An automated pipeline for extracting histological stain area fraction for voxelwise quantitative MRI-histology comparisons ☆

Daniel Z.L. Kor [a,*], Saad Jbabdi [a], Istvan N. Huszar [a], Jeroen Mollink [a], Benjamin C. Tendler [a], Sean Foxley [b], Chaoyue Wang [a], Connor Scott [c], Adele Smart [a,c], Olaf Ansorge [c], Menuka Pallebage-Gamarallage [c], Karla L. Miller [a,1], Amy F.D. Howard [a,1]

[a] Wellcome Centre for Integrative Neuroimaging, FMRIB, Nuffield Department of Clinical Neurosciences, University of Oxford, Headington, Oxford OX3 9DU, , United Kingdom
[b] Department of Radiology, University of Chicago, Chicago, IL, United States of America
[c] Academic Unit of Neuropathology, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, United Kingdom

ARTICLE INFO

ABSTRACT

The acquisition of MRI and histology in the same post-mortem tissue sample enables direct correlation between MRI and histologically-derived parameters. However, there still lacks a standardised automated pipeline to process histology data, with most studies relying on manual intervention. Here, we introduce an automated pipeline to extract a quantitative histological measure for staining density (stain area fraction, SAF) from multiple immuno-histochemical (IHC) stains. The pipeline is designed to directly address key IHC artefacts related to tissue staining and slide digitisation. Here, the pipeline was applied to post-mortem human brain data from multiple subjects, relating MRI parameters (FA, MD, RD, AD, R2*, R1) to IHC slides stained for myelin, neurofilaments, microglia and activated microglia. Utilising high-quality MRI-histology co-registrations, we then performed whole-slide voxelwise comparisons (simple correlations, partial correlations and multiple regression analyses) between multimodal MRI- and IHC-derived parameters. The pipeline was found to be reproducible, robust to artefacts and generalisable across multiple IHC stains. Our partial correlation results suggest that some simple MRI-SAF correlations should be interpreted with caution, due to the co-localisation of other tissue features (e.g., myelin and neurofilaments). Further, we find activated microglia—a generic biomarker of inflammation—to consistently be the strongest predictor of high DTI FA and low RD, which may suggest sensitivity of diffusion MRI to aspects of neuroinflammation related to microglial activation, even after accounting for other microstructural changes (demyelination, axonal loss and general microglia infiltration). Together, these results show the utility of this approach in carefully curating IHC data and performing multimodal analyses to better understand microstructural relationships with MRI.

1. Introduction

Magnetic resonance imaging (MRI) is a powerful tool that can be used to evaluate neurodegenerative disorders in-vivo. MRI techniques have produced quantitative parameters sensitive to macroscopic neuropathological changes (Johnson et al., 2012; Geraldes et al., 2018; Grolez et al., 2016). However, MRI parameters are non-specific and sensitive to multiple factors related to tissue microstructure. Coupled with millimetre resolution, this leads to difficulty in determining the microstructural underpinnings of a given MRI change (e.g., in disease).

Immunohistochemistry (IHC) is a histological staining technique that can be used to address this difficulty. IHC uses primary antibodies to stain target antigens (proteins related to microstructural features-of-interest) with high specificity. Antigens are then commonly visualised using the chromogen 3′3-diaminobenzidine (DAB) to stain the marked proteins brown. To add contextual information and localise stained tissue features, haematoxylin is often used as a counterstain to mark cell nuclei purple. The acquisition of IHC can aid neuropathological diagnosis (Nave and Werner, 2014; Barker et al., 2013; Ulfig et al., 1998; Atik et al., 2014; Schirmer et al., 2011; Korzhevskii and Kirik, 2016; Jurga et al., 2020; Bagnato et al., 2011; Waller et al., 2019), with antibodies targeting either common tissue features (e.g., myelin or neurofilaments) or markers specific to pathology (e.g., activated microglia, a marker of inflammation (Bachiller et al., 2018; Geloso et al., 2017), or aggregated pTDP-43 in ALS (Pallebage-Gamarallage et al., 2018)). Further, the acquisition of IHC and MRI in the same post-mortem tissue

---

sample enables direct correlation of MR and histologically-derived metrics. This methodology is used to validate MR parameters and improve our microstructural interpretation of MRI. Here, post-mortem MRI functions as a crucial intermediary between IHC and in-vivo imaging. Post-mortem MRI shares a common tissue state with IHC, while possessing the same signal forming mechanisms with in-vivo MRI (Tendler et al., 2022).

While many studies relate IHC to MR parameters (Goubran et al., 2015; Fjær et al., 2013; Fjær et al., 2015; Dusek et al., 2017; Yano et al., 2018; Wang et al., 2020; Abe et al., 2019; Mollink et al., 2019; Lazari and Lipp, 2021), there is still a lack of automated pipelines for extracting quantitative metrics from IHC slides of neuronal tissue (Lazari and Lipp, 2021). Most MRI-IHC analyses rely on heavy manual intervention (van der Weijden et al., 2021; Mancini et al., 2020; De Barros et al., 2019; Seewann et al., 2009), with some using subjective metrics, such as staining intensity scores resembling *low, moderate* or *strong* (Seewann et al., 2009; Bulk et al., 2018). Others directly use the DAB channel's stain density to approximate the amount of targeted protein within the tissue (Seewann et al., 2009; Wiggermann et al., 2017; Hametner et al., 2018; Bagnato et al., 2018). However, this interpretation is limited as the densities do not scale linearly with protein density (De Barros et al., 2019; Hametner et al., 2018; van der Loos, 2008). To circumvent this, other pipelines extract the stain area fraction (SAF) i.e., the number of DAB-stained pixels within a given area. In most pipelines, SAF is quantified by manually setting a threshold for the DAB channel to segment microstructural tissue compartments from non-specific background staining in regions-of-interest (ROIs) (Lazari and Lipp, 2021; van der Weijden et al., 2021; Mancini et al., 2020; De Barros et al., 2019).

The manual derivation of SAF has two main issues. First, the manually-set threshold is dependant on the operator's expertise and a single threshold is often applied to all slides in a dataset or a batch of slides which have been processed together (Pallebage-Gamarallage et al., 2018). Although this is time efficient and avoids intra-observer variability, it comes at the expense of optimising thresholds for individual slides, resulting in decreased robustness to histological artefacts, such as slide-to-slide and within-slide staining intensity variation. Slide-to-slide staining intensity variation is introduced due to unintentional variations in tissue sample preparation and staining, which produce artificial (non-biological) differences in stain intensity and colour information. Within-slide staining artefacts include a gradual staining gradient, with stronger staining at one end of the slide progressing to weaker staining at the other, and striping artefacts from slide digitisation. Staining gradients arise due to how the slide is positioned during staining, inconsistent fixation of the tissue sample, and/or uneven application of the reagent (Suvarna et al., 2019). Striping artefacts describes sharp bands of intensity variation across whole slide images, which arise from when stitching together multiple strips of slide during digitization (Farahani et al., 2015). Though these artefacts could be eliminated at source (e.g., through optimised staining protocols, or an improved slide scanner), they are frequently observed in practice (Suvarna et al., 2019; Farahani et al., 2015; Gurcan et al., 2009; McCann et al., 2015; Madabhushi and Lee, 2016). Further, the impact of these artefacts on the SAF may not always be obvious at the point of slide preparation. Once these impacts are observed, re-staining may be infeasible given the considerable time and manpower already invested. When unaccounted for, these artefacts may impact the extracted IHC metrics. Second, these manual workflows are time intensive. This restricts research studies to smaller sample sizes (i.e., less slides and/or subjects), and limits IHC analyses to hand-drawn ROIs, as opposed to analysing voxels from the whole slide.

Here, we propose an automated SAF pipeline, in the context of an end-to-end MRI-histology workflow, to address these challenges. The automated pipeline is able to extract SAF maps from IHC-stained slides for myelin, neurofilaments and microglia. The pipeline was first evaluated on an IHC dataset designed to test the pipeline's reliability. We then applied the pipeline to a second dataset containing co-registered IHC and MRI to correlate SAF on a voxelwise basis with diffusion-weighted MRI (fractional anisotropy: FA; mean, radial, axial diffusivity: MD, RD, AD) and relaxometry (R2*, R1) maps. To account for covariance between stains, we use partial correlation to identify the unique variance in MRI parameters explained by each targeted protein. Finally, we perform multiple regression with all stains to derive a predictive model of each MR parameter, which may be driven by multiple microstructural sources.

2. Data acquisition

We applied the SAF pipeline to two datasets: one with only IHC data, and a second previously published dataset that includes both IHC and co-registered MRI (Pallebage-Gamarallage et al., 2018). Both datasets contain tissue from the same 15 post-mortem brains of patients diagnosed with amyotrophic lateral sclerosis (ALS) and healthy controls (CTL) (12 x ALS, 3 x CTL).

2.1. Immunohistochemistry data

After post-mortem MRI, tissue samples were extracted and stained, as described in (Pallebage-Gamarallage et al., 2018), using primary antibodies against PLP (myelin), SMI312 (neurofilaments), Iba1 (microglia) and CD68 (activated microglia, macrophages). These antibodies have applications in pathological and healthy tissue ((Nave and Werner, 2014; Barker et al., 2013) for PLP, (Ulfig et al., 1998; Atik et al., 2014; Schirmer et al., 2011) for SMI312, (Korzhevskii and Kirik, 2016; Jurga et al., 2020; Bagnato et al., 2011; Waller et al., 2019) for CD68 and Iba1) and are relevant to ALS neuropathology, which is characterised by neuronal loss and microglial activation in motor neuron areas (Pallebage-Gamarallage et al., 2018). All antibodies were visualised with DAB and sections were counterstained with hematoxylin. Non-specific staining of endogenous peroxidase—an enzyme present in many cells—was minimised through a peroxidise blocking step (Bussolati and Radulescu, 2011; Del Cerro et al., 1981). Residual non-specific staining was generally faint except for occasional darker stained vasculature that was sparsely distributed across the slides. Slides were digitised with the Aperio ScanScope® AT Turbo (Leica Biosystems) at x20 object magnification (0.5 μm/pixel). Prior to analysis, each slide was manually quality-checked (QC) where we excluded slides that would fail analysis due to excessive illumination, tears due to poor tissue sectioning or significant amounts of inconsistent staining. Inconsistent staining refers to atypical staining patterns that are non-biological in origin e.g., extremely strong staining gradients, inverted white/grey matter contrasts (opposite to what we neuroanatomically expect) and patchy or unstained tissue.

2.1.1. Evaluation dataset

The "evaluation dataset" was acquired to evaluate our pipeline's performance in terms of reproducibility and robustness to key histological artefacts. Ideally, this data can distinguish true biological variation from variance related to artefacts and analysis. Consequently, data was collected from adjacent tissue slides, which are separated by the slide thickness (6 μm). We assume that adjacent slides have similar microstructure and that the true biological between-slide variance is low. This a priori assumption is reasonable as we aim to summarise microstructure at MRI resolution (0.5–1 mm), which is considerably larger than the slides' separation. At the tissue boundaries and for proteins-of-interest that sparsely populate the tissue, this assumption may not always be met. For each of the 15 brains, 12–15 adjacent slides were obtained from the primary motor cortex (face region). These slides were separated into groups of 4–5 slides. Each group was stained for CD68, PLP or SMI312. During QC, we removed 49% of the CD68 and 24% of the SMI312 slides. No PLP slides were excluded. CD68 slides were mostly removed due to staining artefacts. This is likely due to some difficulty with this batch's

staining, rather than a general issue with the CD68 stain, given that the multiple-region dataset's CD68 slides all passed QC.

2.1.2. Multiple-region dataset

We applied our SAF pipeline (Section 3.1) on the "multiple-region dataset" (Pallebage-Gamarallage et al., 2018) to generate SAF maps. This dataset includes IHC and co-registered MRI data from multiple brain regions with varying levels of disease pathology. Both the dataset and its SAF maps will be available in a future release of the Digital Brain Bank (Tendler et al., 2022). We consider data from the visual cortex (both hemispheres), anterior cingulate (cingulum bundle, corpus callosum) and hippocampus. Slides were stained to visualise CD68, Iba1, PLP and SMI312. All slides passed QC.

