








Original Research Articles

Assessing optimisation of 3T MRI protocol across multi-site, multi-vendor settings for studying mild Traumatic Brain Injury: mTBI Predict Consortium study

Pradeepa Ruwan Wanni Arachchige, PhD^{1,2} , Yidian Gao, PhD^{3,4}, Iman Idrees, PhD⁵ , Jessikah Fildes, BSc¹, Aliza Finch, BSc^{3,4}, Waheeda Hawa, BSc⁵, Martin Craig, PhD¹, Matthew J. Brookes, PhD¹, Lisa J. Hill, PhD^{6,7,8} , Samuel J.E. Lucas, PhD^{3,9} , James L. Mitchell, PhD¹⁰ , Alexandra J. Sinclair, PhD^{6,7,8} , Jan Novak, PhD⁵ , Andrew P. Bagshaw, PhD^{3,4} , Martin Wilson, PhD^{3,4} , Davinia Fernandez-Espejo, PhD^{3,4} , Karen J. Mullinger, PhD^{1,3,4,11} ^a, United Kingdom mTBI Predict Consortium 

¹ Sir Peter Mansfield Imaging Centre, School of Physics and Astronomy, University of Nottingham, Nottingham, United Kingdom, ² Medical Imaging, Department of Health and Care Professions, Faculty of Health and Life Sciences, University of Exeter, Exeter, United Kingdom, ³ Centre for Human Brain Health, University of Birmingham, Birmingham, United Kingdom, ⁴ School of Psychology, University of Birmingham, Birmingham, United Kingdom, ⁵ Institute of Health and Neurodevelopment, Aston University, Birmingham, United Kingdom, ⁶ Translational Brain Science, College of Medicine and Health, University of Birmingham, Birmingham, United Kingdom, ⁷ Biomedical Sciences, School of Infection, Inflammation and Immunology, College of Medicine and Health, University of Birmingham, Birmingham, United Kingdom, ⁸ NIHR Birmingham Biomedical Research Centre, University Hospitals Birmingham, Birmingham, United Kingdom, ⁹ School of Sport, Exercise and Rehabilitation Sciences, University of Birmingham, Birmingham, United Kingdom, ¹⁰ Translational Brain Science, College of Medicine and Health, University of Nottingham, Nottingham, United Kingdom, ¹¹ NIHR Nottingham Biomedical Research Centre, University of Nottingham, Nottingham, United Kingdom

Keywords: T1 anatomical, neuroimaging, diffusion weighted imaging, echo planar imaging, arterial spin labelling

<https://doi.org/10.52294/001c.157692>

Aperture Neuro

Vol. 6, 2026

Large multi-site, multi-vendor MRI studies offer opportunities to advance clinical practice and neuroscience research by increasing sample sizes, especially in clinical populations. However, variability between scanners or sites can affect the reproducibility of quantitative MRI metrics, which is vital for using MRI beyond visual inspection of structural images. This study aimed to assess the reproducibility of MRI metrics for the mild traumatic brain injury-Predict (mTBI-Predict) consortium, spanning Philips and Siemens scanners across three sites. We conducted a travelling head study (n=1), acquiring repeated measurements from a healthy participant across three sites. Evaluated metrics included cortical thickness and subcortical volume from T1-weighted images, fractional anisotropy (FA) and mean diffusivity (MD) from diffusion-weighted imaging (DWI), temporal signal-to-noise ratio (tSNR) from echo planar imaging, and cerebral blood flow (CBF) from arterial spin labelling (ASL). Image quality metrics (point spread function and motion) were also assessed across all DWI, functional magnetic resonance imaging (fMRI), and pseudo-continuous arterial spin labelling (pCASL) data. Standardised pipelines were used for preprocessing and analysis. Within-subject coefficients of variation (wCV%) and linear regressions assessed repeatability and consistency. Structural measures, including cortical thickness, subcortical volume, FA, and MD, demonstrated excellent reproducibility (wCV $\leq 5\%$, $R^2 \geq 0.85$) within and between-sites. Minimal differences in participant motion were observed across sessions. However, as expected, tSNR and CBF showed greater variability than structural measures, due to their functional nature. Despite this, wCV% remained below 10% for all metrics, a threshold indicating very good agreement for biological measures. These findings align with previous work, demonstrating the feasibility of obtaining comparable MRI measurements across sites for studying brain structure and function. This study lays the groundwork for further evaluation of clinically relevant MRI biomarkers for mTBI.

^a Corresponding Author:
Karen Mullinger, PhD
karen.mullinger@nottingham.ac.uk

INTRODUCTION

Large-scale, multisite, multivendor, and multimodal longitudinal imaging studies are crucial in brain research for identifying imaging biomarkers across a wide range of brain health issues, including neurodegenerative disorders, mental health conditions, and brain injuries.^{1,2} Among these, various brain health challenges stand out as significant global health and societal concerns. Over the past few decades, large-scale imaging initiatives have played a pivotal role in fostering international collaboration in addressing important brain health challenges. Multisite studies enable the inclusion of larger, more diverse sample sizes, enhancing the reliability and generalisability of findings. By integrating multisite data, these studies provide a robust framework for identifying and validating potential imaging biomarkers with greater rigor and statistical power.^{2,3}

Despite the extensive global burden of mortality and morbidity associated with traumatic brain injury (TBI), most previous TBI studies face challenges due to relatively small sample sizes, which hinder scientific advancement and clinical translation because of the heterogeneity of symptoms and representativeness of included patients.⁴ Neuroimaging studies play a vital role in quantifying pathological changes in TBI that can progressively occur over a prolonged period. Mild TBI (mTBI) is the most common type of TBI, but many cases go undiagnosed because symptoms can be subtle and may not appear immediately. Indeed, no obvious changes on standard clinical MRI scans are one of the diagnostic criteria for mTBI.⁵ This underdiagnosis can lead to untreated symptoms, prolonged suffering, and delayed return to previous work or leisure activities.⁶ If undiagnosed or not fully recovered before patients return to work or leisure activities, they may be at greater risk if experiencing a second concussion, also known as Second Impact Syndrome, which can result in fatal outcomes.⁷ Additionally, if a concussion is not properly managed, the risk of developing psychiatric disorders increases over time due to ongoing disruption in brain regions responsible for mood, affect, and executive function.⁸ Among all neuroimaging tools, quantitative analysis of magnetic resonance imaging (MRI) data, moving beyond visual inspection of structural MRI images which are typically performed clinically, has significant potential to improve clinical assessments and guide the management and treatment of patients following mTBI.⁹⁻¹¹ Therefore, large-scale MRI studies are essential, for providing robust, quantitative data needed to improve our understanding of mTBI, facilitate the development of more accurate diagnostic and prognosticating tools, and ultimately enhance clinical outcomes.

