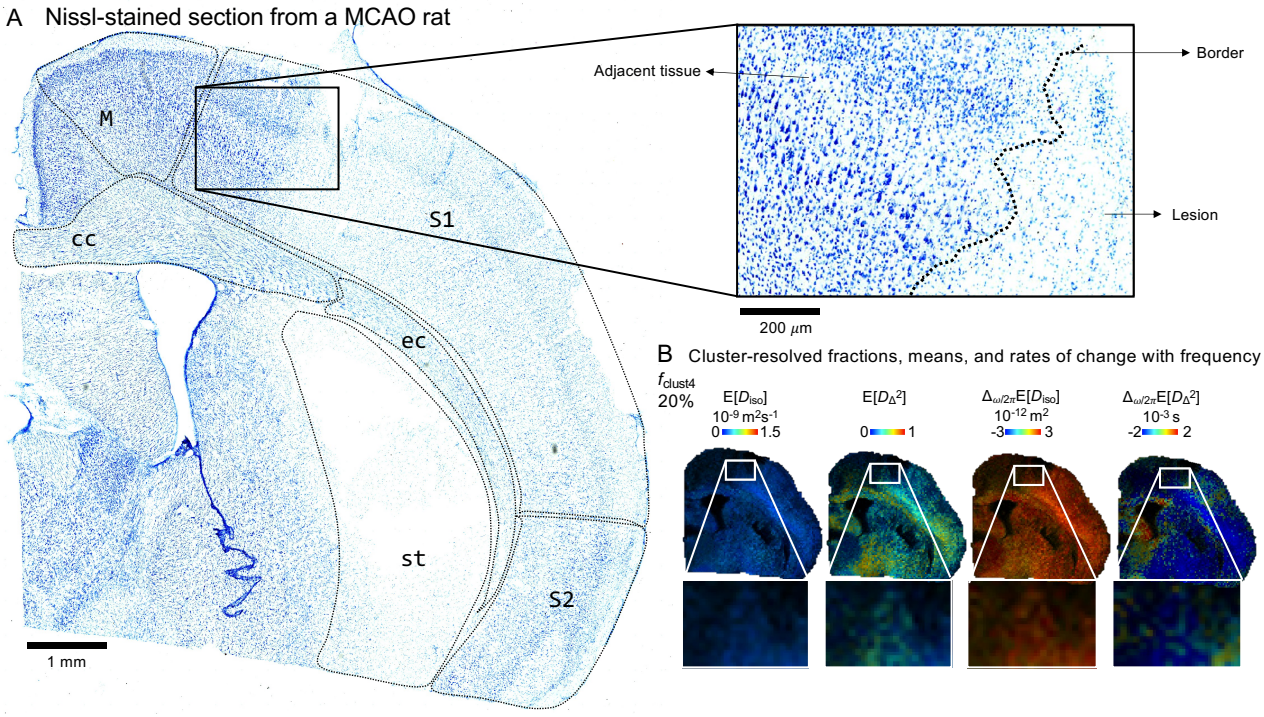

Figure 5: ***Comparison of Nissl-stained section with cluster-resolved parametric maps from a reprasentative MCAO rat.*** ***(A)*** *A Nissl-stained section of a representative MCAO brain showing the overall tissue damage, with the black arrow indicating ischemic lesion.* ***(B)*** *Cluster-resolved ωDTD maps of f_{clust4} of the same brain, which represent the fraction associated with the lesion. The high-magnification of the Nissl photomicrograph* ***(A)*** *and close-ups of the MRI parameter maps* ***(B)*** *display the same lesion boundary region in the somatosensory cortex. The closer view of the Nissl shows the shrinkage and loss of neurons in the lesion typical in ischemia associated neurodegeneration. The labels are M = motor cortex, S1 = primary somatosensory cortex, S2 = Secondary somatosensory cortex, st = striatum, cc = corpus callosum, and ec = external capsule.*

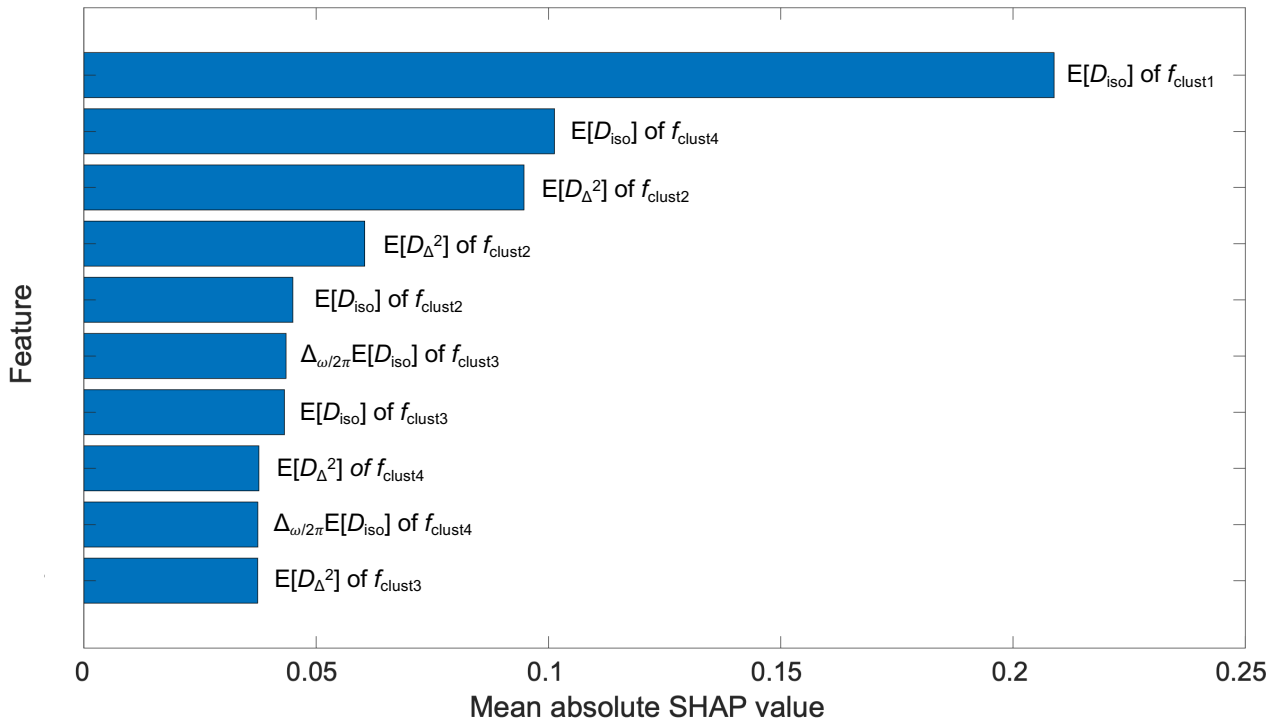

Figure 6: ***Normalised global SHAP mean absolute values for random forest feature importance.*** *Global SHAP values were calculated in the CV LOO folds for out-of-loop subjects for the ωDTD cluster-resolved mean and rates of change per frequency maps.*

The distinct details are also reproduced consistently, such as the layer of increased number of cells in the cortex (**Figure 7B** and **D**). The baseline values of nucleus circularity in the secondary somatosensory cortex, an area that was otherwise characterized by a decreased cell number and nucleus area by the other two histological parameters. The original and predictions of all three histological parameters and all the brains are shown in supplementary figure S11.

DISCUSSION

In this study, we investigated the relationship between *ex vivo* ωDTD parameters and histological parameters in the brain after ischemic damage. We examined the capacity of these MRI parameters in predicting tissue changes seen in the histology. We had four sets of MRI parameters: conventional DTI, ωDTD, bin-resolved ωDTD, and cluster-resolved ωDTD parameters. We found that the ωDTD parameters were highly predictive of changes in the cellularity and morphology, including alterations in nucleus size and shape, of cells, more than conventional MRI parameters. Furthermore, we found that some of the histology metrics could be predicted by using a multivariate non-linear regression modeling approach. Standard voxelwise and ROI-based analyses yielded no significant findings, likely due to the small sample size and heterogeneity of the ischemic lesion. Altogether, these results represent a significant step forward in the interpretation of ωDTD parameters, showing how they potentially reflect the underlying changes in tissue microstructure after ischemic stroke.