2.2. MRI data

The data acquisition and pre-processing of the post-mortem MRI have been previously described (Pallebage-Gamarallage et al., 2018; Wang et al., 2020; Tendler et al., 2020). Whole brains were imaged in a 7T human scanner (Siemens Healthcare, Erlangen, Germany) using a 1Tx/32Rx head coil. $R2^*$ maps were estimated from susceptibility-weighted data acquired with a 3D multi-echo GRE sequence (parameters: TEs = 2, 8.6, 15.2, 21.8, 28.4, 35 ms with monopolar readout and non-selective RF pulse, TR = 38 ms, flip angle = 15°, bandwidth = 650 Hz/pixel, and in-plane resolution = 0.5×0.5 mm^2 (Wang et al., 2020)). R1 maps were estimated from T1-weighted data collected with a multi-TI turbo spin-echo protocol (Subject 1,2: TE = 14.2 ms, TR = 1000 ms, TIs = 30, 60, 120, 240, 480, 935 ms, flip angles = 180°, bandwidth = 130 Hz/pixel, and in-plane resolution = 1.0×1.0 mm^2 (Pallebage-Gamarallage et al., 2018)). Other brains were imaged with slightly different parameters (Table S1). Diffusion tensor maps of FA, MD, AD and RD (Basser et al., 1994) were estimated from a diffusion-weighted steady-state free precession sequence (parameters: TE = 21.0 ms, TR = 28.0 ms, flip angles = 24°, 94°, bandwidth = 393 Hz/pixel, q value = 300 cm^{-1}, number of directions/flip angle = 120, and in-plane resolution = 0.85×0.85 mm^2 (Tendler et al., 2020)).

2.3. Co-registration of MRI and histology

MRI and histology data were previously co-registered using FSL's Tensor Image Registration Library (TIRL), which is a general-purpose image registration framework designed for MRI-histology co-registration (Huszar et al., 2019). We used TIRL to 1) register PLP with structural MRI (2D-3D) and 2) co-register other stains to PLP (2D-2D). PLP was chosen as the reference histology data due to its strong white/grey matter (WM/GM) contrast. All MR parameters maps were first aligned to structural MRI using FLIRT (Jenkinson et al., 2002). The generated warps were combined using TIRL to map these MR parameters to the IHC SAF maps for voxelwise correlations. To investigate relationships in WM voxels, WM masks were derived from maps of the third eigenvalue of the diffusion tensor using FSL-FAST (Zhang et al., 2001). WM masks were similarly resampled into the 2D PLP space.

3. Methods

We describe an end-to-end workflow for MRI-SAF comparisons. At the centre of this workflow is our automated SAF pipeline, which is described in Section 3.1. The pipeline was designed to be generalisable to multiple IHC stains, reproducible and robust to common IHC artefacts, which we evaluate in Section 3.2. The pipeline was then applied to the multiple-region dataset, which includes IHC and co-registered MRI, facilitating voxelwise MRI-SAF comparisons. We describe how we evaluated the quality of the co-registration (Section 3.3) and performed voxelwise MRI-SAF analyses (Sections 3.4, 3.5) to disentangle the contributions of multiple microstructural features stained with IHC to each MR parameter.

3.1. SAF pipeline

The pipeline maps RGB intensity values of high-resolution IHC slides (0.5 μm/pixel) to SAF using three steps: 1) separation of the DAB and haematoxylin stains using colour deconvolution, 2) segmentation of the DAB-stained proteins-of-interest from non-specific DAB using an intensity threshold, and 3) calculation of the SAF map at a specified resolution. The pipeline is automated and data-driven, enabling rapid analysis of many IHC slides across multiple subjects, regions, and stains.

In practice, we found considerable impact of IHC artefacts in some slides. This motivated the development of two "configurations" of the pipeline: the default (Fig. 1) and artefact (Fig. 2) configuration. The default configuration is designed to emulate an expert histologist when deriving SAF (but with data-driven, slide-specific stain separation and thresholding) for IHC slides with no prominent staining gradient and/or striping artefacts. This improves upon a more manual approach, where the adjustment of slide-specific thresholds would be extremely time-consuming. If these artefacts are present, we propose the artefact configuration, which automatically adjusts the local thresholds within-slide to account for the impact of these artefacts. Both configurations are automated, and differ based on whether local or whole-slide methods are used for Steps (1) and (2). We now describe each step of the pipeline in detail.

3.1.1. Colour matrix derivation for stain separation

Colourimetric analysis is based on colour deconvolution, a stain separation method using Beer-Lambert's law (Ruifrok and Johnston, 2001). In Beer-Lambert's law (c.f. Appendix A), the light absorbance (**A**) for RGB channels is linearly related to the stain density (**C**) via the attenuation coefficients (ϵ). The stain (i.e., DAB, hematoxylin) density in a slide can be computed with a matrix inversion:

$$\begin{bmatrix} C_{DAB} \\ C_{hematoxylin} \\ C_{residual} \end{bmatrix} = \begin{bmatrix} \epsilon_{DAB,red} & \epsilon_{DAB,green} & \epsilon_{DAB,blue} \\ \epsilon_{hema,red} & \epsilon_{hema,green} & \epsilon_{hema,blue} \\ \epsilon_{res,red} & \epsilon_{res,green} & \epsilon_{res,blue} \end{bmatrix}^{-1} \begin{bmatrix} A_{red} \\ A_{green} \\ A_{blue} \end{bmatrix}$$

$$= M^{-1} \begin{bmatrix} A_{red} \\ A_{green} \\ A_{blue} \end{bmatrix} \tag{1}$$

where M, termed the colour matrix, is a 3×3 matrix with each element giving the attenuation coefficient ($\epsilon_{i,\lambda}$) for each stain (i) and colour (λ) pair. **M** can be defined from literature or empirically measured from single-stained slides (Ruifrok and Johnston, 2001; Landini et al., 2021). Using an **M** that is non-specific to each slide leads to poor stain separation and consequently, inconsistent interpretation during slide-to-slide comparison (Gurcan et al., 2009; Madabhushi and Lee, 2016; Clarke and Treanor, 2017). We address this problem by employing a k-means clustering approach to derive colour information directly from the IHC data. This approach is tailored according to the different configurations.

Default configuration

To derive a single slide-specific **M**, we used the k-means clustering strategy introduced in (Geijs et al., 2018). We randomly sampled patches (0.5 × 0.5 mm^2; n = 200) from the slide and performed k-means clustering (k = 2) to produce 2 cluster centroids corresponding to the colour vectors (i.e., rows in **M**) of hematoxylin and DAB. We then performed a second k-means clustering (k = 2) on the 200 colour vectors output from the first k-means. This produced two centroids corresponding to the slide-specific DAB and haematoxylin colour vectors of **M**. Stain separation was performed using this single **M** across the whole slide.

Artefact configuration

In slides with a staining gradient, a locally changing **M** is required to account for colour differences from one end of the slide to the other. Consequently, we apply k-means locally to derive **M** along the gradient direction. In principle, **M** can be derived in windows of pixels of any arbitrary shape (i.e., square patch) that are large enough to provide a
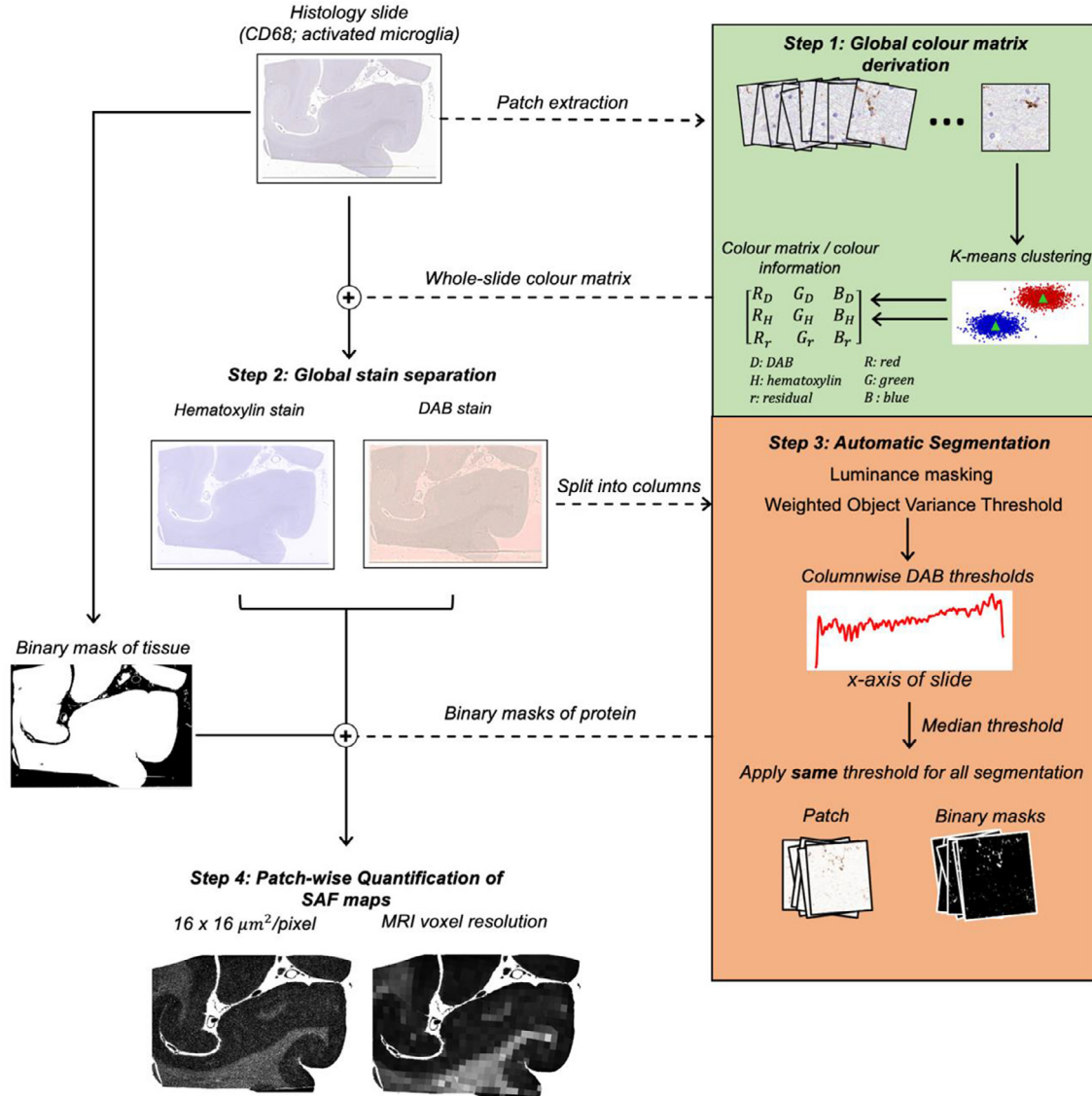
Slides without visible artefacts

Fig. 1. The automated SAF pipeline for stains with little or no intensity gradient and/or visible stitching artefacts (example: visual cortex). This usually includes stains specifying structures that sparsely populate the brain tissue, such as microglia (Iba1, CD68) and neurofilament (SMI312) in some cases. For each slide, the pipeline 1) derives a single, global colour matrix to separate DAB from haematoxylin, 2) performs stain separation to isolate the DAB channel, 3) automatically segments the DAB channel with a single median threshold, and 4) calculates the SAF at variable resolution.

sufficient number of pixels of each class (background versus stain-of-interest) for data-driven analysis. In our datasets, the staining gradient was present along the horizontal axis of the slide. Hence, we applied k-means on a column-wise basis (32-pixels width; height matching the slide) to define $\mathbf{M_{column}}$ that varies along the gradient direction. Stain separation was performed locally using $\mathbf{M_{column}}$.

For stain separation in both configurations, an approach emulating NNLS (non-negative least squares) (c.f. Appendix A), was used to estimate the DAB channel's stain density ($\mathbf{C_{DAB}}$) according to Eq. (1). $\mathbf{C_{DAB}}$ was then converted to intensity values ($\mathbf{I}$) using an equation analogous to the inverse Beer-Lambert's Law: $\mathbf{I} = 10^{-\mathbf{C_{DAB}}}$.