Currently, several global consortia are addressing various issues related to TBI. These include: (i) Transforming Research and Clinical Knowledge in Traumatic Brain Injury (TRACK-TBI)¹⁰; (ii) and Enhancing NeuroImaging Genetics through Meta Analysis consortium¹²; (iii) The Collaborative European NeuroTrauma Effectiveness Research in Traumatic Brain Injury¹³; (iv) Concussion Assessment, Research and Education consortium¹⁴; (v) International initiative for TBI Research¹⁵; and (vi) the Chronic Effects of Neuro-

trauma Consortium.¹⁶ These consortia address different aspects of TBI research by integrating clinical, genetic and neuroimaging data to improve classification and outcome prediction. While all of them are collecting MRI data, there is notably, limited published research on the initial development of their MRI sequences, and most available studies focus on their applications after protocol standardisation. The TRACK-TBI study protocol for 3T MRI, implemented across General Electric, Philips, and Siemens scanners,¹⁰ was adapted from the Alzheimer's Disease Neuroimaging Initiative (ADNI)¹⁷; however, it did not include a dedicated optimisation framework. Similarly, the Chronic Effects of Neurotrauma Consortium was intentionally aligned with both ADNI and TRACK-TBI-2, employing relatively comparable 3T MRI acquisition parameters, though detailed protocol information was not provided.¹⁸ The Concussion Assessment, Research and Education consortium's MRI protocol, developed for Siemens Trio and General Electric scanners, conducted a stability analysis involving 30 and 35 non-contact sport control subjects, and included first two travelling human heads to evaluate site-to-site variation.¹⁹

Multisite and multivendor data acquisition in neuroimaging studies can present significant challenges, including increased variability among scan sessions, differences in MRI sequences across scanners and vendors, and motion biases inherent to different scanners.² These factors can impact the reliability and consistency of the scans, potentially leading to data that is less comparable across sites, which is detrimental to the identification of biomarkers.²⁰ To address these challenges, it is crucial to optimize MRI sequences to minimise between-site differences. This is often done using a "travelling head" paradigm where the same participant(s) are imaged on the different MRI scanners to be used in the main study.²¹ This approach is necessary because MRI sequences often differ inherently between scanner vendors.^{1,22} By employing "travelling head" participants, researchers can assess the impact of scanner differences on key metrics that inform future biomarkers. This enhances the reliability and consistency of the data collected across different sites by reducing the impact of scanner-specific biases.^{22,23}

The mTBI-Predict consortium, of which this work is part, aims to identify the most accurate, and reproducible biomarkers to better identify, in the acute stage (3 weeks) post-mTBI, those at risk of chronic health issues 6 months after the injury. The scale of the mTBI-Predict Consortium's aims, which includes the recruitment of 610 patients, necessitates a multisite approach, similar to other TBI consortia acquiring MRI data across multiple centres. After this optimisation study, we will conduct a variability study across 20 patients with mTBI and 20 control participants.

This study addresses the earliest, essential, stage of protocol development. Specifically, this work focusses on the optimisation of MRI acquisition across different scanner platforms and vendors. Our aims were: 1) To develop 3T MRI data acquisition protocol for imaging sequences harmonised across multiple sites that will be used to generate putative biomarkers for mTBI-Predict; 2) To assess the reproducibility of the chosen MRI sequences across different

Table 1. 3T MRI Scanners and hardware to be used for mTBI-Predict study. The network comprises 3 sites.

Site	Institute	Vendor	Scanner model	Bore size	Gradient performance	Installation date	Software version
1	Sir Peter Mansfield Imaging Centre, University of Nottingham	Philips	Ingenia 3T	70 cm	45 mTm ⁻¹ 200mTm ⁻¹ ms ⁻¹	Mar-2016	R5.3.0 R5.9.0
2	Centre for Human Brain Health, University of Birmingham	Siemens	MAGNETOM Prisma 3T	60 cm	XR Gradients 80 mTm ⁻¹ 200mTm ⁻¹ ms ⁻¹	Dec-2018	VE11C
3	Aston Institute of Health and Neurodevelopment, Aston University	Siemens	MAGNETOM Prisma 3T	60 cm	XR Gradients 80 mTm ⁻¹ 200mTm ⁻¹ ms ⁻¹	Mar-2022	Syngo MR XA30

scanners using a single travelling head before advancing to large-scale data collection. The MRI sequences targeted were: T1-weighted (T1W) MPRAGE, diffusion-weighted imaging (DWI), gradient-echo echo-planar imaging (EPI) to be used for functional MRI (fMRI) and arterial spin labelling (ASL). These sequences have all previously been shown to produce potential imaging biomarkers of mTBI.^{11, 13,24-28} The goal was to ensure the consistency and reliability of these imaging sequences, and of the derived metrics, across the three imaging sites that are part of the mTBI-Predict consortium, thereby establishing a foundation for high-quality data collection. The assessments of reproducibility carried out here, for example on the T1 anatomical and fMRI data, are not the final biomarkers that will be used in mTBI-Predict. Instead, we chose these metrics to provide an indication of the reproducibility and reliability of the underlying images within and between-sites. For the fMRI acquisition for example, this will allow use with multiple tasks.

METHODS

STUDY DESIGN AND IMAGE ACQUISITION

The study was approved by the local university ethical review boards at each institution (University of Birmingham STEM Ethics Committee, University of Nottingham Medical School Ethics Committee, Aston University Health and Life Sciences College Ethics Committee), and the participants provided written informed consent. We began by testing the sequences on a phantom to assess the basic parameters of the images. Prior to the main study, we independently scanned a small number of volunteers at each site to evaluate reproducibility, scan time, and image quality, and iteratively optimised parameters for fMRI acquisition (see below). These preliminary evaluations were not part of the final dataset but helped guide protocol refinement for optimisation for the travelling head scan. Finally, we conducted a travelling head study on a single, extremely experienced MRI participant (the first author, initials P.R.W.A.; female, age 35 years) who has over 7 years of experience working with MRI and has taken part in approximately 20 scanning

hours across multiple research studies. The optimised protocol was used across all sites.

The details of the scanners on which sequences were optimised are shown in [Table 1](#). The scanners were all 3T but varied in hardware features, including bore sizes, gradient strengths, and acceleration capabilities. In brief: Site 1, had a Philips Ingenia wide bore 3T MRI scanner, Sites 2 and 3 had Siemens Prisma narrow (60 cm) bore 3T MRI scanners on different software releases. At all sites, a whole body transmit coil, and 32 channel head receive coil were used. The sequences to be optimised were a T1W, DWI, fMRI-EPI, 2D multi-post label delay 2D-EPI readout pseudo continuous arterial spin labelling (pCASL).

The final sequence parameters to be tested on the travelling head are summarised in [Table 2](#). The initial sequence set-up was based on parameters used by UK Biobank²⁹ on a Siemens Skyra 3T MRI scanner for the T1, DWI and fMRI-EPI sequences. The T1W protocol on the Philips Ingenia required adjustments to repetition time (TR) and echo time (TE), due to vendor-specific sequence implementations, whilst little change was required for the Siemens Prisma ([Table 2](#)). For DWI, we reduced the multiband factor (MB) to two on Philips Ingenia scanner and used a SENSE factor of 1.5 to accommodate platform capabilities. As a result of these changes, combined with different gradient performance ([Table 1](#)), the TR and TE had to be altered to enable the resolution, field of view (FOV) and number of directions to stay consistent between scanners and match those parameters to the UK Biobank protocol.

For fMRI, the effect of a combination of different MB factors was assessed initially on a fBIRN phantom and 7 different participants (different people at different sites). For the Philips Ingenia scanner we assessed MB factors 2, 3 and 4 with SENSE factor set to 2. For the Siemens Prisma scanner we assessed MB factors 4, 5 and 8 with no GRAPPA. As a result of this initial work (data not shown), the optimised sequence employed MB factors of four (Siemens scanners) and two (Philips scanner), reduced from eight in the UK Biobank protocol. We used no GRAPPA (Siemens scanners) and SENSE factor 2 (Philips scanner). This parameter choice minimised artefacts, increased SNR and was in accordance with the scanner capabilities. Compared with the UK Biobank protocol, the FOV was enlarged across all

Table 2. Optimised structural, functional and physiological protocols.