There are numerous studies on using DTI for analyzing ischemic stroke related pathology in both humans and rodents at acute timepoints [52, 53, 54, 55]. Traditionally, DTI is used for detecting the lesion in clinics. MD and FA typically show lower values at an acute timepoint. Our study demonstrates convergent validity in the decrease detected in the lesion in both MD and FA. However, these conventional parameter maps lack specific information on the underlying pathological processes [6, 56]. In studies comparing DTI with more advanced methods, DTI has consistently been found to be less sensitive in detecting subtle microstructural changes and separating specific processes, particularly in cases of early ischemic injury [57, 58, 59].

There are no studies focused on investigating the relationship between ωDTD and quantitative histological parameters in the case of ischemic stroke. Our previous studies show that ωDTD parameter maps reflect distinct cellular properties [30, 37], which can potentially provide specific insights into the cytopathology through differential lesion segmentation, e.g. the $\mathrm{E}[D_{\mathrm{iso}}]$ map in gray matter-like tissues has been linked to cell size. The decrease in values of this parameter seen in the ischemic lesion may be related to the presence of glial cells and neuronal degeneration. The increased values of $\Delta_{\omega 2\pi}\mathrm{E}[D_{\mathrm{iso}}]$ have been linked to increased cellularity in tumors [60, 61]. However, in the case of stroke lesions, which also show increased values in both in bin- and cluster-resolved f_{clust4} and f_{clust6} maps, this may instead be linked to localized debris such as cellular debris and protein aggregation, creating temporary diffusion barriers. Within the framework of multidimensional MRI (MD-MRI), which encompasses advanced techniques like ωDTD, previous studies have shown the stability of both MD-MRI parameters and MD-MRI-derived clusters across subjects [62, 17].

Nissl staining shows a decreased number of cells in the lesion, so another possible reason for the increased values could be changes in the increased extracellular space. The oscillating gradient spin-echo (OGSE) can provide comparison for ωDTD [63], although OGSE probes stick-like tensor shapes at different frequencies, while with ωDTD we can access a wide range of shapes in the tensor distribution space at different frequencies [30, 35]. We found studies on OGSE in evaluating ischemic stroke. These studies state that necrotic tissue often exhibits increased viscosity or heterogeneous environments due to cellular debris/fragmented membranes, which might amplify restriction effects detectable through OGSE [5, 64, 65]. The decreased values in f_{clust2} across all the cluster-resolved ωDTD parameters might be reflective also of the breakdown of structural integrity in the lesion, as this cluster-fraction might potentially describe the dense intracellular networks of neurons [12, 17]. The increased values in $\mathrm{E}[D_{\Delta}^2]$ of f_{clust4} could further support an increase in structural disorder in the lesion indicating the potential of heterogeneous environments in the lesion.

The coefficient of determination R^2 is used to evaluate the model's performance on training data. Based on the R^2 values, our models can explain between 35% - 70% of the variance present in the data depending on the histological and MRI parameter set. We also found that the number of cells parameter was more substantially associated with the MRI data than the area and circularity of the soma. And even though, on average, the difference between the original and predicted circularity of nucleus parameter measured in percentages was the smallest, the small value range of 0 to 1 means even small errors impact the variance and affect the R^2. The consistently positive correlation R in the CV LOO shows that our models repeatedly capture a relationship between the

Histological parameters MRI parameter sets	LOO CV								Model
Number of cells Area of nucleus Circularity of nucleus	E[R]	V[R]	E[Q^2]	V[Q^2]	E[MSE]	V[MSE]	E[MAE]	V[MAE]	R^2
Conventional DTI	0.2275	0.0156	-0.1292	0.0416	103.5989	368.4062	7.8298	0.8398	**0.4853**
	0.1289	0.0014	-0.0894	0.0066	55.4998	298.5853	5.0937	0.4923	**0.3978**
	0.1709	0.0048	-0.0244	0.0034	0.0249	1.7777e-4	0.0772	4.2602e-4	**0.3510**
ω DTD Per-voxel	0.2998	0.0135	-0.0429	0.0349	95.1086	194.6423	7.4462	0.5337	**0.6624**
	0.2324	0.0051	-0.0548	0.0117	53.1447	194.5658	5.0178	0.3939	**0.6013**
	0.2902	0.0084	-0.0310	0.0533	0.0240	1.2312e-4	0.0803	4.8965e-4	**0.5846**
ω DTD Bin-resolved	0.3285	0.0152	0.0074	0.0311	90.7678	226.2984	7.2848	0.6007	**0.6998**
	0.2601	0.0051	-0.0129	0.0111	51.3140	220.1922	4.8888	0.4501	**0.6292**
	0.3054	0.0147	0.0014	0.0620	0.0234	1.2957e-4	0.0798	5.490e-4	**0.6132**
ω DTD Cluster-resolved	0.2973	0.0151	-0.0058	0.0229	92.4096	244.0897	7.3776	0.5053	**0.7289**
	0.2168	0.0060	-0.0172	0.0043	51.9607	264.9770	4.8845	0.5087	**0.6371**
	0.2594	0.0131	-0.0397	0.0995	0.0245	1.4979e-4	0.0855	9.2021e-04	**0.6086**

Table 1: ***Performance metrics for the non-linear RF regression analysis.*** *The mean and variance of R, Q^2, MAE and MSE metrics for the LOO CV calculations as well as R^2 for the final model. The histological parameters are expressed in their original scale. The best predictions for all the histological parameters were achieved with ωDTD cluster-resolved parameter maps as predictors.*

Histological parameter	Comparison	p-value per metric			
		E[R]	E[Q^2]	E[MSE]	E[MAE]
Number of cells	DTI v. ωDTD Per-voxel	**0.0474***	**0.0076****	**0.0021****	**0.0066****
	DTI v. ωDTD Bin-resolved	**0.0067****	**0.0009*****	**0.0009*****	**0.0008*****
	DTI v. ωDTD Cluster-resolved	**0.0335***	**0.0312***	**0.0413***	**0.0306***
Area of Nucleus	DTI v. ωDTD Per-voxel	**0.0032****	**0.1040**	**0.2378**	**0.0934**
	DTI v. ωDTD Bin-resolved	**0.0011****	**0.0019****	**0.0042****	**0.0019****
	DTI v. ωDTD Cluster-resolved	**0.0054***	**0.0054****	**0.0061****	**0.0052****
Circularity of nucleus	DTI v. ωDTD Per-voxel	**0.0011****	0.9483	0.4574	**0.4832**
	DTI v. ωDTD Bin-resolved	**0.0020****	**0.7691**	0.8212	**0.2705**
	DTI v. ωDTD Cluster-resolved	**0.0309***	0.9963	0.1587	**0.8791**

Table 2: ***Two-sided permutation test p-values on random forest LOO CV metrics. Bold*** *indicates instances where ωDTD methods performed better compared to DTI. $^{*}p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$*

input MRI parameters and the target histological parameters. Additionally, the jump of increasing values of R from conventional to ωDTD parameters indicate that the latter MRI parameters are more representative of the histological target parameters. Q^2 is commonly used in context of CV to assess the model's ability to generalize unseen data. These Q^2 values were understandable as considering the highly complex nature of biological data, as the distributions in the target variable values in the new unseen data can be very different compared to the training data. This is especially true in our dataset, as the MCAO lesion