3.1.2. Automatic thresholding for protein segmentation

To segment the DAB-stained protein-of-interest from non-specific DAB, manual analyses require an expert to manually set a threshold. We improve on this by deriving a data-driven threshold. An image segmentation approach is Otsu's method (Otsu, 1979), which computes a

threshold that maximises the image histogram's inter-class variance. Our algorithm is based on the related weighted object variance algorithm (Yuan et al., 2015), which includes a tunable parameter (δ) to separate classes of unequal count and variance. Here, we computed a threshold $\mathbf{t}$ by maximizing:

$$t = \arg\max_t \left[P_0(t)\mu_0(t)^2 + P_1(t)^{1+\delta}\mu_1(t)^2 \right] \qquad (2)$$

Here, $\mathbf{P_j(t)}$ and $\mu_j\mathbf{(t)}$ are the cumulative probabilities and means of the uneven classes $j = 0$ (protein-of-interest) and $j = 1$ (DAB-stained background). δ ranges from -1 to 1 and weights the object variance to shift $\mathbf{t}$ closer to the mean of the smaller ($\delta > 0$) or larger ($\delta < 0$) class 0. $\delta < 0$ was used for densely stained PLP or SMI312 and $\delta > 0$ for sparsely stained slides (CD68, Iba1). To achieve local thresholding, the algorithm was performed on contiguous columns rather than the entire image. Our slides possessed left-to-right gradients in stain density, with vertical bands of intensity (striping). Thresholds $\mathbf{t_i}$ were thus calculated on a column-wise basis (32-pixels width; height matching the slide), with
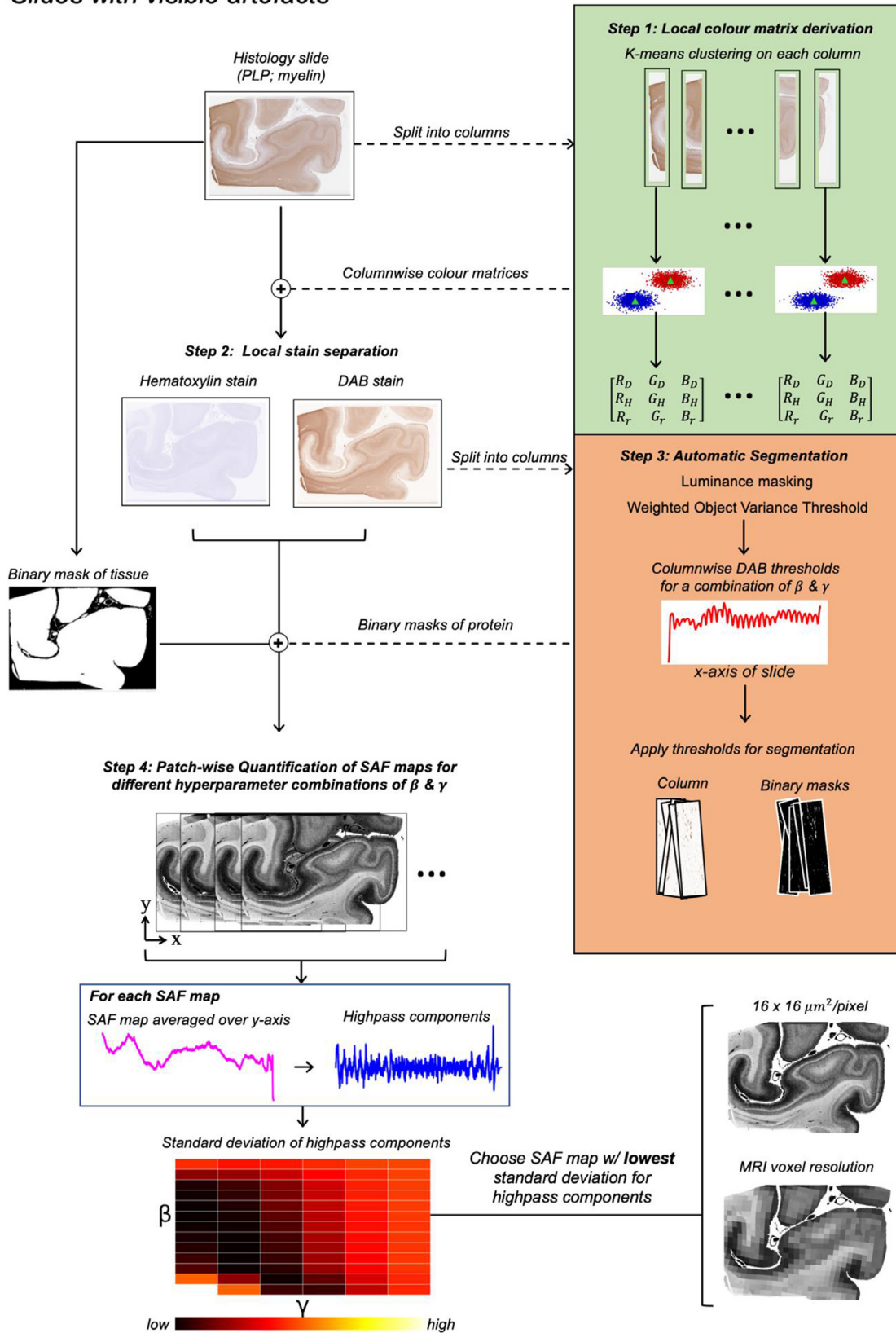
Fig. 2. The automated SAF pipeline for stains confounded by staining gradient and/or visible striping artefacts (visual cortex). For each slide, the pipeline 1) derives multiple, columnar colour matrices from the data, 2) performs stain separation for each column, 3) automatically segments each column's DAB channel and 4) forms an SAF map. Steps 3 and 4 are repeated for a range of hyperparameters β and γ, which are optimised via grid-search to account for within-slide artefacts (Section 3.1.2). β modulates how much correction is needed to offset the impact of the striping artefact, while γ is a smoothing kernel required to prevent abrupt column-to-column correction. For each SAF map, we averaged the map over its height (y-axis) and chose the SAF map with the lowest standard deviation (i.e., least impacted by these artefacts).

i representing the column index. Before calculating the threshold, the DAB channel was masked to remove the non-tissue pixels in the column (luminance < 0.75) to balance the two classes in the DAB channel histogram. This prevented the large number of non-tissue pixels from skewing the histogram and biasing the threshold.

Default configuration

A single threshold was sufficient to segment the entire slide without visible striping artefacts. The fixed threshold was calculated as the median of all local thresholds computed column-wise. Consequently, our data-driven pipeline replaces expert determination of the threshold, **t**, with determination of the hyperparameter, **δ**. Ideally, this hyperparameter is set once by an expert for a stain. Our pipeline will then enable automated, adaptive thresholding on new slides. In this work, we sampled 8–10 patches (0.5×0.5 mm^2) spanning different brain regions, tissue types and subjects to choose a δ (per stain) that produced optimal segmentation. Segmentations were vetted with an expert histologist (MPG).

Artefact configuration

In slides with striping artefacts, a single whole-slide threshold is insufficient for optimal segmentation. Our artefact configuration uses column-wise thresholds that adapt to these artefacts in a data-driven way. Columns affected by striping are characterised by DAB histograms with decreased median absolute deviation (MAD) compared to columns less affected by striping artefacts. This was found by empirically comparing histograms from striping and non-striping regions and associating it with an observed slight blurring of the image along the stitching boundary of the slide scanner. Consequently, we use the MAD from each column (indexed with **i**) to weight the δ to account for how much striping artefact is present:

$$\delta = \alpha(1 + \Delta_i)^\beta \qquad (3)$$

where:

$$\Delta_i = \frac{MAD_i - \sigma(MAD_i)}{\sigma(MAD_i)} \qquad (4)$$

δ_i is the overall exponent used in WOV equation (Equation 2); α is a stain-specific value chosen manually which depends on the relative positive stain and background staining (equivalent to δ in the default pipeline); β is chosen via grid search and depends on amount of striping artefact i.e., it modulates how much δ differs from α due to striping; σ is the kernel size of a 1D Gaussian filter applied to **MAD$_i$** (σ=16); Δ_i represents the **ith** column's change in structure (MAD) relative to neighbouring columns. After calculating the column-wise thresholds using Eqs. 2-4, we performed a final smoothing operation:

$$t_{final} = \gamma(t_i) \qquad (5)$$

where γ is a 1D Gaussian filter that it optimised via grid search that prevents the thresholds (**t$_i$**) from varying abruptly column-to-column.

Figures S1 and S2 demonstrate how the hyperparameters α, β and γ affect the resultant SAF map. α is chosen manually, β and γ are optimised for each slide via grid search to minimise the standard deviation of highpass components (defined below) in the resulting SAF. In our multiple-region dataset, We observed stitching artefacts and/or staining gradients in the PLP slides only. Hence, we applied the artefact configuration on our PLP slides (α=−0.60) and the default configuration on our CD68 (δ=0.05), Iba1 (δ=0.05) and SMI312 (δ=−0.30) slides. Figure S3 shows example segmentations.

3.1.3. Calculating the SAF map

Following segmentation, we computed SAF maps at various resolutions. First, we extracted a tissue mask by applying Otsu's method (Otsu, 1979) on the hematoxylin channel followed by several morphological operations. SAF was defined as the ratio of pixels with positive DAB to those in the tissue mask within a patch. We evaluate the robustness of our pipeline by generating SAF maps at high resolution (16×16

μm^2/patch), which facilitates better identification of high-frequency variations in SAF, and at MRI resolution (0.5×0.5 mm^2/patch), to evaluate the reproducibility of SAF at the resolution used for MRI-SAF analyses.

3.2. Evaluation of SAF pipeline

The pipeline was evaluated with two criteria: robustness to artefacts and reproducibility of SAF. These analyses used the evaluation dataset (c.f. Section 2.1.1). We compared pipeline-derived SAF maps with expert-derived (MPG) SAF maps (Pallebage-Gamarallage et al., 2018) that use a manually-derived dataset-specific colour matrix and a stain-specific segmentation threshold. Each manual threshold was calibrated in at least 10 randomly selected, structurally distinct regions.

Robustness to within-slide artefacts

To test the pipeline's robustness to within-slide artefacts, we calculated the average column-wise SAF. This horizontal SAF profile was filtered using three bandpass filters (modelled from a Butterworth filter (Butterworth, 1930; Virtanen et al., 2020) applied twice) that are sensitive to different artefacts' effects: staining gradients (lowpass; < 3 Hz), striping artefact (bandpass; 3–12 Hz) and bandpass plus all other high frequency noise (highpass; > 3 Hz). The frequency bands were qualitatively chosen based on the spatial variability of the artefacts, which may differ between datasets. Within each frequency band, we computed the relative percent change in standard deviation between the manually- and automatically-derived SAF map (**diffstd**) to quantify the reduction in impact from artefacts:

$$diffstd = \frac{std_{manual} - std_{automated}}{std_{manual}} \cdot 100 \qquad (6)$$

diffstd was observed to be sensitive to artefacts that increase the variance of SAF across the slide: striping artefacts increase **diffstd** in the high frequency band, and staining gradients increase **diffstd** in the low frequency band.

Reproducibility of SAF

Reproducibility was tested by registering all within-subject slides to the subject's first slide via TIRL (Huszar et al., 2019). We compute an SAF difference map for each co-registered pair of slides (SAF$_{1,2}$) and take the median value across pixels:

$$diffsaf = median\left[\frac{SAF_1 - SAF_2}{(SAF_1 + SAF_2)/2} \cdot 100\right] \qquad (7)$$

3.3. Co-registration of MRI and histology

The pipeline was applied to data from the multiple-region dataset (c.f. Section 2.1.2) with co-registered MRI-histology. The quality of the registration was evaluated according to the alignment of contours representing tissue boundaries and WM/GM contrast. Contours were derived from the PLP SAF maps.

3.4. Voxelwise MRI-SAF analyses

MR parameter maps (3D) were resampled into PLP (2D) space for correlation with SAF. As the MR voxels do not align with the 0.5 mm histology grid, the SAF were recalculated for each MR voxel, minimising interpolation effects. Pixels in the SAF maps with values < 5th percentile were identified as non-tissue pixels and discarded. Other outliers were identified using a Huber influence function (tuning coefficient = 2.5), excluding data points with weights < 0.75 (Huber, 2011). We pooled voxelwise data across regions and brains for several linear model analyses between MRI and SAF:

- **Simple correlation**: We correlated each pair of MR (FA, MD, RD, AD, R2*, R1) and SAF (PLP, SMI312, Iba1, CD68) parameters.
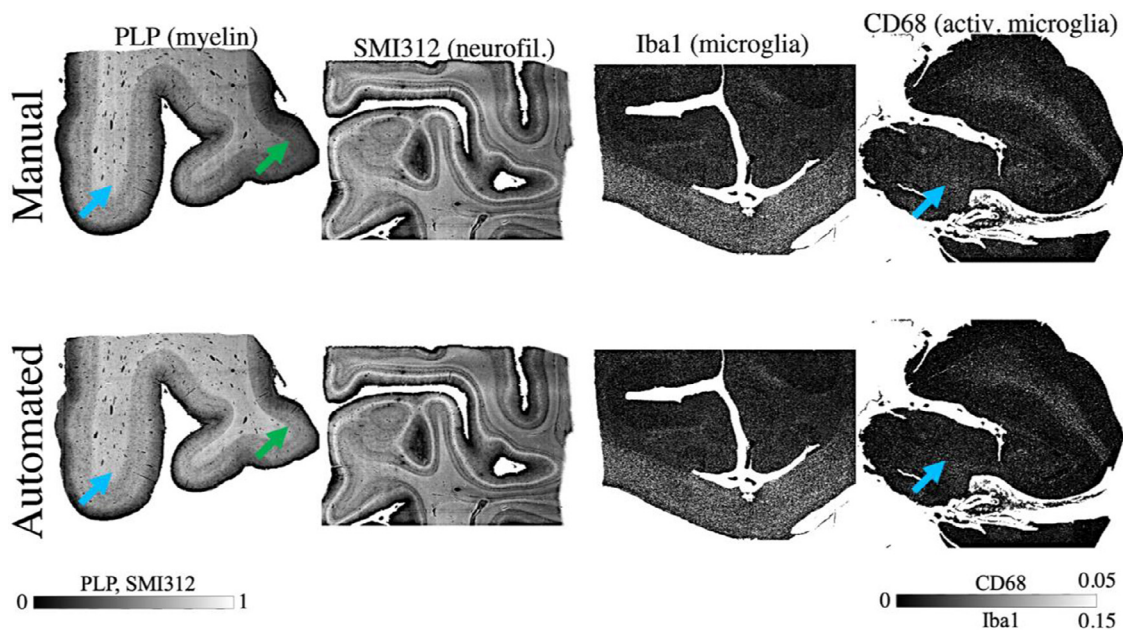
Fig. 3. Within-slide artefacts are reduced using the automated pipeline (bottom row). These include noticeable stitching artefacts (blue arrows) and staining gradients (green arrows) originally seen in manually-derived SAF maps (top row). Examples here show PLP slides processed with the pipeline configuration that corrects for staining gradient artefacts (artefact configuration), while SMI312, Iba1 and CD68 slides were processed with the pipeline's default configuration.