Sequence	Site	Acquisition time (min:sec)	Resolution (mm)	Field of View (mm3)	TR (ms)	TE (ms)	Other
T1 (MPRAGE)	1	04:55	1.0x1.0x1.0	256x256x208	8.1	3.7	$\alpha = 8^\circ$, TFE
	2,3	04:54	1.0x1.0x1.0	256x256x208	2000	2.03	$\alpha = 8^\circ$, TI = 0.88 s
DWI-EPI	1	12:27	2.0x2.0x2.0	224x224x120	6369	97	$\alpha = 90^\circ$, MB=2, SENSE = 1.5 b=0; dir=5, b=1000; dir=50, b=2000; dir=50, blip reversed b = 0
	2,3	06:32	2.0x2.0x2.0	224x224x120	3600	92	$\alpha = 78^\circ$, MB=3, No GRAPPA b=0; dir=6, b=1000; dir=50, b=2000; dir=50, blip-reversed b = 0
fMRI-EPI	1	04:46	2.33x2.33x2.4	224x224x130	2000	40	$\alpha = 85^\circ$, MB=2, SENSE = 2, volumes per run=140
	2,3	04:48	2.3x2.3x2.4	224x224x144	2000	40	$\alpha = 85^\circ$, MB=4, No GRAPPA, volumes per run=138
pCASL-EPI	1	06:00	3.39x3.39x4.95	217x217x119	3100, 3400, 4400, 4300, 4800	14	PLD (s) = 0.2, 0.7, 1.2, 1.7, 2.2; volumes = 12, 12, 12, 20, 30 respectively to the PLD, Label duration = 1.4 s No MB, No SENSE
	2	05:27	3.4x3.4x5.0	217x217x119	3580	19	PLD (s) = 0.2, 0.7, 1.2, 1.7, 2.2; volumes = 12, 12, 12, 20, 30 respectively to the PLD Label duration = 1.4 s MB=1, No GRAPPA

Note: TR=repitition time, TE=echo time, TFE=Turbo Field Echo, TI=Inversion time, α = flip angle, MB= multiband factor, PLD=post label delay, SENSE =SENSitivity Encoding. GRAPPA = GeneRalized Autocalibrating Partially Parallel Acquisitions. The TR in Philips T1 MPRAGE appears shorter due to differences in acquisition strategy but is functionally similar to Siemens T1 in terms of image contrast and timing. In pCASL, Philips acquires separate images for each PLD (multiple TRs), whereas Siemens combines all PLDs into a single image (one TR). See also Table S2 for further DWI parameters.

systems to ensure full-brain coverage. To accommodate these changes in acceleration factors and FOV, it was necessary to increase the TR and TE (Table 2).

The pCASL sequence was based on parameters used in the Human Connectome Project (HCP).³⁰ However, to allow sufficient SNR and to be within the capabilities of the Philips Ingenia wide bore scanner we changed a few parameters relative to the HCP protocol. We increased the voxel size from 2.5x2.5x2.3 mm to 3.4x3.4x5mm (see Table 2) and reduced the FOV in the foot-head direction to 119 mm from 182 mm. The TE was set to 14 ms on the Philips scanner whilst maintaining the HCP value of 19 ms on the Siemens scanners. We standardised the label duration to 1400 ms for all scanners, (compared with 1500 ms in the HCP protocol).

Data were then collected for each of the four harmonised sequences (T1W, DTI, fMRI-EPI, and pCASL) from a single, healthy participant who travelled between-sites. This allowed for initial assessment of consistency between-sites in optimised sequences. For the travelling head participant, wherever possible, two within-site scans were acquired on different days for each of the three sites resulting in two repeated sessions per site. For any scans performed on the same day, the participant left the scanner and was repositioned between acquisitions. The participant refrained from consuming caffeine and smoking before the scan on the same day and between scanning sessions. For our future mTBI studies, EPI data will be acquired during 3 protocols: i) Choice Reaction Task³¹ ii) Resting State³² and iii) Cere-

brovascular Reactivity Task.³³ During the EPI data collected for this optimisation stage, the CRT was performed but is not the focus of our analysis. Here we used measures relevant to all tasks expected to be employed in future work. At the time of sequence optimisation, the ASL sequence was not available at Site 3.

IMAGE PROCESSING

Data from all sites were exported as DICOM images and converted to NIFTI using dcm2nii (V1.0.2.20220720).³⁴ Quantitative metrics that could be compared between scanners were extracted from the data acquired from each imaging sequence. The analysis pipelines all utilised widely adopted neuroimaging tools to ensure ease of use for subsequent mTBI-Predict consortium work, as well as by others in the research community wanting to replicate this work. The metrics and processing pipelines are summarised below.

CORTICAL THICKNESS AND SUBCORTICAL VOLUME

The T1W MPRAGE images were processed to obtain measures of cortical thickness and subcortical volumes using the default FreeSurfer³⁵ v.6.0 pipeline (<https://surfer.nmr.mgh.harvard.edu/fswiki/recon-all>). This pipeline implements all necessary steps to pre-process T1W images to extract cortical thickness and subcortical volume estimations.

In detail, this pipeline consists of the following steps: non-parametric, non-uniformity intensity correction, automated affine transformation from the native T1 space to the MNI305 atlas using Talairach registration, intensity normalization to correct for fluctuations, brain extraction, subcortical segmentation, white matter segmentation, cortical surface reconstruction, cortical parcellation (to assign neuroanatomical labels to each location on the cortical surface) and computation of parcellation statistics for each structure. The cortical parcellation was based on the Desikan-Killiany Atlas, which consists of 31 cortical regions per hemisphere and 7 subcortical structures. This atlas was utilised to measure the average cortical thickness and subcortical volume in each brain region. Bilateral regional values were averaged for subsequent analysis.^{36,37} The ROIs used are given in Supplementary Figure 1 for cortical thickness and Supplementary Figure 3 for subcortical volumes. Additionally, the mean cortical thickness and mean subcortical volume values for each session were used as global measures, providing a comprehensive overview of whole brain morphology.

FRACTIONAL ANISOTROPY AND MEAN DIFFUSIVITY

DWI data were processed to obtain measures of fractional anisotropy (FA) and mean diffusivity (MD). These data were first pre-processed to correct for artefacts. Eddy current correction was performed using the Eddy Tool³⁸ in FSL, which corrects for distortions caused by eddy currents and head motion during diffusion imaging. B0 distortion correction³⁹ was carried out using the TOPUP tool in FSL, using the DWI data from the b=0 acquisition and a b=0 image acquired with reverse phase encoding direction to estimate and correct for susceptibility-induced distortions in the B0 field.

Following pre-processing, diffusion maps such as FA and MD were estimated using DTIFIT in FMRIB's Diffusion Toolbox. DTIFIT fits a tensor model at each voxel of the diffusion data, generating FA and MD maps for further analysis. Skeletonised FA and MD values were then extracted using Tract-Based Spatial Statistics (TBSS)⁴⁰ in FSL. TBSS performs several key steps: first, non-linear registration was applied to align all FA images from each session into a standard space; then, a mean FA image was generated and skeletonised using a threshold of 0.2 to create a representation of the central white matter tracts shared by all sessions from each site. Finally, individual FA data from each session were projected onto the mean FA image. For MD, a separate standard, TBSS FSL pipeline was used for non-FA data to ensure that the skeletonisation was applied correctly to the MD maps.

Global FA and MD values were derived from the FA and MD skeletonised images, respectively. These images were extracted using the mean FA skeleton mask generated by TBSS. The skeletonisation process preserves the centre of the white matter tracts, enhancing the accuracy of subsequent analysis. Global FA and MD values for each session were calculated using FSLMATHS and FSLSTATS. Finally, FA and MD maps were registered to standard space, and mean FA and MD values were calculated for individual

tracts using atlas-based region of interest (ROI) extraction, using the Johns Hopkins University white matter atlas available in FSL. A total of 27 regions were included in the analysis, comprising 6 midline tracts and 21 bilateral tracts, where values from the left and right hemispheres were averaged to obtain a single measure per tract. Tract names are given in Supplementary Figure 5.