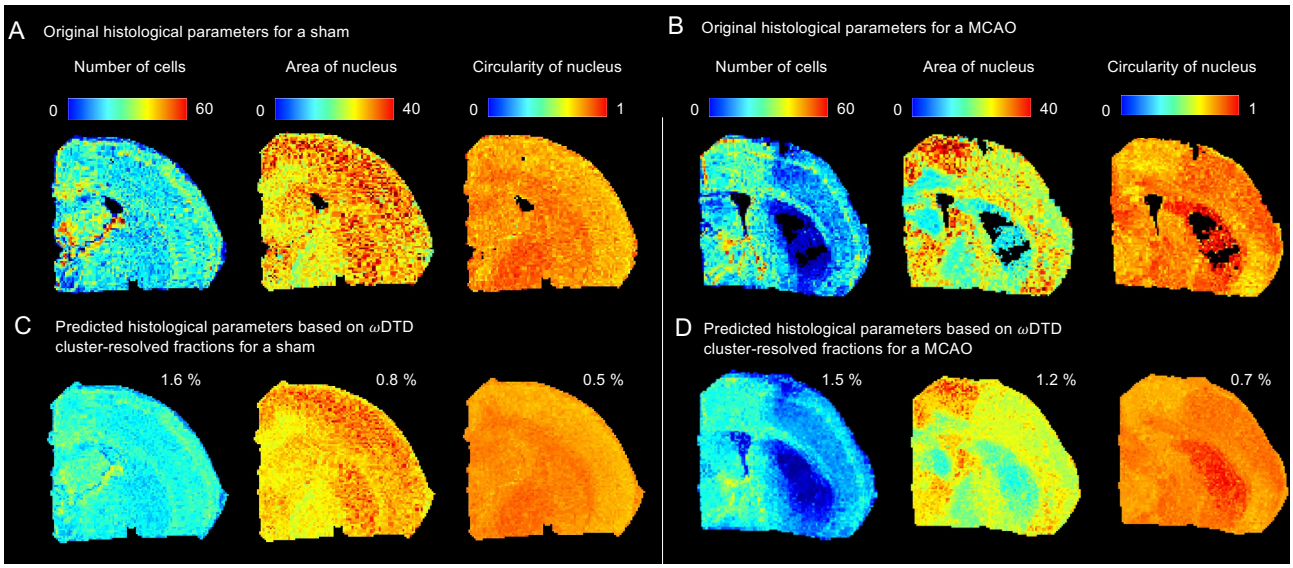

Figure 7: ***Predicted histological parameters of representative sham and MCAO brains based on non-linear RF regression analysis.*** *Comparison of original downsampled and registered histological parameters* ***(A)*** *and* ***(B)*** *with their predicted counterparts* ***(C)*** *and* ***(D)****. The predicted histological parameter maps were derived from training a random RF using ωDTD cluster-resolved mean and rates of change with frequency maps. The percentage indicates average difference between the prediction and original.*

is very heterogeneous in size, location and severity. RFs are not especially prone to overfitting as they include regularization through bagging [66], but the grouped nature of the data (i.e. several datapoints from the same animal) might not be optimal for bagging based regularization as it somewhat violates the independence assumption of data points.

We also calculated the MAE and MSE values for the LOO CV. The spread of both MSE and MAE also suggests some variance in model performance across the folds of the CV. Especially, the variance of MSE indicates that there are large errors in certain CV folds, which are penalized heavily due to the squaring of errors. This implies that the model performs inconsistently across the test folds, which suggests differences between individual animals. The difference in the MSE values associated with the MRI parameter sets shows improved and more consistent performance across the CV folds with less large errors when moving from conventional to ωDTD parameter sets. This phenomenon along with the increase of R^2 values is reasonable, as the performance of a RF model is tied to the information content, but also the amount, of data fed into the algorithm. The increased amount of data helps reduce variance while providing the RF more opportunities to identify robust and generalizable patterns. More descriptive data, i.e. data more strongly related to the target variable, on the other hand allow the decision trees to make more confident splits leading to improved predictive power of the ensemble.

The permutation test results comparing conventional DTI against the three ωDTD methods provide robust statistical evidence that multidimensional diffusion MRI parameters yield systematically higher predictive performance across LOO CV folds. For number of cells, all ωDTD variants significantly outperformed DTI across all four metrics, confirming the markedly stronger relationship between MRI features and neuronal density quantification. Area of nucleus showed a similar pattern with 9 out of 12 significant comparisons favoring ωDTD, whereas circularity exhibited the weakest and most inconsistent MRI-histology coupling with few significant differences. The predominance of comparisons indicating superior ωDTD for R and Q^2 suggests that ωDTD's microstructural sensitivity enhances correlation-based generalization while giving superior absolute prediction accuracy. This inter-subject permutation testing framework validates that the observed performance gains from conventional to ωDTD parameter sets reflect true methodological improvements rather than random variation.

The dominance of $\mathrm{E}[D_{iso}]$ clusters in the global SHAP rankings underscores the model's reliance on isotropic diffusivity as a primary biomarker of stroke-related tissue damage across the whole brain. The leading role of $\mathrm{E}[D_{iso}]$ of f_{clust1} could be explained by this cluster's substantial representation ($\sim$ 20% of voxels per animal) along with its excellent lesion contrast. This suggests the model captures both cytotoxic edema in the lesion core and possibly the decrease of organization. $\mathrm{E}[D_{iso}]$ of f_{clust3} and f_{clust4} provide complementary information on the overall data: f_{clust3}'s oblate, frequency-sensitive profile might reflect disrupted neuronal processes, while

f_{clust4} delineates the primary pathological compartment i.e. edema. Across the entire dataset, $\mathrm{E}[D_{iso}]$ along with these three clusters clearly emerge as the primary drivers of model predictions.

There are some limitations in this study that should be considered. First, for future studies, the sample size should be increased to extract more reliable statistical information. Secondly, the MCAO model produces a quite variable ischemic lesion, both in size and location. This heterogeneity also brought challenges to conventional statistical analyses. Third, the acute 24-h timepoint is a quite short period of time for any secondary changes in distal tissue to be detectable. This timepoint also provides a very limited window for the MRI to show the tissue response to the neuroglobin treatment, particularly given that the treatment's goal at the 24-hour timepoint is primarily neuroprotection [22]. Additionally, infarct maturation beyond 24 h (e.g., toward 72 h or 7 days) could alter ωDTD, reducing insight into temporal robustness.

As a fourth, we have only compared conventional DTI to the ωDTD parameters, without including other advanced post-processing methods, such as soma and neurite density imaging (SANDI) or the standard model [67, 68, 69]. Single diffusion encoding (SDE) data acquired from the same tissue samples as advanced diffusion measurements would enable more rigorous comparisons. This represents an important direction for future work and will be addressed in subsequent studies. Fifth, histology-to-MRI registration was performed using rigid similarity transforms, which may not fully account for local tissue deformations caused by ischemia, fixation, and sectioning. More advanced non-linear or blockface-guided registration strategies with quantitative target registration error assessment could improve the accuracy of voxel-wise MRI-histology correspondence. Sixth, Nissl staining was the only histology method included in our study, and other stains could add more information about the complex changes occurring in the tissue after ischemia, e.g. myelin damage or inflammation. Finally, we used the automated cell detection with default settings provided by QuPath to extract quantitative data from the histology. These algorithms require further development to increase their accuracy and extract more valuable parameters of tissue damage.

Future steps will be to expand the analysis to fully 3D, for both the MRI and histology. Our frequency-dependent MRI data was 2D due to technical constraints of our sequence at the time, which has since been formatted for 3D acquisition for *ex vivo* and *in vivo*. The latest MD-MRI *in vivo* protocols exist for both preclinical and clinical settings, achieving whole-brain coverage in 20-40 minutes depending on parameter selection, ω-ranges, and inclusion of $\mathrm{T}_1/\mathrm{T}_2$ variation. These example protocols for $\mathbf{D}-T_1-T_2$ and $\mathbf{D}(\omega)-T_1-T_2$ presented in [17, 70, 37], establish that streamlined MD-MRI approaches, such as ωDTD, are achievable despite key constraints of scan time and gradient demands. A feasible 40-minute $\mathbf{D}(\omega)-T_1-T_2$ acquisition of the human brain employs varying b-values (0–3×10^9 s/m^{-2}), normalized anisotropy b_Δ (–0.5, 0, and 1), varying repetition time τ_R (620–7000 ms), and echo time τ_E (40–150 ms), orientation (θ,ϕ) and centroid frequency $\omega_{\text{cent}}/2\pi$ within 10–90% percentile (4–9 Hz). We are additionally in the process of analysing recently acquired longitudinal *in vivo* preclinical $\mathbf{D}(\omega)-T_1-T_2$ data to investigate the temporal effects of both the neuroglobin treatment and changes in $\mathbf{D}(\omega)-T_1-T_2$ parameters and -derived clusters for a future publication.

We will also include other stains to help better interpret the ωDTD parameters in terms of histological changes and to understand the relationship between methodologies. With increased histological data, it would be possible to extract more quantitative parameters for our analyses. Additionally, for the analysis, hyperparameter tuning and grid searches for determining the best possible combinations for the RF regression model could also improve the performance [71]. This could improve the robustness of our model and ability to generalize. Incorporating deep learning elements and algorithms could also open new avenues and perspectives in predicting tissue microstructural changes based on imaging outcomes during pathological conditions.

CONCLUSION

In conclusion, our findings show that ωDTD better predicts the changes in quantitative tissue microstructure parameters than conventional DTI after ischemic stroke. And, although our models could not perfectly explain the data, this study indicates that ωDTD can open new avenues in the evaluation of ischemic stroke by offering higher sensitivity to the changes associated with ischemic injury, even in its earliest stages. By capturing subtle cellular level alterations in tissue viability, composition, and microstructure, ωDTD can enhance our understanding of ischemic stroke progression and improve clinical decision-making along with facilitating better monitoring of treatment responses.