- **Partial correlation**: We estimated the unique variance of each MR parameter that is explained by a given stain's SAF, accounting for other stains. Subject ID was used as a covariate to account for between-subject confounds.
- **Multiple linear regression**: We model each MR parameter as a linear combination of all stains (explanatory variables) and the subject ID (confounding variable). We computed each variable's relative importance measure (Grömping, 2015) i.e., the averaged relative contribution of each variable in explaining the variance of the MR parameter after it is added to the model.

As a proof-of-concept, we performed voxelwise analyses on two subjects from the multiple-region dataset (1 x CTL, subject 1; 1 x ALS, subject 2). Analyses were done using WM+GM or WM voxels, where the latter demonstrates sensitivity to subtle microstructural changes rather than gross WM/GM tissue differences

3.5. Application to multiple subjects for MRI-PLP

This pipeline can be rapidly applied across many subjects. We demonstrate this by applying the workflow (Sections 3.1, 3.3, 3.4) to 41 PLP slides from ten additional subjects. We correlated MRI with PLP only, as co-registration between other stains and MR data was a work-in-progress. We compared the ten subjects' output with the two subjects previously analysed. Subject ID was modelled as a covariate.

4. Results

4.1. Evaluation of SAF pipeline

Robustness to within-slide artefacts

The evaluation dataset slides were processed with either the default (CD68, Iba1) or artefact (PLP, SMI312) pipeline. Fig. 3 compares SAF maps from both our automated and manual pipeline. Both configurations of the automated pipeline reduce the impact of stitching artefacts. In the default configuration, this can be attributed to the use of a slide-specific colour matrix, which results in better stain separation compared to the non-specific colour matrices used in manual analyses. The artefact

configuration also reduced staining gradient artefacts. In slides with less noticeable artefacts (Iba1), the manual and automated pipelines produce similar results.

In almost all PLP and SMI312 slides (Fig. 4), the **diffstd** for each frequency band (Eq. (6)) was positive, suggesting a reduction in artefacts for the automated pipeline. There was high similarity in highpass and bandpass **diffstd**. This implies the striping artefact's major contribution to overall noise in these slides, and how its impact is especially mitigated with the automated pipeline. The results for CD68 are less conclusive. The automated pipeline reduces the impact of the striping artefact (bandpass **diffstd** is mostly positive) but performs worse than the manual pipeline for the highpass and lowpass filters.

Reproducibility of SAF

Reproducibility of SAF maps (Fig. 5) was quantified with **diffsaf** (Eq. (7)) between within-subject adjacent slides. Manual and automated pipelines were found to have similar reproducibility with a median **diffsaf** of around 5% (PLP), 9% (SMI312) and 20% (CD68). As fewer slides were used when evaluating CD68 (9/15 subjects; 37/73 slides) than in SMI312 (13/15 subjects; 55/72 slides) and PLP (15/15 subjects; 73/73 slides) after manual quality control (c.f. Section 2.1.1), Figs. 4,5 have fewer datapoints.

4.2. Co-registration of MRI and histology

Fig. 6 shows good alignment of the tissue mask (green contour) and WM/GM interface (red contour) for subject 1. Subject 2 had similar results (Figure S4). Of the remaining subjects, 10 had good quality registrations (Figure S5) and 3 were excluded from subsequent analysis due to substantial misalignment. Additional quantitative evaluation of the dataset registration can be found in (Huszar et al., 2019).

4.3. Voxelwise MRI-SAF analyses

Eight IHC slides from different regions (Figs. 6, S5) in subjects 1 and 2 were mapped onto MRI data. For each MR parameter, up to 3% of voxels were classified as outliers and removed.
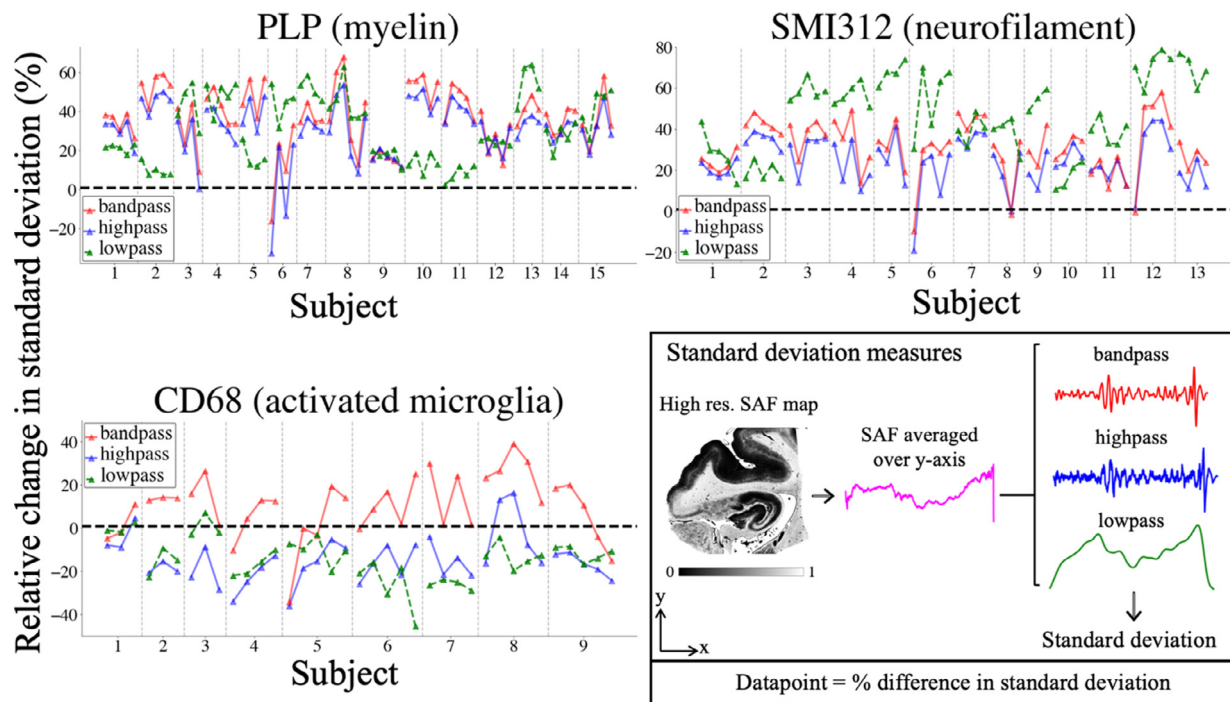
Fig. 4. Quantitatively comparing the impact of within-slide artefacts in the manual and automated processing pipelines, as measured using the relative change in standard deviation (100 x (std$_{manual}$ - std$_{automatic}$) / std$_{manual}$). Standard deviation measures (box) are derived by 1) averaging the high-resolution (16 μm/pixel) SAF map along the y-axis, 2) filtering it to produce 3 components (bandpass, highpass and lowpass), and 3) computing the components' standard deviation. The lowpass filter isolates the staining gradient artefact, whilst the stitching artefact only is captured with the bandpass filter. The highpass filter combines the same striping artefact with all high frequency noise. Each IHC slide is represented as a single datapoint, and slides from the same subject are grouped together and connected by lines. A positive (negative) value indicates that the automatically-derived SAF map is less (more) affected by the associated artefact than the manually-derived SAF map.

4.3.1. Simple correlation

Simple correlation was used to relate each MRI-SAF pair (Fig. 7). While correlations with Iba1 appear low (|r| = 0.035–0.28), CD68 correlated well with FA (r = 0.56) and MD (r = −0.39). The scatterplots also suggest nonlinear trends for FA with PLP and SMI312. Similar analysis of WM only is shown in Figure S6, where FA shows a strong correlation with CD68 (r = 0.33) and R2* with Iba1 (r = −0.30). Scatterplots with AD and RD are shown in Figure S7. In WM+GM, AD and RD correlated negatively with all stains. When analysing the WM only, notable results include the correlation of AD with Iba1 (r = 0.27) and RD with CD68 (r = −0.21).

4.3.2. Partial correlation

For each stain, correlations were calculated after regressing out one other stain or all other stains as covariates. We show correlation coefficients when including all voxels (WM+GM) (Fig. 8A), or WM only (Fig. 8B). Similar data for AD and RD are shown in Figure S8.

WM and GM

When accounting for all other stains (Figs. 8A and S8A, bottom rows), FA is best explained by CD68 (r = 0.38), MD/AD/RD/R2* by PLP (r = −0.42/r = −0.65/r = −0.68/r = 0.30) and R1 by SMI312 (r = 0.36). We observed two main results: the shared explained variance between PLP (myelin) and SMI312 (neurofilaments), and between Iba1 (all microglia) and CD68 (activated microglia).

First, the correlation coefficient between FA and SMI312 is greatly reduced after accounting for PLP (r = 0.35 to 0.026), whilst the correlation coefficient between FA and PLP is marginally reduced after accounting for SMI312 (r = 0.54 to 0.44). Similar effects are observed between MD and PLP/SMI312, and between R2* and PLP/SMI312. The opposite effect is seen in R1: the correlation coefficient between R1 and PLP

decreases significantly when covarying for SMI312 (r = 0.46 to 0.19), but the correlation coefficient between R1 and SMI312 is marginally reduced when accounting for PLP (r = 0.54 to 0.36). Second, we saw that the high correlations of FA/RD with CD68 marginally decrease after accounting for Iba1 (r$_{FA}$ = 0.56 to 0.5; r$_{RD}$ = −0.50 to −0.45). Conversely, the apparent correlations with Iba1 are minimal after accounting for CD68 (r$_{FA}$ = 0.28 to 0.061; r$_{RD}$ = −0.24 to −0.029). We note similar trends for correlations of MD with CD68/Iba1.

WM only

Figs. 8B and S8B show the same analysis for voxels found in the WM only. When accounting for all other stains, FA/MD/RD are best explained by CD68 (r = 0.26/r = −0.18/r = −0.22), AD by PLP (r = 0.10), R2*/R1 by Iba1 (r = −0.11/ r = −0.23). When compared to our previous results for WM+GM voxels, the correlation coefficient between most MR parameters and PLP/SMI312 are greatly reduced and/or close to zero. This behaviour is not seen for CD68 and Iba1. In WM, CD68 explains the most unique variance in FA and RD. Notably, these correlations remain relatively unchanged when accounting for other stains (r$_{FA}$ = 0.33 to 0.26; r$_{RD}$ = −0.29 to −0.22). Similar behaviour is observed when relating CD68 with MD, and Iba1 with R1/R2*.