TSNR MAPPING

To establish the quality of EPI data across scanners the tSNR of data acquired on each scanner was compared. EPI data from each acquisition were first pre-processed to perform correction of susceptibility distortions using TOPUP in FSL.³⁹ Data were then motion corrected using MCFLIRT in FSL.⁴¹ tSNR maps were generated using an in-house MATLAB script which, for each voxel, calculated the mean signal divided by the standard deviation of the signal, both calculated over the full duration of the scan. To exclude outliers, a threshold of 5% of the maximum signal intensity was set and any voxel with signal below this threshold was not included in tSNR calculations. This approach ensured that the calculated tSNR values accurately reflect the true signal variability within the ROI. The grey matter mask that was generated by the processing of the T1 anatomical data was registered to the EPI space using FLIRT in FSL.⁴² In addition, ROI masks of the frontal lobe, cingulate gyrus, motor gyrus, occipital lobe, and parietal lobe (derived from Harvard Oxford atlas) were moved, via the T1 MPRAGE image, to the native EPI space using FNIRT in FSL.⁴³ Mean tSNR values were calculated for the whole of grey matter and the grey matter within the five ROIs. Average tSNR values over the 3 runs within each session were then calculated, to give a mean tSNR value per ROI per session which were used for comparison between-sites and sessions.

CEREBRAL BLOOD FLOW

To establish the reproducibility of pCASL data across sites, cerebral blood flow (CBF) was estimated. pCASL data from all five post-label delays (PLDs) were concatenated and motion-corrected using MCFLIRT in FSL. The average tag-control difference image for each PLD, and the M0 calibration images, were input into the Bayesian Inference for Arterial Spin Labeling (BASIL) toolbox in FSL.⁴⁴⁻⁴⁶ Parameters used in the kinetic model to estimate CBF were: bolus duration=1.4 s, other input parameters were kept at their default values including tissue $T_1=1.3$ s, arterial blood $T_1=1.65$ s, T_2 for blood=0.15s, labelling efficiency=0.85. Partial volume correction was applied within the BASIL analysis pipeline using the T1 MPRAGE.^{45,47} Quantified CBF maps (in ml/min/100 g of tissue) were used to extract mean CBF values for the global grey matter CBF and the same five regional ROIs used in the EPI analysis (frontal lobe, cingulate gyrus, motor gyrus, occipital lobe, and parietal lobe), using an in-house MATLAB script.

IMAGE QUALITY AND MOTION ASSESSMENT

To evaluate site-specific differences in EPI-based data (DWI, fMRI, and pCASL), the effective echo spacing (EES), total readout time (TRT), and apparent point-spread function were quantified. Effective echo spacing and total readout time were derived from DICOM timing parameters, accounting for in-plane (SENSE or GRAPPA) and MB acceleration factors. Spatial smoothness was estimated using AFNI 3dFWHMx with the autocorrelation function to obtain full width at half-maximum (FWHM) estimates of the data.

For diffusion data, we also considered sequence timing differences relevant to inter-vendor reproducibility. Diffusion-timing parameters (Δ and δ) were not provided in vendor exports. Therefore, values were estimated using the Stejskal–Tanner⁴⁸ relationship and sequence echo times, assuming $\Delta \approx TE / 2$. For the Siemens Prisma $\Delta \approx 46$ ms and $\delta \approx 26$ ms, whereas for the Philips Ingenia $\Delta \approx 48.5$ ms and $\delta \approx 18$ ms (see also Supplementary Table 2). These parameters were considered when interpreting inter-vendor diffusion reproducibility.

Head motion was estimated for DWI, fMRI-EPI, and pCASL data using an FSL-based framework. For DWI data, motion estimates were derived from eddy outputs, including root-mean-square (RMS) displacement between diffusion volumes and six rigid-body motion parameters (three translations and three rotations).³⁸ From these, the mean RMS, as well as mean translation and rotation, were calculated per session to assess participant stability within and across sites. For fMRI-EPI and pCASL data both framewise displacement (FD) and RMS values were extracted using MCFLIRT and summarised across sites and sessions

STATISTICAL ANALYSIS

To assess the repeatability and reproducibility of each MRI metric for the travelling head, we calculated the within-subject coefficient of variation (wCV%) across both sessions and sites, along with conducting linear regression analyses. The wCV% analysis was performed for both global and ROI values for all metrics. Linear regression analyses were only conducted on the ROI values.

WITHIN-SITE ANALYSIS

COEFFICIENT OF VARIATION

For the within-site analysis, the wCV% was calculated using the data from the two sessions conducted at the same site for each MRI metric. wCV% was calculated for each ROI for a given metric. To calculate wCV% Eq. 1 was used:

$$wCV\% = \frac{\sigma_{data}}{\mu_{data}} \times 100 \quad (1)$$

where σ_{data} is the standard deviation of data points (in the within-site analysis this was taken within each ROI across sessions within a site), and μ_{data} is the corresponding mean of the data points. In line with previous literature,^{49–53} we defined reproducibility as: 1) excellent when wCV% was less than 5%; 2) very good when wCV% was between 5 and 10

%; 3) good when wCV% was between 10% to 20% and 4) a wCV% greater than 20% indicated poor reproducibility. To provide an indication of statistical precision, 95% confidence intervals (CI) were calculated from the regional measures using a bootstrapping approach in SPSS.

LINEAR REGRESSION ANALYSIS

For each imaging metric, a linear regression analysis was conducted across the ROIs, comparing either sessions or sites.

In the within-site analysis, the regression was performed using session data from the two sessions conducted at the same site. For each regression model, the gradient, intercept, and R^2 values were calculated. Generally, an R^2 value of 0.90 or greater indicates a strong relationship between measures, while an R^2 value between 0.70 and 0.90 signifies a good relationship. An R^2 value ranging from 0.50 to 0.70 indicates a moderate relationship.⁵³ In addition, deviations in the gradient of the line from 1, combined with the intercept deviating from 0 provided an indication of systematic biases between datasets acquired. Together, these statistical measures allowed us to evaluate the consistency of the metrics within the same scanner and understand what was driving low wCV% values.

BETWEEN-SITE ANALYSIS

For the between-site analysis, the measurements from the two sessions at each site were averaged. This averaging prevented bias from selecting a specific session for between-site comparisons. Here the wCV% from Eq1 was calculated where the standard deviation across sites and corresponding mean were used.

For the between-site regression analysis, the averaged values from the two sessions for each site were used. The outcome metrics from these were the same as the within-site analysis. The results from these analyses aided the identification of any systematic differences and in assessing the overall reproducibility of the MRI metrics. To quantify precision, 95% CI for the between-site wCV% were derived from the distribution of regional wCV% values across brain regions using a bootstrapping approach in SPSS.

RESULTS

Table 3 and Supplementary Table 1 summarise the wCV% to assess reproducibility over all MRI metrics considered. Table 3 shows the results for the relevant whole brain ROIs, whilst Supplementary Table 1 shows the results when averaging wCV% values calculated from the regional ROIs.

Tables 4 and 5 summarise the results of the linear regression analyses conducted for both within-site and between-site assessments, respectively, of all the MRI metrics.

Below we evaluate each of the MRI metrics separately.

Table 3. Coefficient of variation within-site and between-site calculated from global MRI metrics.

MRI metric	Within-site wCV%			Between-site wCV%
	Site 1	Site 2	Site 3	
Cortical thickness	0.26	0.60	0.03	1.03
Subcortical volume	0.86	0.15	1.53	2.60
FA	0.05	0.72	0.41	4.39
MD	0.37	0.11	1.76	4.95
tSNR	1.62	7.09	3.75	3.97
CBF	7.03	1.59		8.53

Note: wCV% values are below the threshold of 10%.⁴⁹ *pCASL data not available for Site 3.