DATA AND CODE AVAILABILITY

The md-dmri MATLAB toolbox for preprocessing and Monte-Carlo data inversion, along with codes for Gaussian mixture model clustering are freely available at https://github.com/UEF-Multiscale-Imaging. The data used in this study is available upon request.

FUNDING

This study has been funded by the Molecular Medicine Doctoral Programme from University of Finland (UEF), Kuopio, Finland, Vilho, Yrjö, and Kalle Väisälä Foundation of the Finnish Academy of Science and Letters, Finnish Cultural Foundation, Flagship of Advanced Mathematics for Sensing Imaging and Modeling from the Research Council of Finland (#358944), project funding from the Research Council of Finland (#323385 and #361370), the Spanish Ministry for Science, Innovation and Universities (the Spanish Ministry for Science, Innovation and Universities (Grant BFU2016-80316-R funded by MICIU/AEI/10.13039/501100011033 and by "ERDF A way of making Europe" and the Spanish Ministry for Science and Innovation (Grant PID2021-123791OB-100, funded by MICIU/AEI/10.13039/501100011033 and by "ERDF/EU"), the Swedish Foundation for Strategic Research (Stiftelsen för Strategisk Forskning; grant no. ITM17-0267), and the Swedish Research Council (Vetenskapsrådet; grant nos. 2018-03697, 2022-04422_VR, 21073).

DECLARATION OF COMPETING INTERESTS

The authors declare that this research was conducted without any commercial or financial relationship that could be considered as a potential conflicts of interest.

ACKNOWGELEDGEMENTS

The computational analyses were performed with the servers of the Bioinformatics Center, University of Eastern Finland, Finland, and the supercomputers at CSC — IT Center for Science, Finland. This study was carried out with the support of the Kuopio Biomedical Imaging Unit, University of Eastern Finland, Kuopio, Finland (part of the Finnish Biomedical Imaging Node, Euro-BioImaging), and the technical and human support provided by the *Centro de Instrumentación Científico-Técnica* (CICT) and by the *Centro de Producción y Experimental Animal* (CPEA) — *Servicios Centrales de Apoyo a la Investigación* (SCAI) — Universidad de *Jaén (UJA, MICINN, Junta de Andalucía, FEDER*).

References

[1] World Stroke Organization (WSO). 2022. URL: https://www.world-stroke.org/assets/downloads/WSO_Global_Stroke_Fact_Sheet.pdf.

[2] S.E. Lakhan, A. Kirchgessner, and M. Hofer. "Inflammatory mechanisms in ischemic stroke: therapeutic approaches". In: *Journal of Translational Medicine* 7.97 (2009). DOI: 10.1186/1479-5876-7-97.

[3] C. Qin, S. Yang, Y-H. Chu, H. Zhang, X-W. Pang, L. Chen, L-Q. Zhou, M. Chen, D-S. Tian, and W. Wang. "Signaling pathways involved in ischemic stroke: molecular mechanisms and therapeutic interventions". In: *Signal Transduction and Targeted Therapy* 7.215 (2022). DOI: 10.1038/s41392-022-01064-1.

[4] B.J. Kim, H.G. Kang, H-J. Kim, S-H. Ahn, N.Y. Kim, S. Warach, and D-W. Kang. "Magnetic Resonance Imaging in Acute Ischemic Stroke Treatment". In: *Journal of Stroke* 16.3 (2014). DOI: 10.5853/jos.2014.16.3.131.

[5] E.V.R. DiBella, A. Sharma, L. Richards, V. Prabhakaran, J.J. Majersik, and S.K. HashemizadehKolowri. "Beyond diffusion tensor MRI methods for improved characterisation of brain after ischemic stroke: Review". In: *American Journal of Neuroradiology* 43.5 (2025). DOI: 10.3174/ajnr.A7414.

[6] L. Emsell, W. Van Hecke, and J. Tournier. *Introduction to Diffusion Tensor Imaging*. Diffusion Tensor Imaging: A Practical Handbook. Springer, 2016.

[7] P.A. Barber. "Magnetic resonance imaging of ischemia viability thresholds and the neurovascular unit". In: *Magnetic Resonance Sensors* 13.6 (2013). DOI: 10.3390/s130606981.

[8] J.J. Nukovic, V. Opancina, E. Ciceri, M. Muto, N. Zdravkovic, A. Altin, P. Altaysoy, R. Kastelic, D.M. Velazquez Mendivil, J.A. Nukovic, N.V. Markovic, M. Opancina, T. Prodanovic, M. Nukovic, J. Kostic, and N. Prodanovic. "Neuroimaging modalities used for ischemic stroke diagnosis and monitoring". In: *Medicina* 59.11 (2023). DOI: 10.3390/medicina59111908.

[9] D. Topgaard. "Diffusion tensor distribution imaging". In: *NMR in Biomedicine* 32.5 (2019). DOI: 10.1002/nbm.4066.

[10] M. Yon, J.P. de Almeida Martins, Q. Bao, M.D. Budde, L. Frydman, and D. Topgaard. "Diffusion tensor distribution imaging of an in vivo mouse brain at ultrahigh magnetic field by spatiotemporal encoding". In: *NMR in Biomedicine* 33.11 (2020). DOI: 10.1002/nbm.4355.

[11] M. Yon, O. Narvaez, J. Martin, H. Jiang, D. Bernin, E. Forssell-Aronsson, F. Laun, A. Sierra, and D. Topgaard. "Frequency-dependence in multidimensional diffusion-relaxation correlation MRI of the brain: overfitting or meaningful parameter?" In: *bioRxiv* (2024). DOI: 10.1101/2024.04.29.591586.

[12] O. Narvaez, L. Svenningsson, M. Yon, A. Sierra, and D. Topgaard. "Massively Multidimensional Diffusion-Relaxation Correlation MRI". In: *Frontiers in Physics* 9 (2022). DOI: 10.3389/fphy.2021.793966.

[13] M. Aggarwal, M.V. Jones, P.A. Calabresi, S. Mori, and J. Zhang. "Probing mouse brain microstructure using oscillating gradient diffusion MRI". In: *Magnetic Resonance in Medicine* 67.1 (2012). DOI: 10.1002/mrm.22981.

[14] H. Lundell, M. Nilsson, T.B. Dyrby, G.J.M. Parker, P.L. Hubbard Cristinacce, F-L. Zhou, D. Topgaard, and S. Lasič. "Multidimensional diffusion MRI with spectrally modulated gradients reveals unprecedented microstructural detail". In: *Scientific Reports* 9.9026 (2019). DOI: 10.1038/s41598-019-45235-7.

[15] D. Topgaard. "Multidimensional diffusion MRI". In: *Journal of Magnetic Resonance* 275 (2017). DOI: 10.1016/j.jmr.2016.12.007.

[16] J.P. de Almeida Martins, C.M.W. Tax, F. Szczepankiewicz, D.K. Jones, C-F. Westin, and D. Topgaard. "Transferring principles of solid-state and Laplace NMR to the field of in vivo brain MRI". In: *Magnetic Resonance* 1.1 (2020). DOI: 10.5194/mr-1-27-2020.

[17] O. Narvaez, M. Yon, R.A. Salo, J. Kyyriäinen, M. Estela, E. Paasonen, D. Topgaard, and A. Sierra. "Data-driven classification of tissue water populations by massively multidimensional diffusion-relaxation correlation MRI". In: *bioRxiv* (2025). DOI: 10.1101/2025.03.28.645938.

[18] S. Blanco, E. Martínez-Lara, E. Siles, and M.Á. Peinado. "New Strategies for Stroke Therapy: Nanoencapsulated Neuroglobin". In: *Pharmaceutics* 14.8 (2022). DOI: 10.3390/pharmaceutics14081737.

[19] S. Blanco, S. Peralta, M.E. Morales, E. Martínez-Lara, J.R. Pedrajas, H. Castán, M.Á. Peinado, and M.A. Ruiz. "Hyaluronate Nanoparticles as a Delivery System to Carry Neuroglobin to the Brain after Stroke". In: *Pharmaceutics* 12.1 (2020). DOI: 10.3390/pharmaceutics12010040.