4.3.3. Multiple linear regression

We show the regression and fitted correlation coefficients (i.e., square root of the coefficient of determination) when including voxels from WM+GM, or WM only (Tables 1A, S2A). Using multiple stains better explains the variance in MR parameters than individual stains. We also show the relative importance of each predictor (Tables 1B, S2B). The subject predictor had the highest relative importance in predicting R1 in all voxels, and MD/R2*/R1 in WM voxels only. This implies that between-subject confounds substantially influence MR parameters.
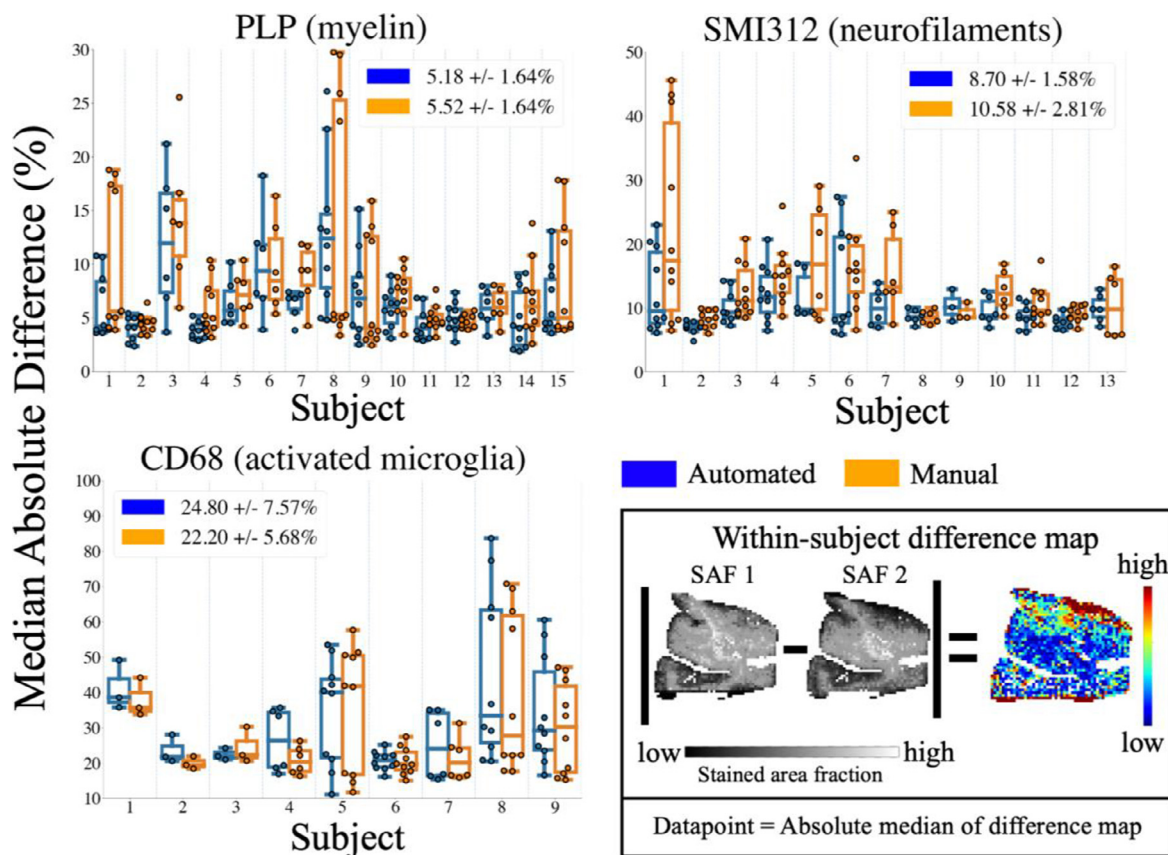
Fig. 5. The reproducibility of SAF values across adjacent tissue sections, as measured by the median absolute difference. The median absolute difference (box) was computed by 1) taking the absolute difference of a pair of SAF maps produced from IHC slides extracted from the same region and subject, 2) normalising this difference map by the mean of both SAF maps, and 3) taking the median to represent one datapoint. We used this measure to compare reproducibility of both manually- (orange) and automatically-derived (blue) SAF maps for PLP, SMI312 and CD68. The median and median absolute variance of data points are also shown for each method (legends). Note the change in scale-bar for CD68, where the difference values were generally larger when compared to PLP and SMI312.

4.4. Application to multiple subjects for MRI-PLP

Similar regression coefficients were seen for the larger group ($n = 10$, Figure S9 right) compared to the original analysis ($n = 2$, Figure S9 left), with the estimated effect (β) being highly consistent for FA, MD and R2*, but less consistent for R1.

5. Discussion

To facilitate high-throughput MRI-histology analyses, we introduce an automated pipeline to extract stain area fraction (SAF) from immuno-histochemical (IHC) stains. Using high-quality co-registration, we performed whole-slide voxelwise MRI-SAF comparisons. The pipeline was applied to post-mortem human brain data from multiple subjects, relating slides stained for myelin (PLP), neurofilaments (SMI312), microglia (Iba1) and activated microglia (CD68) to MR parameters (FA, MD, RD, AD, R2*, R1).

Our approach has three advantages over previous studies. First, while most MRI-SAF studies use manually-derived SAF, our pipeline deploys data-driven algorithms (Geijs et al., 2018; Yuan et al., 2015). This enables the pipeline to be applied rapidly to larger datasets without expert intervention. Second, our pipeline can handle key artefacts (non-biological sources of variation) which impact SAF maps. This is an improvement over previous automated pipelines (Grussu et al., 2017; Khodanovich et al., 2019; Praet et al., 2018). Third, the pipeline is generalisable to multiple IHC stains, allowing for analysis spanning multiple microstructural sources. Combined with recent advances in MRI-histology co-registration (Huszar et al., 2019), the pipeline may enable

standardised analyses for voxelwise MRI-SAF comparisons (Lazari and Lipp, 2021).

The pipeline facilitates voxelwise MRI-SAF analyses that inspire more confidence than ROI-based analyses, which are biased by the choice of ROI and/or dilute localised effects-of-interest. Further, most studies perform simple correlations with only a single or a few histological markers (Yano et al., 2018; Seewann et al., 2009; Peters et al., 2019). We extend this analysis to partial correlation and multiple linear regression relating multimodal MRI to multiple IHC stains. This presents two key benefits. First, partial correlations can demonstrate sensitivity to specific microstructural changes such as microglial activation (changes in CD68 but not Iba1), demyelination (changes in PLP but not SMI312) and axonal loss (simultaneous changes in SMI312 and PLP). This may aid with disease diagnosis and/or staging. Here, our pipeline is applied to a dataset that was specifically acquired for future investigation into ALS neuropathology, as motivated in our previous publication that presented the study design (Pallebage-Gamarallage et al., 2018). Currently, we refrain from making any strong conclusions with respect to ALS pathology since our MRI-SAF results are from only two subjects, where co-registration for other subjects is still a work-in-progress. However, future work will draw on MRI-SAF results from the entire multi-subject dataset to provide an in-depth evaluation of the cellular bases of MRI in the context of ALS neuropathology. Second, both the partial correlation and multiple linear regression analyses can account for confound variables, which may otherwise influence our results. Subject-specific confounds contributed to significant variation in R1/MD/R2* (relative importance in Table 1B). This may be due to different post-mortem intervals (3 versus 2 days) and/or fixation times (45 versus 139 days), both
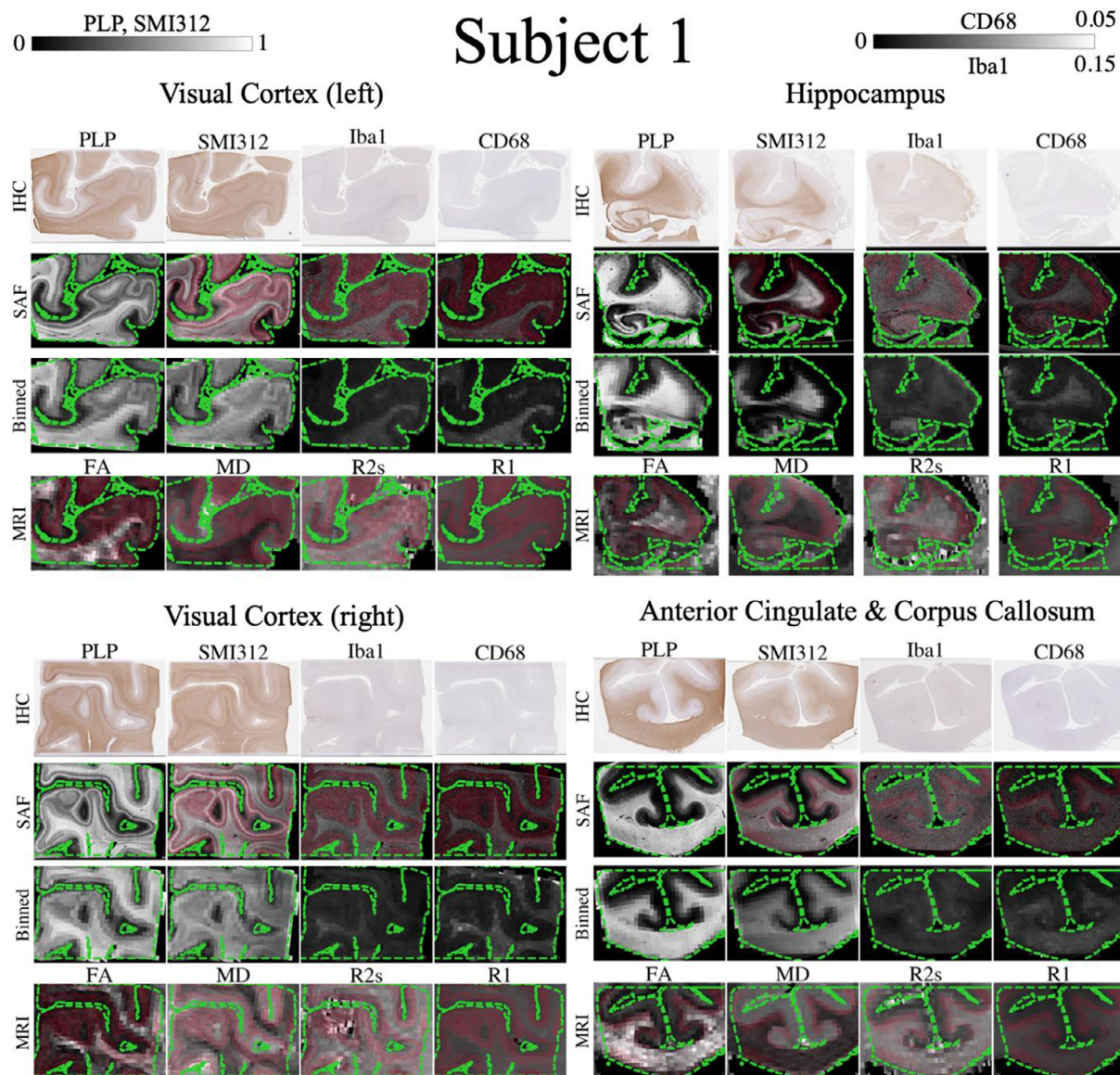
Fig. 6. Registration evaluation for all brain regions in subject 1. In each brain region, contours of the tissue mask (green dashed) are overlaid on the co-registered SAF maps at high resolution (first row, 1 pixel here represents SAF calculated in a $16 \times 16\ \mu m^2$ patch), SAF maps matching MRI resolution (second row) and MR parameter maps (third row). The white and grey matter interface is shown in red. The tissue boundaries are closely aligned and the high registration accuracy enables us to perform meaningful voxelwise MRI-histology correlations.

of which are known to influence MR parameters (R1, R2*: (Shatil et al., 2018; McNab et al., 2009; Foxley et al., 2014); MD: (McNab et al., 2009; Foxley et al., 2014)). Notably, FA-microstructure mappings were independent of subject confounds, given that post-mortem tissue exhibits similar patterns of FA when pooling voxels from WM+GM (Sun et al., 2005; Sun et al., 2003; Roebroeck et al., 2019). A more in-depth analysis of these confounds (e.g., whether post-mortem interval or fixation time is the primary driver of between-subject variance) is left for future work within the larger multiple-subject dataset.

5.1. Evaluation of SAF pipeline

Our results show the pipeline's robustness to common IHC artefacts, while maintaining similar reproducibility to manually-derived SAF values (Figs. 3-5). This is especially in PLP slides, which typically displayed within-slide artefacts. Conversely, CD68 appears less reproducible for both manual and automated processing. CD68 stains sparse, small, activated microglia (soma diameter~10 µm (Kozlowski and

Weimer, 2012)). Here, the **diffstd** metric may be unsuitable as the assumption of low biological variability across slides may not be met. Inspection of the CD68 segmentation revealed that the automated pipeline correctly segmented more activated microglia relative to the manual method. The inclusion of more sparsely distributed microglia may add biological variability to the SAF, resulting in higher low- and high-pass **diffstd**. Staining gradients were not visibly present in CD68 slides, though some striping was visible on the manually-derived SAF maps. The automated pipeline reduced this striping, as reflected in the improved bandpass **diffstd**.