Table 4. Results of Linear Regression Analysis of MRI Metrics within-sites (comparing session 1 and 2 for each site).

MRI metrics	Within-site	Gradients	Intercept	R ²
Cortical Thickness	Site1	0.93	0.16	0.98
	Site2	0.94	0.16	0.98
	Site3	1.08	-0.19	0.98
Subcortical Volume	Site1	0.97	62.57	1.00
	Site2	0.97	81.45	1.00
	Site3	0.94	99.53	0.99
FA	Site1	0.98	0.01	0.99
	Site2	1.00	0.00	1.00
	Site3	0.98	0.02	0.99
MD	Site1	1.01	0.00	0.98
	Site2	1.00	0.00	0.99
	Site3	1.04	0.00	0.96
tSNR	Site1	0.76	10.33	0.79
	Site2	0.83	2.77	0.95
	Site3	1.06	1.62	0.96
CBF	Site1	1.50	-21.82	0.91
	Site2	1.44	-28.21	0.90

Note: The values reflect the results of linear regression analyses across all relevant ROIs for the MRI metrics. See also [Figures 1-6\(a\)](#). *pCASL data was not available for Site 3.

CORTICAL THICKNESS AND SUBCORTICAL VOLUME

Cortical thickness measures were found to have excellent reproducibility (wCV% < 5%) and were strongly related ($R^2 > 0.90$) across sessions and sites from our global measures, shown in [Table 3](#) and Supplementary Table 1. When considering individual brain regions, as expected, greater variability was seen both within and between-sites (Supplementary Figures 1 and 2). The majority of ROIs interrogated showed a wCV% below 5% for cortical thickness when considering both within and between-sites. However, the entorhinal area showed an elevated wCV% of 7.2 % likely due to the location in the brain and the small volume of this region. When considering all ROIs together, the overall consistency (R^2 values) of measures across regions was high ($R^2 > 0.90$) both within-sites ([Table 4](#)) and between-sites ([Table 5](#)), with no clear systematic biases, as shown by the gradient and/or intercept of the line of best fit from the linear regression ([Figure 1](#)).

Regarding subcortical volumes, as expected due to the small volumes considered in many structures, the wCV% were in general larger than for the cortical thicknesses ([Table 3](#) and Supplementary Table 1). When considering subcortical volumes ([Table 3](#)), the wCV% remained excellent (< 5%) showing high reproducibility. In general, for each ROI the wCV% both within and between-sites was also very good (<10%). The only exception was the nucleus accumbens where Site 3 had higher between scan variability (wCV% = 17.27%) which likely drove the between-site variability for this region (Supplementary Figures 3 and 4). Notably, the nucleus accumbens was the smallest volume measured (bottom left point on the graphs in [Figure 2](#)). Whilst the reproducibility of this structure was the lowest (Supplementary Table 1) it does sit on the line of best fit suggesting there is no systematic errors generated in measuring this volume within or between-sites. The overall consistency of measures from the linear regression across regions was excellent both within ([Table 4](#) and [Figure 2a](#))

Table 5. Results of Linear Regression Analysis of MRI Metrics between-sites.

MRI metrics	Between-site	Gradients	Intercept	R ²
Cortical Thickness	Site 1 vs Site 2	1.07	-0.14	0.93
	Site 1 vs Site 3	1.09	-0.19	0.96
	Site 2 vs Site 3	1.01	-0.00	0.99
Subcortical Volume	Site 1 vs Site 2	1.08	-94.52	1.00
	Site 1 vs Site 3	1.04	-23.35	1.00
	Site 2 vs Site 3	0.96	70.71	1.00
FA	Site 1 vs Site 2	0.95	0.00	0.97
	Site 1 vs Site 3	0.94	0.01	0.96
	Site 2 vs Site 3	0.99	0.01	1.00
MD	Site 1 vs Site 2	1.04	0.00	0.85
	Site 1 vs Site 3	1.02	0.00	0.85
	Site 2 vs Site 3	0.98	0.00	0.99
tSNR	Site 1 vs Site 2	0.88	6.11	0.93
	Site 1 vs Site 3	0.84	9.99	0.90
	Site 2 vs Site 3	0.95	4.18	0.96
CBF	Site 1 vs Site 2	1.24	-0.09	0.79

Note: The values reflect the results of linear regression analyses across all relevant ROIs for the MRI metrics for the between-site analysis. See also [Figures 1-6\(b\)](#). *pCASL data was not available for Site 3.

and between-sites ([Table 5](#) and [Figure 2b](#)) with no clear systematic biases indicated by a change in the gradient and/or intercept of the line of best fit.

These findings show the high reproducibility of structural measures of cortical thickness and subcortical volumes.

FRACTIONAL ANISOTROPY AND MEAN DIFFUSIVITY

When considering white matter measures, FA exhibited excellent reproducibility (wCV% <5%) both across the whole brain white matter tracts ([Tables 3](#) and Supplementary Table 1) and all regions included in the ROI analysis, within sites (Supplementary Figure 5) whilst very good reproducibility (wCV% <10%) of FA across sites within ROIs was achieved (Supplementary Figure 6). [Figure 3](#) and [Table 4](#) show that there was very good consistency in FA values across ROIs between-sessions within a site (high R², gradient close to 1 and intercept close to 0). However, there was some variation in values collected at Site 1 compared with Sites 2 and 3 ([Figure 3b](#)) which was likely to be driving the higher wCV% seen between-sites ([Table 3](#) and Supplementary Figure 6).

The MD measures showed a similar pattern of excellent reproducibility between-sessions within a site for the global measures and the average of all the ROI measures (wCV% <5%). Whilst very good reproducibility between-sites (wCV% <10%) was achieved for both the global measures ([Table 3](#) and Supplementary Table 1) and individual ROIs (Supplementary Figures 7 and 8). Again, differences between Site 1 and Sites 2 and 3 clearly drove the lower reproducibility between-sites compared to within-site ([Figure 4](#)).

Together these findings show that excellent (within-site) and very high (between-site), reproducibility of FA and MD measures derived from our optimised DWI sequence were achieved.

TEMPORAL SIGNAL TO NOISE RATIO

When considering tSNR across the whole of grey matter, in general we achieved excellent reproducibility both within-site and between-site (wCV% <5%, [Table 3](#)). The only exception to this was Site 2, where very good between session reproducibility was achieved (wCV% <10%, [Table 3](#)). When considering individual ROIs, there was a wider spread of values for the wCV% than seen in general for structural measures (Supplementary Figures 9 and 10). Mirroring the global values the between session reproducibility for individual ROIs was lowest (highest wCV% values) in Site 2 for the majority of regions considered, but a good reproducibility was reached in this site (wCV% <20%). When considering the consistency (R²) between-sessions across regions, [Figure 5a](#) shows that this was the lowest for Site 1 (lowest R², mostly driven by the two regions with the lowest tSNR, which were the frontal lobe and the occipital lobe). [Figure 5b](#) shows that between-sites, where values from the two within-site sessions had been averaged, the consistency across the ROIs was very high with R² > 0.9.