[20] S. Peralta, S. Blanco, R. Hernández, H. Castán, E. Siles, E. Martínez-Lara, M.E. Morales, M.Á. Peinado, and M.A. Ruiz. "Synthesis and characterization of different sodium hyaluronate nanoparticles to transport large neurotherapeutic molecules through blood brain barrier after stroke". In: *European Polymer Journal* 112 (2019). DOI: 10.1016/j.eurpolymj.2019.01.030.

[21] N. Burmester and T. Hankeln. "What is the function of neuroglobin?" In: *Journal of Experimental Biology* 212.10 (2009). DOI: 10.1242/jeb.000729.

[22] M.Á. Peinado, D. Ovelleiro, M.L. Del Moral, R. Hernández, E. Martínez-Lara, E. Siles, J.R. Pedrajas, M.L. García-Martín, C. Caro, S. Peralta, M.E. Morales, M.A. Ruiz, and S. Blanco. "Biological Implications of a Stroke Therapy Based in Neuroglobin Hyaluronate Nanoparticles. Neuroprotective Role and Molecular Bases". In: *International Journal of Molecular Sciences* 23.1 (2022). DOI: 10.3390/ijms23010247.

[23] F. Liu, W. Yuan, Z. Li, Y. Cui, Y. Li, and Y. Wang. "TTC, Fluoro-Jade B and NeuN staining confirm evolving phases of infarction induced by middle cerebral artery occlusion". In: *Journal of Neuroscience Methods* 179.1 (2009), pp. 1–8. DOI: 10.1016/j.jneumeth.2008.12.019.

[24] M. D. Linnik, J. A. Miller, J. R. Sprinkle-Carnum, J. R. Faraday, and S. J. Hardwick. "Apoptotic DNA fragmentation in the rat cerebral cortex induced by permanent middle cerebral artery occlusion". In: *Molecular Brain Research* 32.1 (1995), pp. 116–124. DOI: 10.1016/0169-328X(95)00062-F.

[25] J. Chen, M. Nagayama, K. Jin, R. A. Stetler, L. Xie, J. Zhu, S. H. Graham, and R. P. Simon. "Induction of caspase-3-like protease is responsible for apoptosis in rat after transient focal cerebral ischemia". In: *Journal of Neurochemistry* 69.1 (1997), pp. 232–245. DOI: 10.1046/j.1471-4159.1997.69010232.x.

[26] R.L.C. Barrett, D. Cash, C. Simmons, E. Kim, T.C. Wood, R. Stones, A.C. Vernon, M. Catani, and F. Dell'Acqua. "Tissue optimization strategies for high-quality ex vivo diffusion imaging". In: *NMR in Biomedicine* 36.3 (2023). DOI: 10.1002/nbm.4866.

[27] Bruker BioSpin. *ParaVision: Advanced Imaging Software for Preclinical MRI*. Version X.X, Bruker BioSpin, Germany. Available: https://www.bruker.com. 2024.

[28] J. Veraart, E. Fieremans, and D.S. Novikov. "Diffusion MRI noise mapping using random matrix theory". In: *Magnetic Resonance in Medicine* 76.5 (2016). DOI: 10.1002/mrm.26059.

[29] J. Veraart, B. Dhital, V.G. Kiselev, and M. Reisert. "Gibbs-ringing artifact removal based on local subvoxel-shifts". In: *Magnetic Resonance in Medicine* 76.5 (2016). DOI: 10.1002/mrm.26054.

[30] O. Narvaez, M. Yon, H. Jiang, D. Bernin, E. Forssell-Aronsson, A. Sierra, and D. Topgaard. "Nonparametric distributions of tensor-valued Lorentzian diffusion spectra for model-free data inversion in multidimensional diffusion MRI". In: *Journal of Chemical Physics* 161.8 (2024). DOI: 10.1063/5.0213252.

[31] M. Prange and Y-Q. Song. "Quantifying uncertainty in NMR T_2 spectra using Monte Carlo inversion". In: *Journal of Magnetic resonance* 196.1 (2009). DOI: 10.1016/j.jmr.2008.10.008.

[32] M. Nilsson, F. Szczepankiewicz, and A. Lampinen B. ans Ahlgren. "An Open-Source Framework for Analysis of Multidimensional Diffusion MRI Data Implemented in MATLAB". In: *Proceedings International Society Of Magnetic Resonance In Medicine, 2018 Jun 16–21; Paris, France* (2018).

[33] M. Aggarwal. *Chapter 4. Restricted Diffusion and Spectral Content of the Gradient Waveforms*. New Developments in NMR. Royal Society of Chemistry, 2020. DOI: 10.1039/9781788019910-00103.

[34] A. Arbabi, J. Kai, A.R. Khan, and C.A. Baron. "Diffusion dispersion imaging: Mapping oscillating gradient spin-echo frequency dependence in the human brain". In: *Magnetic Resonance in Medicine* 83.6 (2020). DOI: 10.1002/mrm.28083.

[35] M. Yon, O. Narvaez, D. Topgaard, and A. Sierra. "In vivo rat brain mapping of multiple gray matter water populations using nonparametric D(ω)-R1-R2 distributions MRI". In: *NMR in Biomedicine* 38.1 (2025). DOI: 10.1002/nbm.5286.

[36] E. Manninen, S. Bao, B.A. Landman, Y. Yang, D. Topgaard, and D. Benjamini. "Variability of multidimensional diffusion–relaxation MRI estimates in the human brain". In: *Imaging Neuroscience* 2 (2024). DOI: 0.1162/imag_a_00387.

[37] J.T.E. Johnson, M.O. Irfanoglu, E. Manninen, T.J. Ross, Y. Yang, F.B. Laun, J. Martin, D. Topgaard, and D. Benjamini. "In vivo disentanglement of diffusion frequency-dependence, tensor shape, and relaxation using multidimensional MRI". In: *Human Brain Mapping* 45.7 (2024). DOI: 10.1002/hbm.26697.

[38] C-F. Westin, H. Knutsson, O. Pasternak, F. Szczepankiewicz, E. Özarslan, D. van Westen, C. Mattisson, M. Bogren, L.J. O'Donnell, M. Kubicki, D. Topgaard, and M. Nilsson. "Q-space trajectory imaging for multidimensional diffusion MRI of the human brain". In: *Neuroimage* 135 (2016). DOI: 10.1016/j.neuroimage.2016.02.039.

[39] S. Eriksson, S. Lasič, M. Nilsson, C-F. Westin, and D. Topgaard. "NMR diffusion-encoding with axial symmetry and variable anisotropy: Distinguishing between prolate and oblate microscopic diffusion tensors with unknown orientation distribution". In: *Journal of Chemical Physics* 142.10 (2015). DOI: 10.1063/1.4913502.

[40] J.P. de Almeida Martins and D. Topgaard. "Multidimensional correlation of nuclear relaxation rates and diffusion tensors for model-freeinvestigations of heterogeneous anisotropic porous materials". In: *Scientific reports* 8.2488 (2018). DOI: 10.1038/s41598-018-19826-9.

[41] L. Breiman. "Random forests". In: *Machine learning* 45 (2001).

[42] P. Bankhead, M.B. Loughrey, J.A. Fernández, Y. Dombrowski, D.G. McArt, P.D. Dunne, S. McQuaid, R.T. Gray, L.J. Murray, H.G. Coleman, J.A. James, M. Salto-Tellez, and P.W. Hamilton. "QuPath: Open source software for digital pathology image analysis". In: *Scientific Reports* 7 (2017).

[43] D. Mattes, D.R. Haynor, H. Vesselle, T. Lewellen, and W. Eubank. "Non-rigid multimodality image registration". In: *Medical Imaging 2001: Image Processing*. SPIE, 2001.

[44] W. Wienholt. *Entwurf neuronaler Netze: ein datenbasiertes Verfahren zur Modellierung nichtlinearer Systeme mit Hilfe evolutionarer Algorithmen und unscharfer Logik*. Harri Deutsch, 1996.

[45] M. Styner, C. Brechbuehler, G. Székely, and G. Gerig. "Parametric estimate of intensity inhomogeneities applied to MRI". In: *IEEE Transactions on Medical Imaging* 19.3 (2000). DOI: 10.1109/42.845174.

[46] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, 2009.

[47] I. San Martín Molina, R.A. Salo, O. Gröhn, J. Tohka, and A. Sierra. "Histopathological modeling of status epilepticus-induced brain damage based on in vivo diffusion tensor imaging in rats". In: *Frontiers in Neuroscience* 16 (2022). DOI: 10.3389/fnins.2022.944432.

[48] H. Martens and M. Martens. *Multivariate Analysis of Quality*. Elsevier, 2012.