5.2. Voxelwise MRI-SAF analyses

Our correlation results agree with previous studies for myelin (Lazari and Lipp, 2021; Mancini et al., 2020; Seewann et al., 2009; Bagnato et al., 2018) and neurofilaments (Grussu et al., 2017). Activated microglia correlates positively with FA and negatively with MD/RD, similar to that reported in the human spinal cord (Grussu et al., 2017).
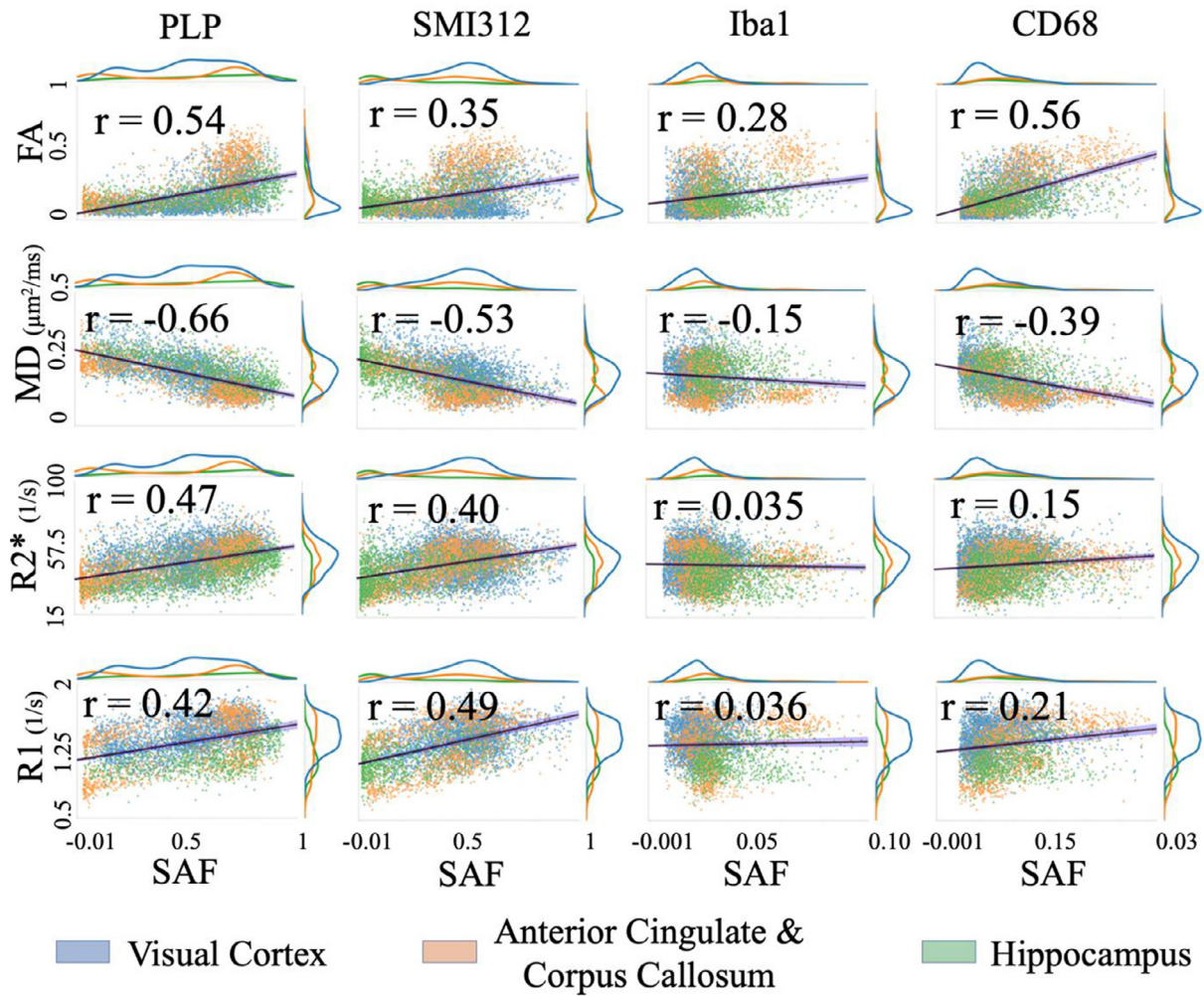
Fig. 7. Correlating MR parameters (DTI FA, MD, R2* and R1) with IHC SAF (WM+GM). The line of best fit (black line) and corresponding Pearson correlation coefficients, r, are overlaid. Confidence intervals (99%) are shown in blue shade. The visual cortex (blue), anterior cingulate (orange) and hippocampus (green) provide good dynamic range for the MR parameters and SAFs.

Table 1
Multiple linear regression predicting MR parameters using multiple IHC stains. Values are computed from WM+GM, or from WM only. A "subject" variable was also included to account for confounds such as post-mortem interval and age effects. As the predictors are unitless, all offsets and regression slopes are given in units of MR parameters. A: The regression coefficients and correlation coefficients r_{fit}. B: The relative importance of each stain describes the amount of variance it can explain in an MR parameter, averaged over all permutations of multiple regressions that include the specific stain. Values are normalised across stains to get a unit percentage.

|	All voxels/WM only (A) Multiple regression offset and slopes						
	Offset	PLP	SMI312	Iba1	CD68	Subject	r_{fit}
FA	−0.00142*/0.255	0.173/−0.173	0.0429/0.144	0.252/0.410*	10.5/8.81	−0.0112/−0.0247	0.645/0.353
MD [μm²/ms]	0.286/0.118	−0.112/0.0610	−0.070/−0.0323	−0.0629*/−0.0531*	−2.55/−1.40	0.0366/0.0310	0.745/0.429
R2* [1/s]	39.2/61.2	16.5/−2.22*	9.04/−0.889*	−61.2/−55.2	80.7/19.6*	−3.73/−5.54	0.536/0.455
R1 [1/s]	1.04/1.52	0.220/−0.217	0.382/0.0878	−2.48/−2.71	1.34/1.89	0.234/0.268	0.710/0.637
	(B) Relative Importance [%]						
	PLP	SMI312	Iba1	CD68		Subject	
FA	36.0/4.32	11.9/8.75	7.04/15.1	44.5/69.2		0.55/2.66	
MD	45.0/6.43	25.7/2.40	1.69/4.99	14.5/16.2		13.1/70.0	
R2*	51.3/0.47	30.1/1.97	2.90/22.6	3.86/3.62		11.8/71.3	
R1	18.7/4.54	30.6/2.39	3.15/5.88	2.93/0.443		44.6/86.8	

*not significant ($p>0.05$)
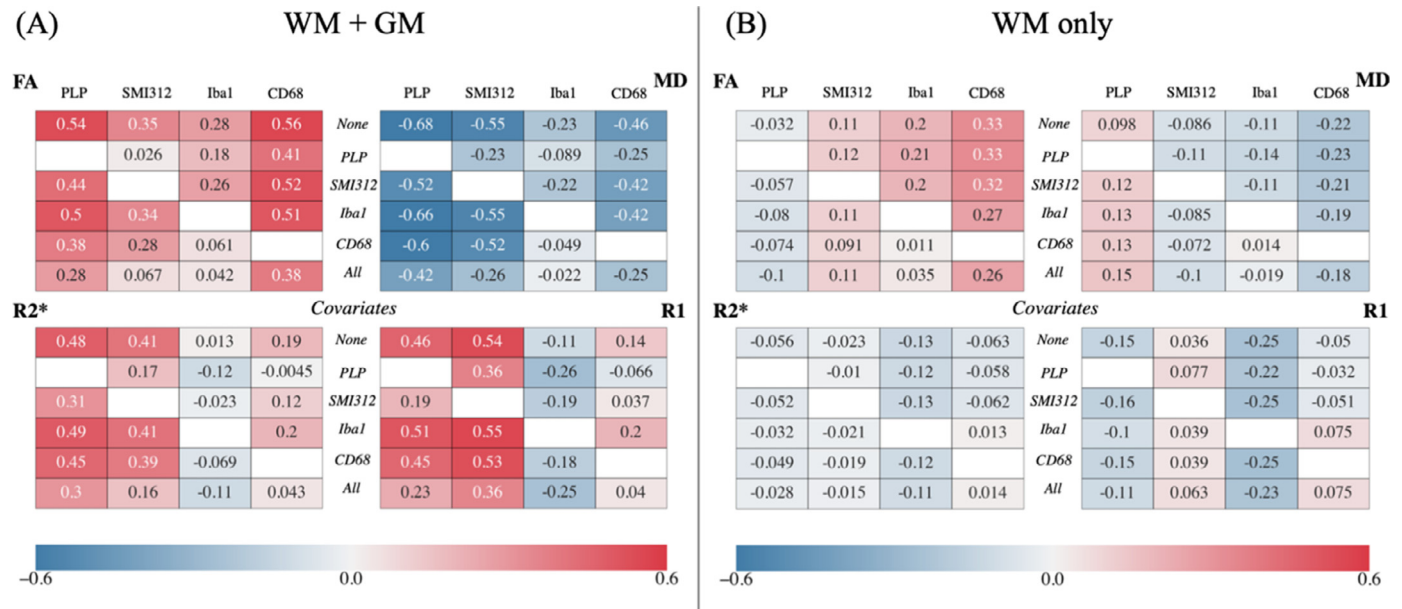
Fig. 8. Partial correlation analysis between MR parameters (FA, MD, R2* and R1) and IHC stains for (A) both WM+GM voxels and (B) for WM only. In each quadrant for either (A) or (B), the top row ("None") gives the correlation coefficients when accounting for subject ID only. The middle rows give the partial correlation coefficient controlling for one of the other stains and subject ID (italicised labels in the second to fifth rows). The bottom row ("All") gives the partial correlation coefficient controlling for all stains and subject ID. Note that correlations between FA/MD with CD68 remain relatively high, even after accounting for all other stains. *PLP (myelin); SMI312 (neurofilaments); Iba1 (microglia); CD68 (activated microglia).*

5.2.1. Myelin and neurofilaments

Our partial correlations for PLP (myelin) and SMI312 (neurofilaments) demonstrate pitfalls in performing pairwise MRI-histology correlations without controlling for other tissue features. Our simple correlations from WM+GM voxels reveal an apparent relationship between neurofilaments and FA, MD, R2*; however, our partial correlation results that account for myelin find no such relationship. Thus, one may erroneously conclude from simple correlations that anisotropy is affected by myelin and neurofilament (axons, dendrites) load, whereas our partial correlations show that anisotropy is primarily related to myelin density, with no unique variance attributed to dendritic load. As we do not account for microstructure orientation dispersion—where the orientational coherence of axons and dendrites may drive variations in diffusion anisotropy (Beaulieu, 2002)—we cannot determine whether the FA-PLP correlation reflects a direct relationship with myelin density or an indirect relationship with microstructure dispersion. These results imply that the portion of microstructure dispersion that affects FA is unrelated to dendritic load; if it were, we expect a high correlation between FA and dendritic load, with microstructure dispersion being the main confound in driving this relationship. This is something we hope to investigate further in the future. MD and R2* are similarly affected by myelin load, though some specificity to unmyelinated axons and/or dendrites remains.

Conversely, SMI312 explains unique variance in R1, even after accounting for PLP. This implies that R1 is sensitive to neurites in general, rather than myelin alone. This is noteworthy, given the studies correlating myelin with R1 (Stüber et al., 2014; Hakkarainen et al., 2016; Warntjes et al., 2017) without considering the impact of unmyelinated axons and/or dendrites.

Finally, in WM only, all PLP/SMI312 associations are reduced, suggesting WM/GM contrast as the primary driver. Taken together, our results may explain the wide range of MRI-myelin correlation coefficients reported in two reviews (Figs. 4, 5 in (Lazari and Lipp, 2021); Fig. 4 in (Mancini et al., 2020)).

5.2.2. Microglia

Our results suggest a relationship between CD68 (activated microglia, a biomarker for neuroinflammation (Bachiller et al., 2018; Geloso et al., 2017)) and FA/MD/RD, after accounting for all other stains (e.g., demyelination, axonal loss or general microglia infiltration). This relationship is present when considering all voxels (WM+GM) and WM only, demonstrating MR sensitivity to subtle microstructural changes rather than gross WM/GM differences. This result may be linked to microstructural changes in response to neuroinflammation that are unaccounted for in our histology metrics, or non-trivial changes in "activated" versus "non-activated" microglia morphology. For example, activated microglia often have more retracted, thicker processes and different soma sizes and shapes compared to their non-activated counterparts (Davis et al., 2017). This may contribute to increased restriction and/or hinderance of diffusion in the extra-axonal space, leading to an observed reduction in RD and a resultant increase in diffusion anisotropy (FA). While some report decreased FA in diseases associated with neuroinflammation (Chiang et al., 2017; Samara et al., 2020), they do not measure microglial activation and generally attribute the changes in FA to WM damage. Future research into measuring microglial activation using MRI will benefit from the development of advanced biophysical models (Taquet et al., 2019; Garcia-Hernandez et al., 2022; Zhang et al., 2012; Rafipoor et al., 2022) that can simultaneously account for multiple changes in the microstructure (demyelination, axonal damage, microglial infiltration, exchange effects etc.) and provide superior specificity to the DTI metrics.