CEREBRAL BLOOD FLOW

When assessing CBF across the entirety of grey matter, we observed within-site wCV% values <10%, indicating very good global CBF reproducibility. Notably, Site 1 exhibited higher wCV% values between the two sessions than Site 2, as shown in [Table 3](#). Additionally, the between-site wCV%

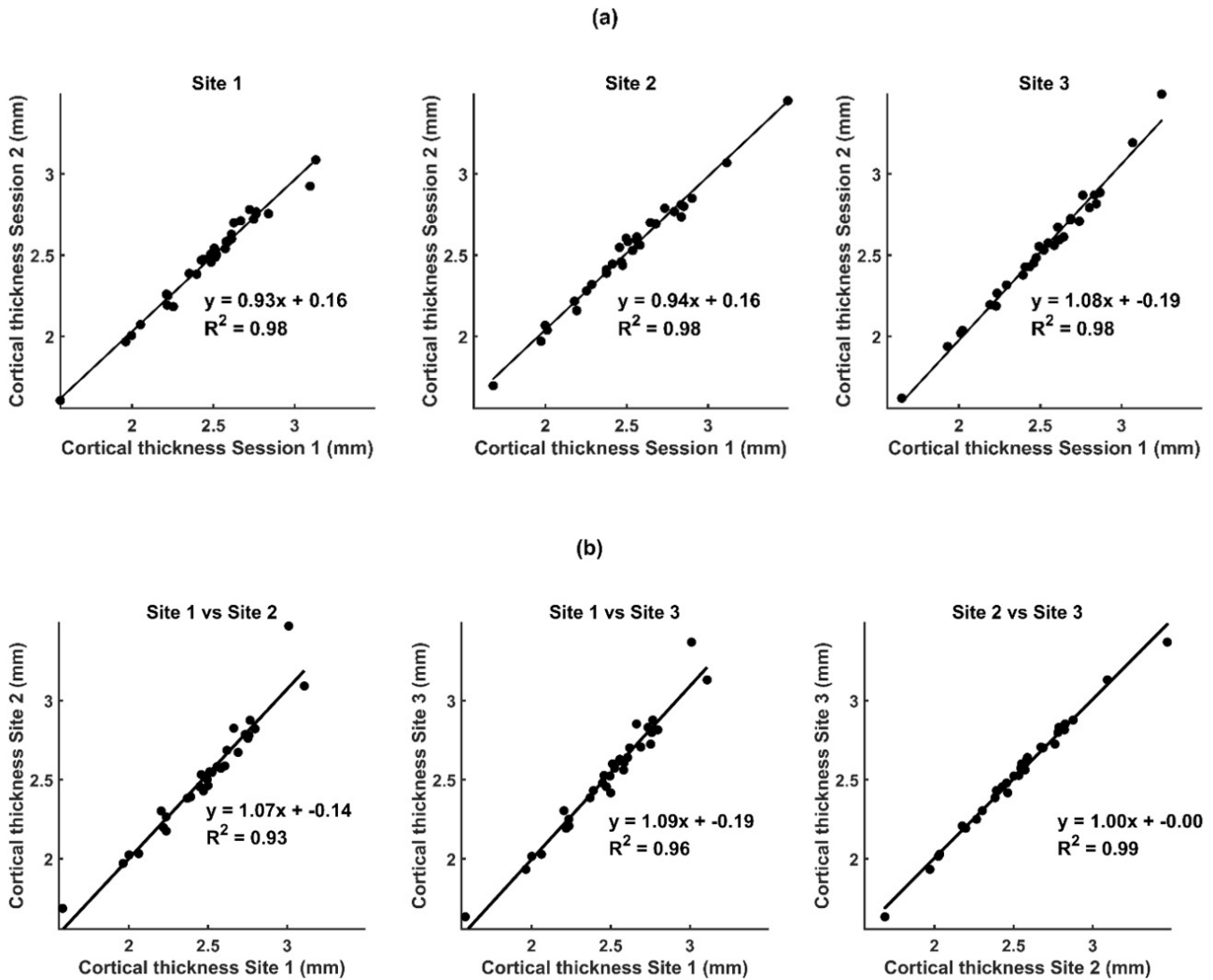


Figure 1. Regression analysis of cortical thickness a) between-sessions at each site and b) between-sites. Each data point represents an ROI included in the analysis, names of each of the ROIs are given in Supplementary Figure 1.

was <10%, reinforcing the overall very good reproducibility of this physiological measurement.

As expected, examination of individual ROIs revealed higher wCV% values compared to structural measures, as shown in Supplementary Figures 11 and 12. Interestingly, the regional variation in wCV% values was different for the two sites, with lower reproducibility in parietal and occipital lobes for Site 2, whilst the reproducibility was lower between-sessions in the motor and frontal lobes for Site 1 (Supplementary Figure 11). The between-session reproducibility was similar across sites for the cingulate gyrus, which interestingly had the lowest between-site reproducibility (Supplementary Figure 12). This between-site reproducibility was the lowest of all the measures considered in this work.

Despite the relatively low wCV% values (generally 10-20%) for each region, when considered over all the regions, the consistency (R^2) between-sessions across these regions was very good for both Site 1 and Site 2, as indicated in Table 5 and Figure 6(a). However, in line with the reproducibility of the measures across regions between-sites, the between-site consistency was good but lower

compared to all other between-site comparisons of MRI metrics made in this work, as illustrated in Figure 6(b).

APPARENT POINT-SPREAD FUNCTION

For DWI, fMRI and pCASL data, the FWHM estimates were lower for the Philips system (Site 1) compared with the Siemens scanners (Sites 2 and 3), as shown in Supplementary Table 3. The magnitude of this difference between vendors varied between image acquisitions with the largest difference for fMRI data and the smallest for DWI data. Importantly, the largest differences in EES and TRT estimates between vendors were also seen in the fMRI acquisition (see Supplementary Table 3).

MOTION PARAMETERS

DWI DATA

Motion estimates demonstrated higher overall total displacement at the Philips site (Site 1) compared with both Siemens sites (Sites 2 and 3). However, inspection of the motion parameters over time showed that the majority of