[49] G. James, D. Witten, T. Hastie, and R. Tibshirani. *An Introduction to Statistical Learning: With Applications in R*. Springer, 2013.

[50] D. Chicco, M.J. Warrens, and G. Jurman. "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation". In: *PeerJ - Computer Science* 7 (2021). DOI: 10.7717/peerj-cs.623.

[51] Scott M. Lundberg, Gabriel Erion, Hugh Chen, Alex DeGrave, Jordan M. Prutkin, Bala Nair, Ronit Katz, Jonathan Himmelfarb, Nisha Bansal, and Su-In Lee. *Explainable AI for Trees: From Local Explanations to Global Understanding*. 2019. DOI: 10.48550/arXiv.1905.04610. arXiv: 1905.04610 [cs.LG].

[52] C.H. Sotak. "The Role of diffusion tensor imaging in the evaluation of ischemic brain injury – a review". In: *NMR in biomedicine* 15 (2002). DOI: 10.1002/nbm.786..

[53] M. Pitkonen, U. Abo-Ramadan, I. Marinkovic, E. Pedrono, K.M. Hasan, D. Strbian, A. Durukan, and T. Tatlisumak. "Long-term evolution of diffusion tensor indicates after temporary experimental ischemic stroke in rats". In: *Brain Research* 1445 (2012). DOI: 10.1016/j.brainres.2012.01.043.

[54] M. Zhou, R. Stobbe, F. Szczepankiewicz, M. Budde, B. Buck, M. Kate, M. Lloret, P. Fairall, K. Butcher, A. Shuaib, D. Emery, M. Nilsson, C-F. Westin, and C. Beaulieu. "Tensor-valued diffusion MRI of human acute stroke". In: *Magnetic resonance in medicine* 91.5 (2024). DOI: 10.1002/mrm.29975..

[55] F. Christidi, D. Tsiptsios, A. Fotiadou, S. Kitmeridou, S. Karatzetzou, K. Tsamakis, A. Sousanidou, E.A. Psatha, E. Karavasilis, I. Seimenis, C. Kokkotis, N. Aggelousis, and K. Vadikolias. "Diffusion tensor imaging as a prognostic tool for recovery in acute and hyperacute stroke". In: *Neuron intentional* 14.4 (2022). DOI: 10.3390/neurolint14040069.

[56] Y. Liao, S. Coelho, J. Chen, B. Ades-Aron, M. Pang, V. Stepanov, R. Osorio, T. Shepherd, Y.W. Lui, D.S. Novikov, and E. Fieremans. "Mapping tissue microstructure of brain white matter in vivo in health and disease using diffusion MRI". In: *Imaging Neuroscience* 2 (2024). DOI: 10.1162/imag_a_00102.

[57] J. Veraart, D.H.J. Poot, W. Van Hecke, I. Blockx, A. Van der Linden, M. Verhoye, and J. Sijbers. "More Accurate Estimation of Diffusion Tensor Parameters Using Diffusion Kurtosis Imaging". In: *Magnetic Resonance in Medicine* 65.1 (2011).

[58] A.M. Auriat, M.R. Borich, N.J. Snow, K.P. Wadden, and L.A. Boyd. "Comparing a diffusion tensor and non-tensor approach to white matter fiber tractography in chronic stroke". In: *NeuroImage: Clinical* 7 (2015).

[59] Z. Wang, S. Zhang, C. Liu, Y. Yao, J. Shi, J. Zhang, Y. Qin, and W. Zhu. "A study of neurite orientation dispersion and density imaging in ischemic stroke". In: *Magnetic Resonance Imaging* 57 (2019).

[60] A. Reymbaut, Y. Zheng, S. Li, W. Sun, H. Xu, I. Daimiel Naranjo, S. Thakur, K. Pinker-Domenig, S. Rajan, V.K. Vanugopal, V. Mahajan, H. Mahajan, J. Critchley, G. Durighel, M. Sughrue, K. Bryskhe, and D. Topgaard. *Clinical Research with Advanced Diffusion Encoding Methods in MRI in Advanced Diffusion Encoding Methods in MRI*. The Royal Society of Chemistry, 2020.

[61] L.S. Hu, F. D'Angelo, T.M. Weiskittel, F.P. Caruso, S.P. Fortin Ensign, M.R. Blomquist, M.J. Flick, L. Wang, C.P. Sereduk, K. Meng-Lin, G. De Leon, A. Nespodzany, J.C. Urcuyo, A.C. Gonzales, L. Curtin, E.M. Lewis, K.W. Singleton, T. Dondlinger, A. Anil, N.B. Semmineh, T. Noviello, R.A. Patel, P. Wang, J. Wang, J.M. Eschbacher, A. Hawkins-Daarud, P.R. Jackson, I.S. Grunfeld, C. Elrod, G.L. Mazza, S.C. McGee, L. Paulson, K. Clark-Swanson, Y. Lassiter-Morris, K.A. Smith, P. Nakaji, B.R. Bendok, R.S. Zimmerman, C. Krishna, D.P. Patra, N.P. Patel, M. Lyons, M. Neal, K. Donev, M.M. Mrugala, A.B. Porter, S.C. Beeman, T.R. Jensen, K.M. Schmainda, Y. Zhou, L.C. Baxter, C.L. Plaisier, J. Li, H. Li, A. Lasorella, C.C. Quarles, K.R. Swanson, M. Ceccarelli, A. Iavarone, and N.L. Tran. "Integrated molecular and multiparametric MRI mapping of high-grade glioma identifies regional biologic signatures". In: *Nature Communications* 14.6066 (2023). DOI: 10.1038/s41467-023-41559-1.

[62] J.S. Park, E. Manninen, Y. Yang, and D. Benjamini. "Informed Dictionary-Guided Monte Carlo Inversion for Robust and Reproducible Multidimensional MRI". In: *Magn Reson Med* (2025). DOI: 10.1002/mrm.70228.

[63] J. Xu. "Probing neural tissues at small scales: Recent progress of oscillating gradient spin echo (OGSE) neuroimaging in humans". In: *Journal of Neuroscientific Methods* 349 (2023). DOI: 10.1016/j.jneumeth.2020.109024.

[64] C.A. Baron and C. Beaulieu. "Oscillating gradient spin-echo (OGSE) diffusion tensor imaging of the human brain". In: *Magnetic Resonance in Medicine* 72.3 (2014). DOI: 10.1002/mrm.24987.

[65] C.A. Baron, M.P. Kate, L.C. Gioia, K. Butcher, D. Emery, M.D. Budde, and C. Beaulieu. "Oscillating Gradient Spin-Echo (OGSE) DTI Yields Mechanistic Insights in Human Stroke". In: 2013. URL: https://api.semanticscholar.org/CorpusID:46648512.

[66] A. Geron. *Hands-on Machine Learning with Scikit-Learn, Keras and TensorFlow*. Springer, 2009.

[67] M. Palombo, A. Ianus, M. Guerreri, D. Nunes, D.C. Alexander, N. Shemesh, and H. Zhang. "SANDI: A compartment-based model for non-invasive appartent soman and neurite imaging by diffusion MRI". In: *NeuroImage* 215 (2020). DOI: 10.1016/j.neuroimage.2020.116835.

[68] D.S. Novikov, E. Fieremans, S.N. Jespersen, and V.G. Kiselev. "Quantifying brain microstructure with diffusion MRI: Theory and parameter estimation". In: *NMR in Biomedicine* 32.4 (2019). DOI: 10.1002/nbm.3998.

[69] S. Coelho, S.H. Baete, G. Lemberskiy, B. Ades-Aron, G. Barrol, J. Veraart, D.S. Novikov, and E. Fieremans. "Reproducibility of the Standard Model of diffusion in white matter on clinical MRI systems". In: *NeuroImage* 257 (2020). DOI: 10.1016/j.neuroimage.2022.119290.

[70] J. Martin, A. Reymbaut, M. Schmidt, A. Doerfler, M. Uder, F.B. Laun, and D. Topgaard. "Nonparametric D-R1-R2 distribution MRI of the living human brain". In: *Neuroimage* (2021). DOI: 10.1016/j.neuroimage.2021.118753..

[71] A. Geron. *Hands-on Machine Learning with Scikit-Learn, Keras and TensorFlow*. Springer, 2009.