Activated microglia is known to colocalise with iron (Kwan et al., 2012; Kaunzner et al., 2019) and correlate positively with R1/R2* (Wiggermann et al., 2017; Hametner et al., 2018; Roebroeck et al., 2019; Langkammer et al., 2010). However, CD68 is found to weakly correlate with R1/R2* after accounting for subject-specific confounds (Fig. 8, Table 1). Our results also suggest a negative correlation between non-activated microglia (Iba1 after accounting for CD68) and R1/R2*. Future studies may incorporate iron-stained IHC to elucidate the contributions of iron to relaxometry measures.

5.3. Limitations

There are limitations to this study. First, we use SAF—a semi-quantitative metric defined via a segmentation threshold—to quantify MRI-histology relationships. Depending on how the SAF threshold is derived, this may result in subtly different SAF maps. Further, SAF may not scale linearly with protein density, as anything above the stain threshold is considered positive stain. While our results show key trends between MR parameters and proteins, the exact MRI-SAF slopes may not be directly comparable to other studies using different thresholding methods or stain densities. Here, a key advantage of our data-driven SAF is that it is less sensitive to non-biological staining variations both within and between slides, which if unaccounted for, can bias stain density metrics. Third, we do not incorporate other histological metrics that may relate to MRI. For example, we do not include the fibre orientation dispersion and microglial cell morphology, which may explain variance in FA or other MR parameters. Finally, our pipeline does not address smaller-scale within-slide artefacts, such as artefactual staining of vasculature and folding artefacts at tissue edges (Suvarna et al., 2019). These might bias estimated SAF. Since our analysis occurs at the slide-level, manually excluding all smaller-scale artefacts would be very time-intensive. Following a qualitative check of the slides, we observed the staining of vasculature and folding artefacts to be relatively sparse, with a tendency to appear on the tissue edges. We do not anticipate the effect size of these artefacts to be significant.

6. Conclusion

We introduce an automated pipeline that extracts SAF maps from IHC slides. By design, the pipeline is generalisable to multiple IHC stains and found to be robust to artefacts from tissue staining and/or slide digitisation. Pipeline-derived SAF values were found to be as reproducible as expert-derived SAF estimates. The pipeline was applied to co-registered MRI and IHC data from post-mortem human brains, where we perform multiple linear analyses (simple correlation, partial correlation and multiple linear regression) to investigate voxelwise relationships between MRI- and IHC-derived parameters. Our results emphasise the need to simultaneously analyse multiple stains when validating MRI, as to avoid misleading inference due to the spatial covariance of multiple microstructural features. Interestingly, we found several diffusion-weighted metrics' sensitivity to activated microglia–a biomarker of neuroinflammation. This result held after accounting for spatial covariance of other stains (myelin, neurofilaments and non-activated microglia), suggesting MR specificity to aspects of neuroinflammation related to microglial activation, irrespective of other microstructural changes (axonal loss, demyelination or general microglia infiltration). Finally, subject ID was found to be the strongest predictor of some MR parameters, highlighting the need to consider tissue processing confounds when comparing post-mortem MRI across subjects. Together, our results demonstrate the pipeline as a valuable tool for IHC analyses and future investigations relating MRI to disease pathology.

Authors' contributions

DZLK conceived the study design, developed the histology processing pipeline, implemented the MRI-histology workflow, conducted data analyses, and drafted the manuscript. **SJ** conceived the study design, contributed to the design of the MRI-histology workflow and data analyses. **INH** developed the TIRL registration platform, performed the MRI-histology co-registration, designed the acquisition of the reproducibility dataset, and advised on data analyses. **JM** contributed to the histology processing pipeline. **BCT** analysed the diffusion MRI and T1 data. **SF** developed the MRI protocols and acquired the MRI data. **CW** analysed the R2* data. **CS** contributed to the acquisition of the multiple-region dataset. **AS** performed the histological processing, curated histological data, and provided expert feedback on histology analysis. **OA** designed

the brain sampling strategy and provided expert feedback on histological analysis. **MPG** optimised histological protocols, designed the brain sampling strategy, performed histology analysis, evaluated histological segmentations, and provided expert feedback on histological analysis. **KLM** conceived the study design, designed the post-mortem MRI protocol, contributed to the MRI-histology workflow, contributed to all data analysis, and helped draft the manuscript. **AFDH** conceived the study design, contributed to the MRI-histology workflow, contributed to all data analysis, and helped draft the manuscript.

Declaration of Competing Interest

None.

Data availability

The code for this work is publicly available at https://git.fmrib.ox.ac.uk/spet4877/ihcpy. The data will soon be available on https://open.win.ox.ac.uk/DigitalBrainBank/.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2022.119726.

References

Johnson, K.A., Fox, N.C., Sperling, R.A., Klunk, W.E., 2012. Brain Imaging in Alzheimer Disease. Cold Spring Harb. Perspect. Med. 2 (4), a006213. doi:10.1101/cshperspect.a006213.

Geraldes, R., Ciccarelli, O., Barkhof, F., et al., 2018. The current role of MRI in differentiating multiple sclerosis from its imaging mimics. Nat. Rev. Neurol. 14 (4), 199–213. doi:10.1038/nrneurol.2018.14.

Grolez, G., Moreau, C., Danel-Brunaud, V., et al., 2016. The value of magnetic resonance imaging as a biomarker for amyotrophic lateral sclerosis: a systematic review. BMC Neurol. 16 (1), 155. doi:10.1186/s12883-016-0672-6.

Nave, K.A., Werner, H.B., 2014. Myelination of the Nervous System: mechanisms and Functions. Annu. Rev. Cell Dev. Biol. 30 (1), 503–533. doi:10.1146/annurev-cell-bio-100913-013101.

Barker, R., Wellington, D., Esiri, M.M., Love, S., 2013. Assessing white matter ischemic damage in dementia patients by measurement of myelin proteins. J. Cereb. Blood Flow Metab. 33 (7), 1050–1057. doi:10.1038/jcbfm.2013.46.

Ulfig, N., Nickel, J., Bohl, J., 1998. Monoclonal antibodies SMI 311 and SMI 312 as tools to investigate the maturation of nerve cells and axonal patterns in human fetal brain. Cell Tissue Res. 291 (3), 433–443. doi:10.1007/s004410051013.

Atik, A., Cheong, J., Harding, R., Rees, S., De Matteo, R., Tolcos, M., 2014. Impact of daily high-dose caffeine exposure on developing white matter of the immature ovine brain. Pediatr. Res. 76 (1), 54–63. doi:10.1038/pr.2014.55.

Schirmer, L., Antel, J.P., Brück, W., Stadelmann, C., 2011. Axonal loss and neurofilament phosphorylation changes accompany lesion development and clinical progression in multiple sclerosis. Brain Pathol. 21 (4), 428–440. doi:10.1111/j.1750-3639.2010.00466.x.

Korzhevskii, D.E., Kirik, O.V., 2016. Brain Microglia and Microglial Markers. Neurosci. Behav. Physiol. 46 (3), 284–290. doi:10.1007/s11055-016-0231-z.

Jurga, A.M., Paleczna, M., Kuter, K.Z., 2020. Overview of General and Discriminating Markers of Differential Microglia Phenotypes. Front. Cell Neurosci. 14. Accessed August 30, 2022 https://www.frontiersin.org/articles/10.3389/fncel.2020.00198.

Bagnato, F., Hametner, S., Yao, B., et al., 2011. Tracking iron in multiple sclerosis: a combined imaging and histopathological study at 7 Tesla. Brain 134 (12), 3599–3612. doi:10.1093/brain/awr278.

Waller, R., Baxter, L., Fillingham, D.J., et al., 2019. Iba-1-/CD68+ microglia are a prominent feature of age-associated deep subcortical white matter lesions. PLoS One 14 (1), e0210888. doi:10.1371/journal.pone.0210888.

Bachiller, S., Jiménez-Ferrer, I., Paulus, A., et al., 2018. Microglia in Neurological Diseases: a Road Map to Brain-Disease Dependent-Inflammatory Response. Front. Cell Neurosci. 12, 488. doi:10.3389/fncel.2018.00488.

Geloso, M.C., Corvino, V., Marchese, E., Serrano, A., Michetti, F., D'Ambrosi, N, 2017. The Dual Role of Microglia in ALS: mechanisms and Therapeutic Approaches. Front. Aging Neurosci. 9, 242. doi:10.3389/fnagi.2017.00242.

Pallebage-Gamarallage, M., Foxley, S., Menke, R.A.L., et al., 2018. Dissecting the pathobiology of altered MRI signal in amyotrophic lateral sclerosis: a post mortem whole brain sampling strategy for the integration of ultra-high-field MRI and quantitative neuropathology. BMC Neurosci. 19 (1), 11. doi:10.1186/s12868-018-0416-1.

Tendler, B.C., Hanayik, T., Ansorge, O., et al., 2022. The Digital Brain Bank, an open access platform for post-mortem imaging datasets. Elife 11, e73153. doi:10.7554/eLife.73153.

Goubran, M., Hammond, R.R., de Ribaupierre, S., et al., 2015. Magnetic resonance imaging and histology correlation in the neocortex in temporal lobe epilepsy. Ann. Neurol. 77 (2), 237–250. doi:10.1002/ana.24318.

Fjær, S., Bø, L., Lundervold, A., et al., 2013. Deep Gray Matter Demyelination Detected by Magnetization Transfer Ratio in the Cuprizone Model. PLoS One 8 (12), e84162. doi:10.1371/journal.pone.0084162.

Fjær, S., Bø, L., Myhr, K.M., Torkildsen, Ø., Wergeland, S., 2015. Magnetization transfer ratio does not correlate to myelin content in the brain in the MOG-EAE mouse model. Neurochem. Int. 83-84, 28–40. doi:10.1016/j.neuint.2015.02.006.

Dusek, P., Bahn, E., Litwin, T., et al., 2017. Brain iron accumulation in Wilson disease: a post mortem 7 Tesla MRI - histopathological study. Neuropathol. Appl. Neurobiol. 43 (6), 514–532. doi:10.1111/nan.12341.

Yano, R., Hata, J., Abe, Y., et al., 2018. Quantitative temporal changes in DTI values coupled with histological properties in cuprizone-induced demyelination and remyelination. Neurochem. Int. 119, 151–158. doi:10.1016/j.neuint.2017.10.004.

Wang, C., Foxley, S., Ansorge, O., et al., 2020. Methods for quantitative susceptibility and R2* mapping in whole post-mortem brains at 7T applied to amyotrophic lateral sclerosis. Neuroimage 222, 117216. doi:10.1016/j.neuroimage.2020.117216.

Abe, Y., Komaki, Y., Seki, F., Shibata, S., Okano, H., Tanaka, K.F., 2019. Correlative study using structural MRI and super-resolution microscopy to detect structural alterations induced by long-term optogenetic stimulation of striatal medium spiny neurons. Neurochem. Int. 125, 163–174. doi:10.1016/j.neuint.2019.02.017.

Mollink, J., Hiemstra, M., Miller, K.L., et al., 2019. White matter changes in the perforant path area in patients with amyotrophic lateral sclerosis. Neuropathol. Appl. Neurobiol. 45 (6), 570–585. doi:10.1111/nan.12555.

Lazari, A., Lipp, I., 2021. Can MRI measure myelin? Systematic review, qualitative assessment, and meta-analysis of studies validating microstructural imaging with myelin histology. Neuroimage 230, 117744. doi:10.1016/j.neuroimage.2021.117744.

van der Weijden, C.W.J., García, D.V., Borra, R.J.H., et al., 2021. Myelin quantification with MRI: a systematic review of accuracy and reproducibility. Neuroimage 226, 117561. doi:10.1016/j.neuroimage.2020.117561.

Mancini M., Karakuzu A., Cohen-Adad J., Cercignani M., Nichols T.E., Stikov N. An Interactive Meta-Analysis of MRI Biomarkers of Myelin. Jbabdi S, Baker CI, Jbabdi S, Does M, eds. eLife. 2020;9:e61523. doi:10.7554/eLife.61523

De Barros, A., Arribarat, G., Combis, J., Chaynes, P., Péran, P., 2019. Matching ex vivo MRI With Iron Histology: pearls and Pitfalls. Front. Neuroanat. 13. Accessed January 18, 2022 https://www.frontiersin.org/article/10.3389/fnana.2019.00068.

Seewann, A., Vrenken, H., van der Valk, P., et al., 2009. Diffusely Abnormal White Matter in Chronic Multiple Sclerosis: imaging and Histopathologic Analysis. Arch. Neurol. 66 (5), 601–609. doi:10.1001/archneurol.2009.57.