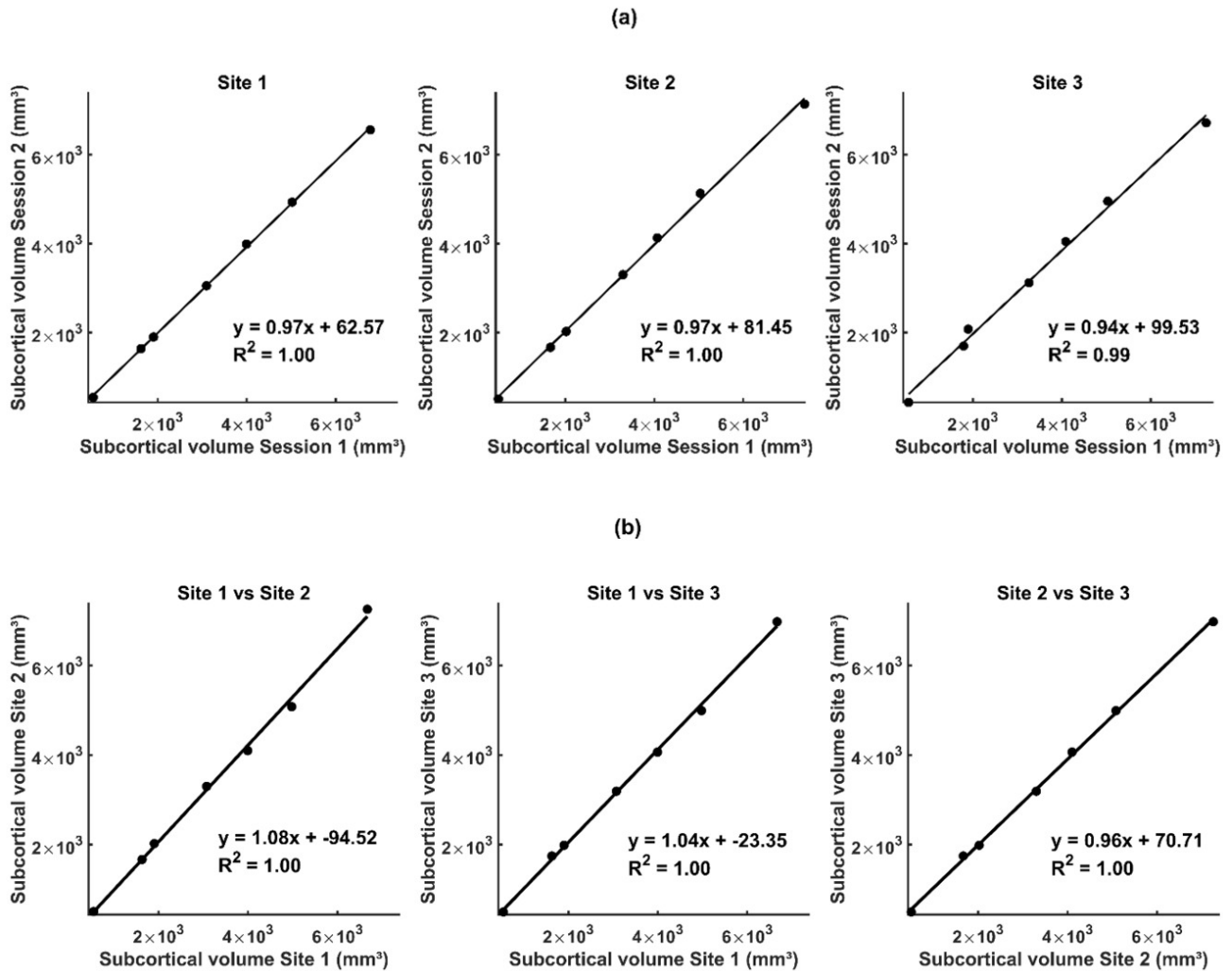


Figure 2. Regression analysis of subcortical volume a) between-sessions at each site b) between-sites. Each data point represents an ROI included in the analysis, names of each of the ROIs are given in Supplementary Figure 3.

the “motion” was linear drift due to heating of the gradient coils and passive shims which accumulated over time, due to the longer acquisition time at Site 1⁵⁴ (see Supplementary Figure 11). Looking at motion per unit time, the values were similar across sites, as expected (Supplementary Table 4). Importantly, all sessions remained well below the diffusion MRI motion thresholds (≈ 1 mm RMS; half-voxel criterion) generally considered acceptable for reliable tensor and fibre-tracking analysis.⁵⁵⁻⁵⁷

FMRI DATA

Both framewise displacement and RMS values remained low across sessions and sites, with only a slight increase in motion observed during the first session at Site 2, RMS = 0.44 ± 0.20 (mean \pm SD) (see Supplementary Table 5). Overall, motion was small and did not vary systematically between sites.

PCASL DATA

Motion was minimal across sessions for these data. There was marginally higher fractional displacement observed in

the Siemens site (Site 2, session 1) compared with all other sessions (see Supplementary Table 5). All sessions were below established thresholds for head motion (FD < 0.2 mm; RMS < 0.5 mm), consistent with accepted stability criteria for EPI-based acquisitions.^{57,58}

DISCUSSION

Imaging techniques that yield consistent longitudinal measurements across different MRI scanners and vendors have become increasingly important as many clinical and neuroscience studies, including clinical trials, adopt multi-site approaches to data collection. The need for MRI metrics that are quantitative and reproducible across vendors, scanners and sites is vital for any measure that may become a clinically useful biomarker. Importantly, this is needed to move beyond the classical clinical use of MRI, which is to visually inspect and report on the structural images that are acquired in a clinical setting.⁵⁹ In this work, we investigated the reproducibility and consistency of MRI measures from optimised sequences and standard analysis pipelines with the view of using them to derive biomarkers for mTBI.

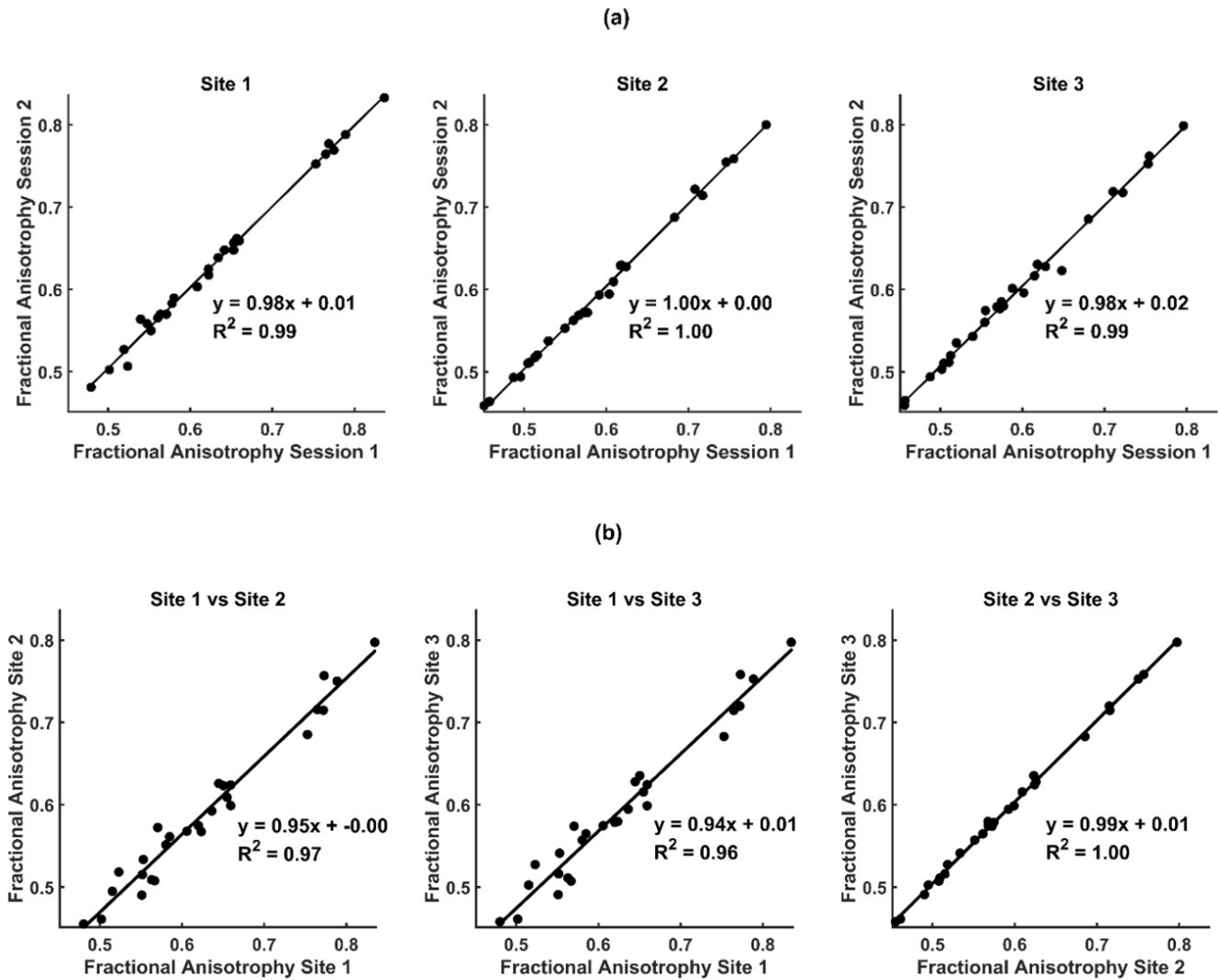


Figure 3. Regression analysis of FA a) between-sessions at each site and b) between-sites. Each data point represents an ROI included in the analysis, names of each of the ROIs are given in Supplementary Figure 5.