SUPPLEMENTARY MATERIAL

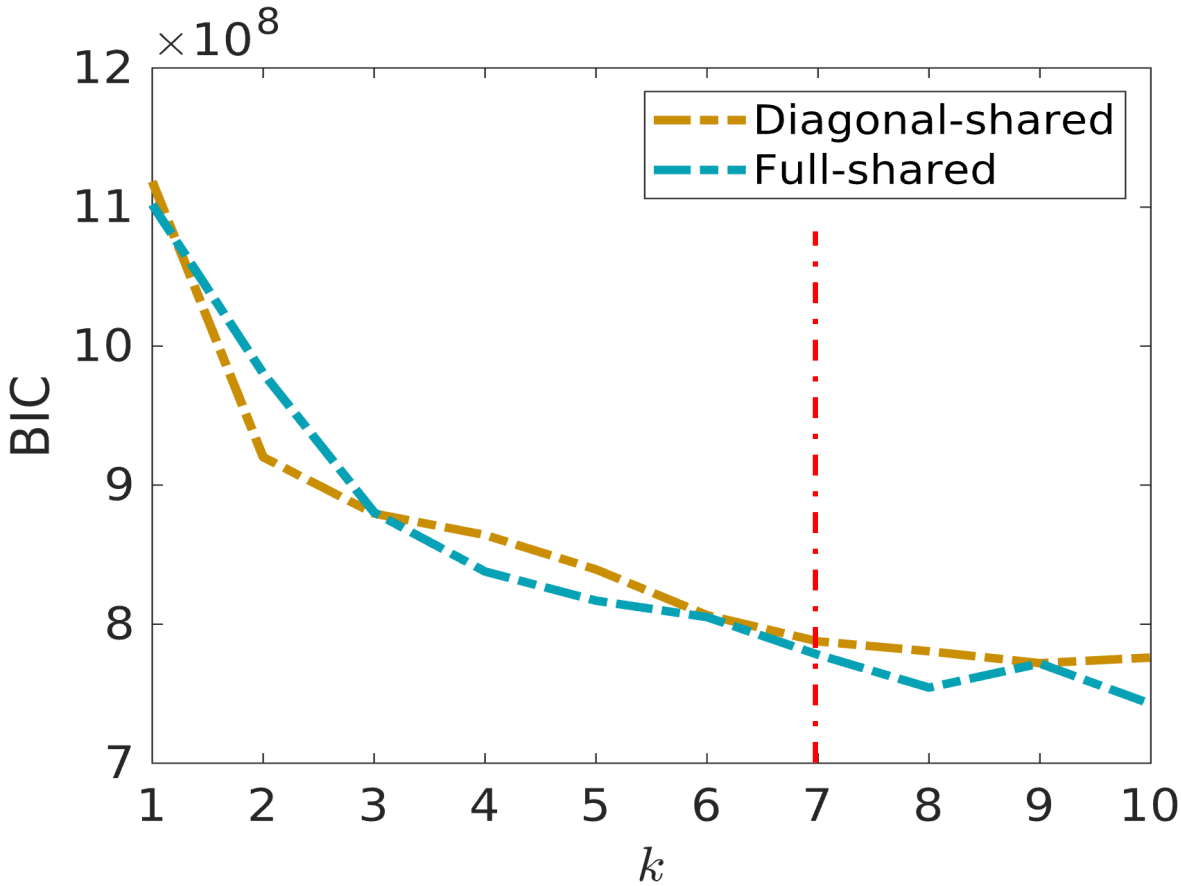

Figure S8: ***Bayesian information criteria (BIC) in aiding the selection of optimal value for k in unsupervised clustering of D(ω) distributions.*** *The red dotted line indicates the optimum k, chosen as the point of stabilization. Diagonal and full shared refer to the allowed shapes of the clusters.*

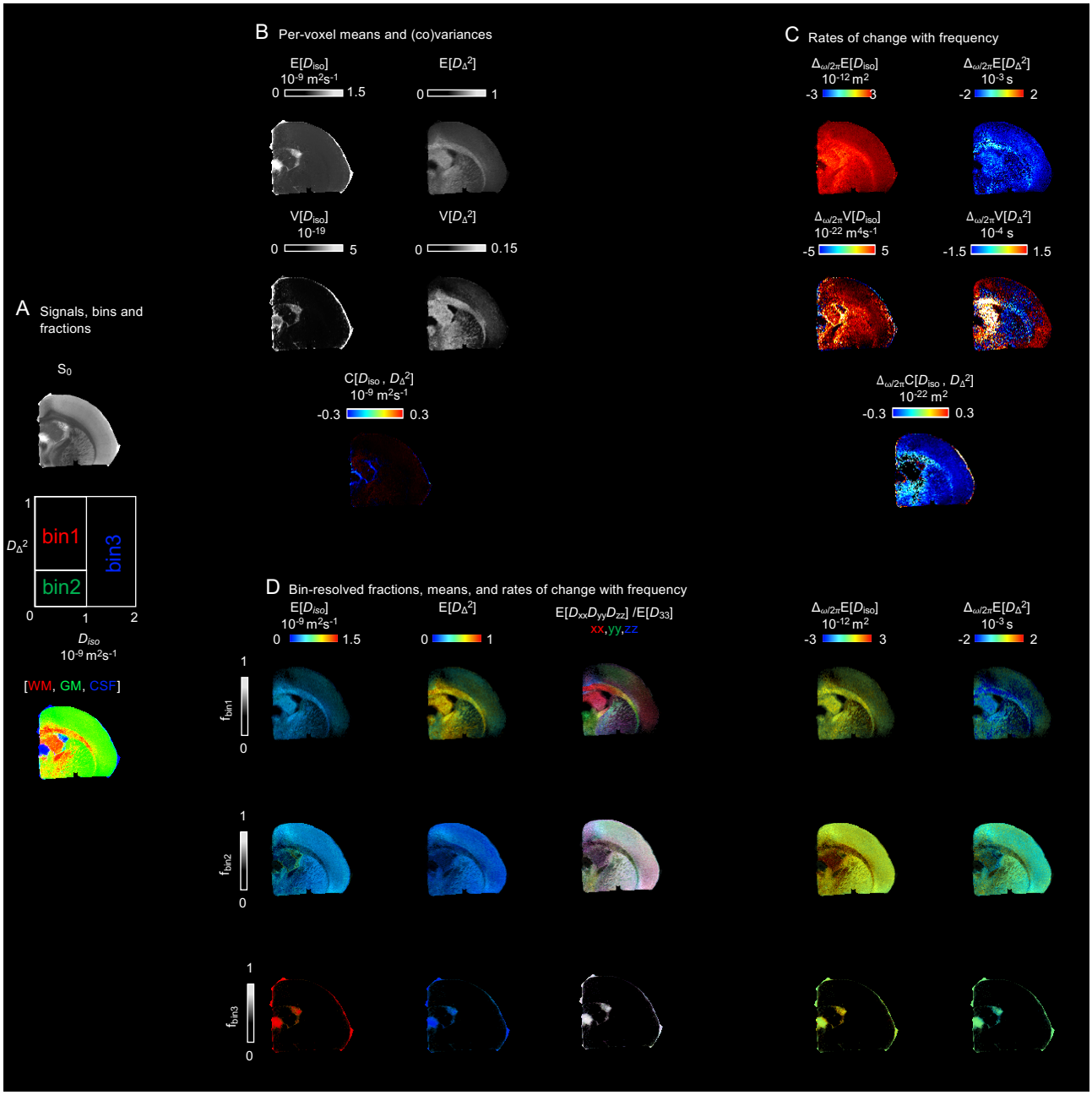

Figure S9: ***Parameter maps derived from the per-voxel D(ω) distributions for a representative sham brain.*** ***(A)*** *S_0 map in grayscale, the diagram depicting the division of the 2D D_{iso} - D_δ^2 projection into three bins (bin_1, bin_2, and bin_3) and the resulting signal fractions (f_{bin1}, f_{bin2}, and f_{bin3}) coded into RGB color. (B) The per-voxel means E[X], variances V[X], and covariances C[X] at a selected encoding frequency at $\omega/2\pi$ = 34.4 Hz.* ***(C)*** *Parameter maps of the rate of change with frequency $\Delta_{\omega/2\pi}$ of the per-voxel means, variances, and covariances of D_{iso} and D_δ^2 calculated with $\omega/2\pi$ = 34.4-115 Hz. (D) The bin resolved maps of E[X] and $\Delta_{\omega/2\pi}E[X]$. The brightness of the bin-resolved maps reflects signal fraction, while the color indicates parameter values.*