Bulk, M., Kenkhuis, B., van der Graaf, L.M., Goeman, J.J., Natté, R., van der Weerd, L., 2018. Postmortem T2*- Weighted MRI Imaging of Cortical Iron Reflects Severity of Alzheimer's Disease. J. Alzheimers Dis. 65 (4), 1125–1137. doi:10.3233/JAD-180317.

Wiggermann, V., Hametner, S., Hernández-Torres, E., et al., 2017. Susceptibility-sensitive MRI of multiple sclerosis lesions and the impact of normal-appearing white matter changes. NMR Biomed. 30 (8). doi:10.1002/nbm.3727.

Hametner, S., Endmayr, V., Deistung, A., et al., 2018. The influence of brain iron and myelin on magnetic susceptibility and effective transverse relaxation - A biochemical and histological validation study. Neuroimage 179, 117–133. doi:10.1016/j.neuroimage.2018.06.007.

Bagnato, F., Hametner, S., Boyd, E., et al., 2018. Untangling the R2* contrast in multiple sclerosis: a combined MRI-histology study at 7.0 Tesla. PLoS One 13 (3), e0193839. doi:10.1371/journal.pone.0193839.

van der Loos, C.M., 2008. Multiple immunoenzyme staining: methods and visualizations for the observation with spectral imaging. J. Histochem. Cytochem. 56 (4), 313–328. doi:10.1369/jhc.2007.950170.

Suvarna, S.K., Layton, C., Bancroft, J.D., 2019. Bancroft's Theory and Practice of Histological Techniques, 8th edition Elsevier, Oxford.

Farahani, N., Parwani, A.V., Pantanowitz, L., 2015. Whole slide imaging in pathology: advantages, limitations, and emerging perspectives. PLMI 7, 23–33. doi:10.2147/PLMI.S59826.

Gurcan, M.N., Boucheron, L., Can, A., Madabhushi, A., Rajpoot, N., Yener, B., 2009. Histopathological Image Analysis: a Review. IEEE Rev. Biomed. Eng. 2, 147–171. doi:10.1109/RBME.2009.2034865.

McCann, M.T., Ozolek, J.A., Castro, C.A., Parvin, B., Kovacevic, J., 2015. Automated Histology Analysis: opportunities for signal processing. IEEE Signal Process. Mag. 32 (1), 78–87. doi:10.1109/MSP.2014.2346443.

Madabhushi, A., Lee, G., 2016. Image analysis and machine learning in digital pathology: challenges and opportunities. Med. Image Anal. 33, 170–175. doi:10.1016/j.media.2016.06.037.

Bussolati, G., Radulescu, R.T., 2011. Blocking Endogenous Peroxidases in Immunohistochemistry: a Mandatory. Yet Also Subtle Measure. Appl. Immunohistochem. Mol. Morphol. 19 (5), 484. doi:10.1097/PAI.0b013e318219a6e6.

Del Cerro, M., Cogen, J.P., Del Cerro, C., 1981. Retrospective demonstration of endogenous peroxidase activity in plastic-embedded tissues conventionally prepared for electron microscopy. J. Histochem. Cytochem. 29 (7), 874–876. doi:10.1177/29.7.7021672.

Tendler, B.C., Foxley, S., Hernandez-Fernandez, M., et al., 2020. Use of multi-flip angle measurements to account for transmit inhomogeneity and non-Gaussian diffusion in DW-SSFP. Neuroimage 220, 117113. doi:10.1016/j.neuroimage.2020.117113.

Basser, P.J., Mattiello, J., LeBihan, D., 1994. MR diffusion tensor spectroscopy and imaging. Biophys. J. 66 (1), 259–267. doi:10.1016/S0006-3495(94)80775-1.

Huszar I.N., Pallebage-Gamarallage M., Foxley S., et al. Tensor Image Registration Library: automated Non-Linear Registration of Sparsely Sampled Histological Specimens to Post-Mortem MRI of the Whole Human Brain.; 2019:849570. doi:10.1101/849570

Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage 17 (2), 825–841. doi:10.1016/s1053-8119(02)91132-8.

Zhang, Y., Brady, M., Smith, S., 2001. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Trans. Med. Imaging 20 (1), 45–57. doi:10.1109/42.906424.

Ruifrok, A.C., Johnston, D.A., 2001. Quantification of histochemical staining by color deconvolution. Anal. Quant. Cytol. Histol. 23 (4), 291–299.

Landini, G., Martinelli, G., Piccinini, F., 2021. Colour deconvolution: stain unmixing in histological imaging. Bioinformatics 37 (10), 1485–1487. doi:10.1093/bioinformatics/btaa847.

Clarke, E.L., Treanor, D., 2017. Colour in digital pathology: a review. Histopathology 70 (2), 153–163. doi:10.1111/his.13079.

Geijs, D.J., Intezar, M., der, Laak JAWM van, Litjens, G.J.S, 2018. Automatic color unmixing of IHC stained whole slide images. In: Medical Imaging 2018: Digital Pathology, 10581. SPIE, pp. 165–171. doi:10.1117/12.2293734.

Otsu, N., 1979. A Threshold Selection Method from Gray-Level Histograms. IEEE Trans. Syst. Man Cybern. 9 (1), 62–66. doi:10.1109/TSMC.1979.4310076.

Yuan, X., Wu, L., Peng, Q., 2015. An improved Otsu method using the weighted object variance for defect detection. Appl. Surf. Sci. 349, 472–484. doi:10.1016/j.apsusc.2015.05.033.

Butterworth, S., 1930. On the Theory of Filter Amplifiers. Exp. Wirel. Wirel. Engineer 7, 536–541.

Virtanen, P., Gommers, R., Oliphant, T.E., et al., 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat. Methods 17 (3), 261–272. doi:10.1038/s41592-019-0686-2.

Huber, P.J., 2011. Robust Statistics. In: Lovric, M (Ed.), International Encyclopedia of Statistical Science. Springer, pp. 1248–1251. doi:10.1007/978-3-642-04898-2_594.

Grömping, U., 2015. Variable importance in regression models. Wiley Interdiscip. Rev. Comput. Stat. 7 (2), 137–152. doi:10.1002/wics.1346.

Grussu, F., Schneider, T., Tur, C., et al., 2017. Neurite dispersion: a new marker of multiple sclerosis spinal cord pathology? Ann. Clin. Transl. Neurol. 4 (9), 663–679. doi:10.1002/acn3.445.

Khodanovich, M., Pishchelko, A., Glazacheva, V., et al., 2019. Quantitative Imaging of White and Gray Matter Remyelination in the Cuprizone Demyelination Model Using the Macromolecular Proton Fraction. Cells 8 (10), 1204. doi:10.3390/cells8101204.

Praet, J., Manyakov, N.V., Muchene, L., et al., 2018. Diffusion kurtosis imaging allows the early detection and longitudinal follow-up of amyloid-β-induced pathology. Alzheimer's Res. Ther. 10 (1), 1. doi:10.1186/s13195-017-0329-8.

Peters, J.M., Struyven, R.R., Prohl, A.K., et al., 2019. White matter mean diffusivity correlates with myelination in tuberous sclerosis complex. Ann. Clin. Transl. Neurol. 6 (7), 1178–1190. doi:10.1002/acn3.793.

Shatil, A.S., Uddin, M.N., Matsuda, K.M., Figley, C.R., 2018. Quantitative Ex Vivo MRI Changes due to Progressive Formalin Fixation in Whole Human Brain Specimens: longitudinal Characterization of Diffusion, Relaxometry, and Myelin Water Fraction Measurements at 3T. Front. Med. (Lausanne) 5. Accessed January 19, 2022 https://www.frontiersin.org/article/10.3389/fmed.2018.00031.

McNab, J.A., Jbabdi, S., Deoni, S.C.L., Douaud, G., Behrens, T.E.J., Miller, K.L., 2009. High resolution diffusion-weighted imaging in fixed human brain using diffusion-weighted steady state free precession. Neuroimage 46 (3), 775–785. doi:10.1016/j.neuroimage.2009.01.008.

Foxley, S., Jbabdi, S., Clare, S., et al., 2014. Improving diffusion-weighted imaging of post-mortem human brains: SSFP at 7 T. Neuroimage 102 (2), 579–589. doi:10.1016/j.neuroimage.2014.08.014, Pt.

Sun, S.W., Neil, J.J., Liang, H.F., et al., 2005. Formalin fixation alters water diffusion coefficient magnitude but not anisotropy in infarcted brain. Magn. Reson. Med. 53 (6), 1447–1451. doi:10.1002/mrm.20488.

Sun, S.W., Neil, J.J., Song, S.K., 2003. Relative indices of water diffusion anisotropy are equivalent in live and formalin-fixed mouse brains. Magn. Reson. Med. 50 (4), 743–748. doi:10.1002/mrm.10605.

Roebroeck, A., Miller, K.L., Aggarwal, M., 2019. Ex vivo diffusion MRI of the human brain: technical challenges and recent advances. NMR Biomed. 32 (4), e3941. doi:10.1002/nbm.3941.

Kozlowski, C., Weimer, R.M., 2012. An Automated Method to Quantify Microglia Morphology and Application to Monitor Activation State Longitudinally In Vivo. PLoS One 7 (2), e31814. doi:10.1371/journal.pone.0031814.

Beaulieu, C., 2002. The basis of anisotropic water diffusion in the nervous system – a technical review. NMR Biomed. 15 (7–8), 435–455. doi:10.1002/nbm.782.

Stüber, C., Morawski, M., Schäfer, A., et al., 2014. Myelin and iron concentration in the human brain: a quantitative study of MRI contrast. Neuroimage 93 (1), 95–106. doi:10.1016/j.neuroimage.2014.02.026, Pt.

Hakkarainen, H., Sierra, A., Mangia, S., et al., 2016. MRI relaxation in the presence of fictitious fields correlates with myelin content in normal rat brain. Magn. Reson. Med. 75 (1), 161–168. doi:10.1002/mrm.25590.

Warntjes, J.B.M., Persson, A., Berge, J., Zech, W., 2017. Myelin Detection Using Rapid Quantitative MR Imaging Correlated to Macroscopically Registered Luxol Fast Blue–Stained Brain Specimens. AJNR Am. J. Neuroradiol. 38 (6), 1096–1102. doi:10.3174/ajnr.A5168.

Davis, B.M., Salinas-Navarro, M., Cordeiro, M.F., Moons, L., De Groef, L., 2017. Characterizing microglia activation: a spatial statistics approach to maximize information extraction. Sci. Rep. 7 (1), 1576. doi:10.1038/s41598-017-01747-8.

Chiang, P.L., Chen, H.L., Lu, C.H., et al., 2017. White matter damage and systemic inflammation in Parkinson's disease. BMC Neurosci. 18 (1), 48. doi:10.1186/s12868-017-0367-y.

Samara, A., Murphy, T., Strain, J., et al., 2020. Neuroinflammation and White Matter Alterations in Obesity Assessed by Diffusion Basis Spectrum Imaging. Front. Hum. Neurosci. 13. https://www.frontiersin.org/article/10.3389/fnhum.2019.00464.

Taquet, M., Jankovski, A., Rensonnet, G., et al., 2019. Extra-axonal restricted diffusion as an in-vivo marker of reactive microglia. Sci. Rep. 9, 13874. doi:10.1038/s41598-019-50432-5.

Garcia-Hernandez, R., Cerdán Cerdá, A., Trouve Carpena, A., et al., 2022. Mapping microglia and astrocyte activation in vivo using diffusion MRI. Sci. Adv. 8 (21), eabq2923. doi:10.1126/sciadv.abq2923.

Zhang, H., Schneider, T., Wheeler-Kingshott, C.A., DC, Alexander, 2012. NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain. Neuroimage 61 (4), 1000–1016. doi:10.1016/j.neuroimage.2012.03.072.

Rafipoor, H., Zheng, Y.Q., Griffanti, L., Jbabdi, S., Cottaar, M., 2022. Identifying microstructural changes in diffusion MRI; How to circumvent parameter degeneracy. Neuroimage 260, 119452. doi:10.1016/j.neuroimage.2022.119452.

Kwan, J.Y., Jeong, S.Y., Van Gelderen, P., et al., 2012. Iron accumulation in deep cortical layers accounts for MRI signal abnormalities in ALS: correlating 7 tesla MRI and pathology. PLoS One 7 (4), e35241. doi:10.1371/journal.pone.0035241.

Kaunzner, U.W., Kang, Y., Zhang, S., et al., 2019. Quantitative susceptibility mapping identifies inflammation in a subset of chronic multiple sclerosis lesions. Brain 142 (1), 133–145. doi:10.1093/brain/awy296.

Langkammer, C., Krebs, N., Goessler, W., et al., 2010. Quantitative MR Imaging of Brain Iron: a Postmortem Validation Study. Radiology 257 (2), 455–462. doi:10.1148/radiol.10100495.