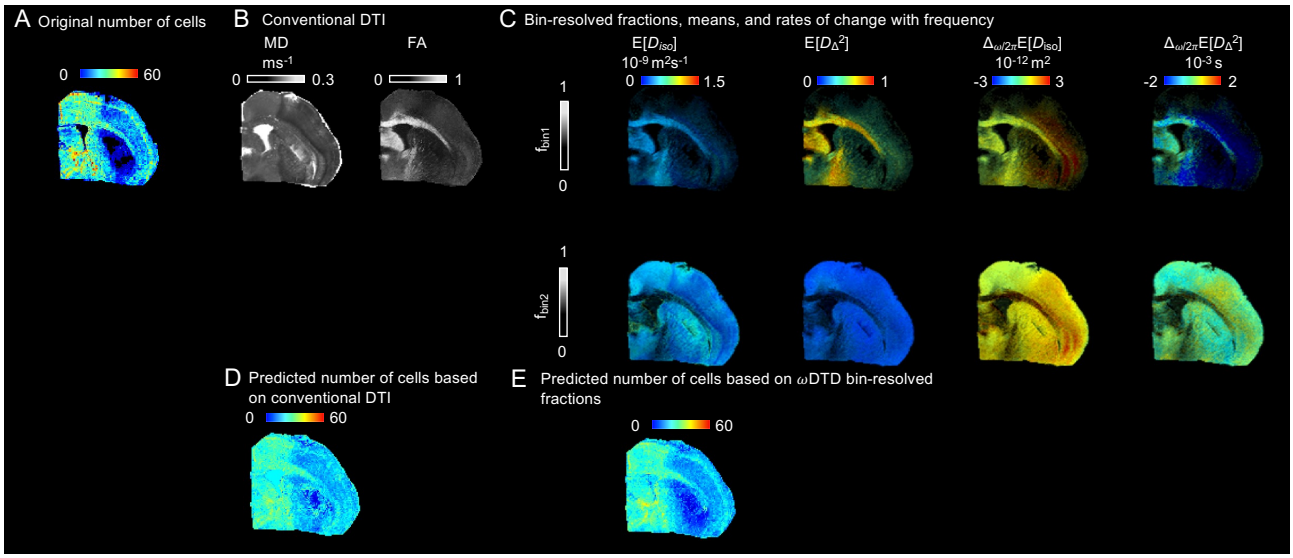

Figure S10: ***Example of the difference between the non-linear regression predictions for number of cells per voxel histological parameter based on two different MRI parameter sets.*** *The two MRI parameter set are shown in* ***(B)****: Conventional MRI and* ***(C)****: ωDTD bin-resolved mean and rates of change per frequency maps (****equations*** (4),(6),(7)***, and*** (8)*). The predicted number of cells per voxel based on the conventional MRI parameter set* ***(D)*** *demonstrates near identical lesion details as can be seen in the MD map in* ***(B)****, highlighting a scenario where the regression has clearly failed to capture the nuances of the histological tissue structure. This prediction also lacks the distinct details of the lesion as seen in the original map in* ***(A)****, in addition to underestimating the values. In comparison, the predicted number of cells per voxel based on the ωDTD bin-resolved fractions* ***(E)****, resembles more closely the original map in* ***(A)*** *both in lesion details and values, while not directly copying information from any singular MRI parameter.*

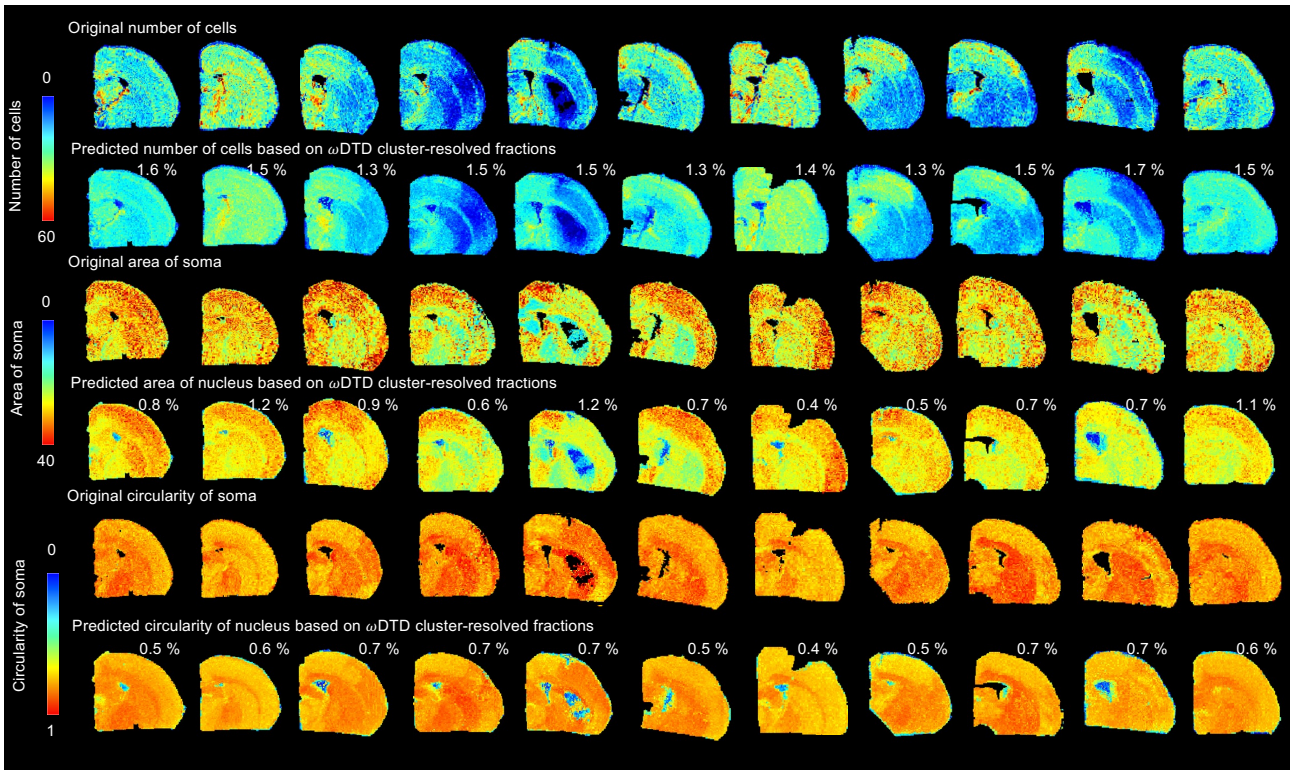

Figure S11: ***Predicted histological parameters based on non-linear RF regression analysis for all brains.*** *Comparison of original downsampled and registered histological parameters (rows) for the sham (first column) and all MCAO brains (columns 2-11) with their predicted counterparts. The predicted histological parameter maps were derived from training a RF model using ωDTD cluster-resolved mean and rates of change with frequency maps. The figure displays data across all the brain samples and histological parameters, providing a comprehensive comparison between the observed and predicted histological features. The percentage indicates average difference between the prediction and original.*

Histological parameter	Comparison	p-value per metric			
		E[R]	E[Q^2]	E[MSE]	E[MAE]
Number of cells	ωDTD Per-voxel v. ωDTD Bin-resolved	**0.0091****	**0.0022****	**0.0050****	**0.0018****
	ωDTD Per-voxel v. ωDTD Cluster-resolved	0.8770	**0.3936**	**0.6575**	**0.4336**
	ωDTD Bin-resolved v. ω DTD Cluster-resolved	0.0211*	0.7008	0.4889	0.5241
Area of Nucleus	ωDTD Per-voxel v. ωDTD Bin-resolved	**0.0179***	**0.0058****	**0.0078****	**0.0176***
	ωDTD Per-voxel v. ωDTD Cluster-resolved	0.5164	**0.1392**	**0.0485***	**0.3234**
	ωDTD Bin-resolved v. ω DTD Cluster-resolved	0.0475*	0.8556	**0.9514**	0.5186
Circularity of nucleus	ωDTD Per-voxel v. ωDTD Bin-resolved	**0.3839**	**0.2461**	**0.7251**	**0.1799**
	ωDTD Per-voxel v. ωDTD Cluster-resolved	0.3909	0.8524	0.1893	0.5608
	ωDTD Bin-resolved v. ω DTD Cluster-resolved	0.0600	0.2639	0.1088	0.1485

Table S3: ***Two-sided permutation test p-values on random forest LOO CV metrics. Bold*** *indicates instances where the 2nd ωDTD method performed better compared to the 1st ωDTD method. $^{*}p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$*