Overall, motion remained low and stable across all sites, likely due to the considerable familiarity the participant had with the MRI environment. This low motion supported the subsequent between-site comparisons. The low levels of motion for this experienced participant suggest that the reproducibility between scanners of the sequences represents a best-case scenario; in patients, reproducibility is likely to be poorer. All of our measures at a global level exhibited at least very good reproducibility ($wCV\% < 10\%$) with most of them categorised as excellent ($wCV\% < 5\%$). The structural metrics were in general more reproducible and consistent than the measures derived from MR sequences designed to study aspects of brain function, as might be expected given the dynamic and integrative nature of brain function. Overall, we found excellent reproducibility ($wCV\%$) and consistency (R^2) between-sessions and sites for both the cortical thickness and subcortical volume measures. When considering individual regions all regions had at least very good ($wCV\% < 10\%$) reproducibility. The highest $wCV\%$ (denoting poorest reproducibility) was in the entorhinal region for cortical thickness (Supplementary Figures 1 and 2) and nucleus accumbens for subcor-

tical volume (Supplementary Figures 3 and 4). These regions are small and located near the inferior frontal portion of the brain, and prone to susceptibility artifacts due to the tissue-bone interface. These artifacts can make it challenging to accurately differentiate anatomical boundaries in these areas, potentially increasing the variability in cortical thickness and subcortical volume measurements.^{36,37} However, despite these challenges we achieve very good reproducibility in these areas. Our results align with those of McGuire et al. (2017), Fujita et al. (2019) and Bano et al (2024) who reported poor reproducibility in the entorhinal region.^{2,37,60} Similarly, the variation we observed in subcortical volume measurements in the nucleus accumbens area is generally consistent with other studies.^{37,60} However, in general, these results show that we have successfully optimised our T1 anatomical sequence across scanners and vendors.

When considering the metrics derived from the DWI acquisition, FA and MD global measures showed excellent reproducibility both between and within-sites (Table 3). The $wCV\%$ values for FA within sites were considerably better than those reported in many other studies.^{2,49,61}

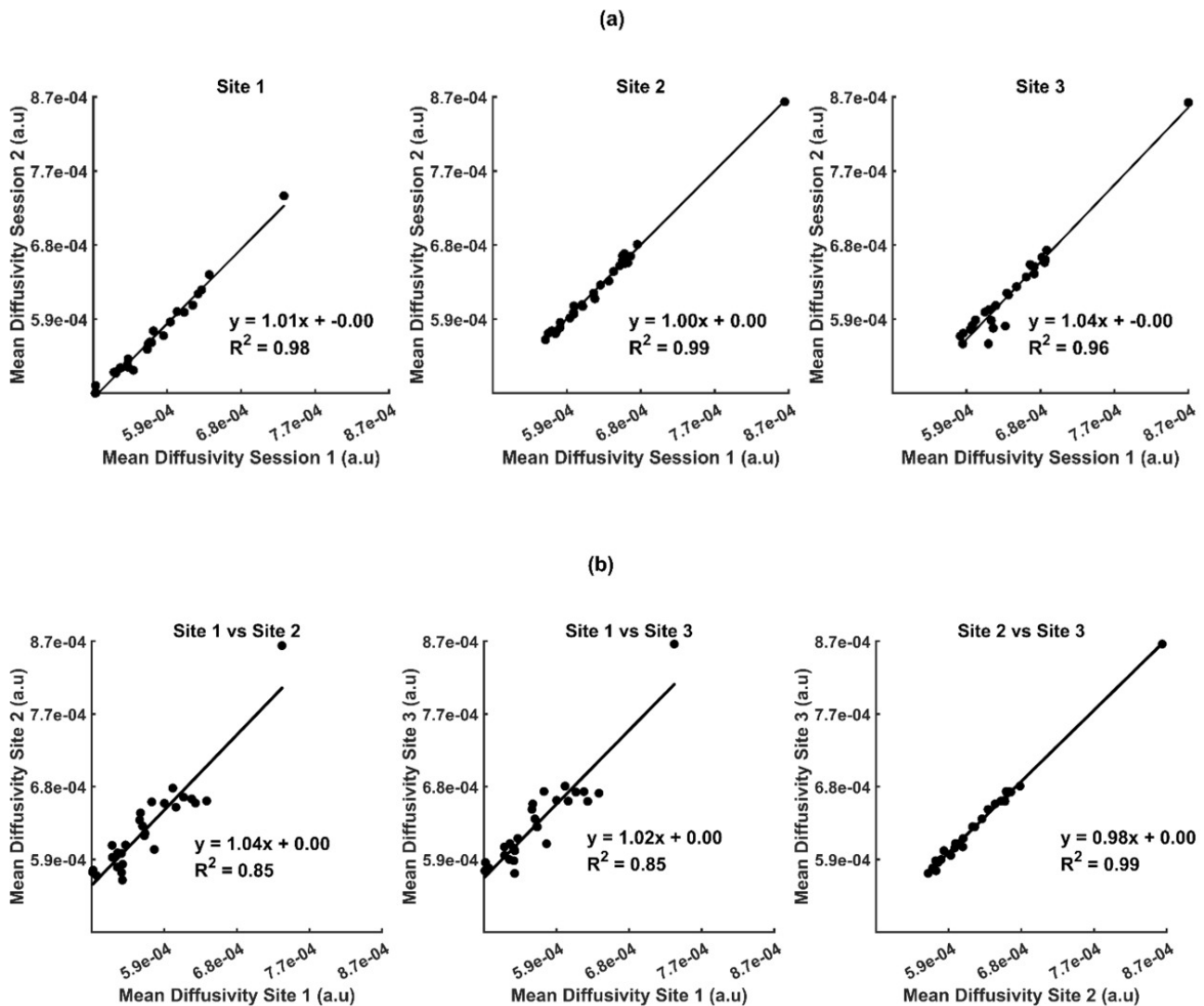


Figure 4. Regression analysis of a) MD between-sessions at each site and b) between-sites. Each data point represents an ROI included in the analysis, names of each of the ROIs are given in Supplementary Figure 7.

The mean global white matter FA between-sites achieved a wCV% <5%, similar to a multi-scanner study by Grech-Sollars and colleagues.⁶² The within-site reproducibility of the MD measures globally was still excellent (Table 3), but MD showed greater variability when measured in a number of specific white matter tracts (Supplementary Figures 7 and 8) compared with the FA values (Supplementary Figures 5 and 6). The reproducibility of MD between-sites was very good (showing wCV% 5 - 10%). Notably, the genu of the callosum exhibited a relatively higher between-site wCV%, but this remains below 10% indicating very good reproducibility between-sites. The overall pattern observed here, as well as the specific region with the highest wCV%, is in agreement with previous work.^{51,63} While these results highlight the care that must be taken in interpreting metrics from different regions, in general they support the excellent reproducibility of DWI metrics.^{60,61}

When considering the between-site reproducibility specifically, it is clear that the reduction in reproducibility (R^2 values) that is seen for the MD and FA values compared with within-site values is driven by differences between-

site 1 compared with Sites 2 and 3 (Figures 3 & 4b). Notably Site 1 had a Philips Ingenia scanner whilst Sites 2 and 3 both had a Siemens Prisma (Table 1). Although EES and TRT were well matched between vendors, differences in gradient performance and bore size could not be avoided. Vendor-related differences in diffusion timing, specifically the diffusion gradient separation time (Δ) and the diffusion gradient duration (δ), and readout characteristics were found with slightly longer values for the Philips system than the Siemens system, likely contributing to the observed variability.⁶⁴ Consistent with these vendor-related differences, the apparent point-spread function (FWHM) values were generally lower for the Philips system compared with the Siemens scanners, indicating sharper EPI spatial profiles across sessions for Philips DWI data. In addition, the Philips DWI sequence required a longer TR due to the hardware and safety constraints. The longer TR permits greater T_1 recovery between diffusion volumes, potentially altering baseline signal and SNR and introducing subtle inter-vendor bias in tensor fitting.^{64,65} However, from our data we observed that the gradients and intercepts of

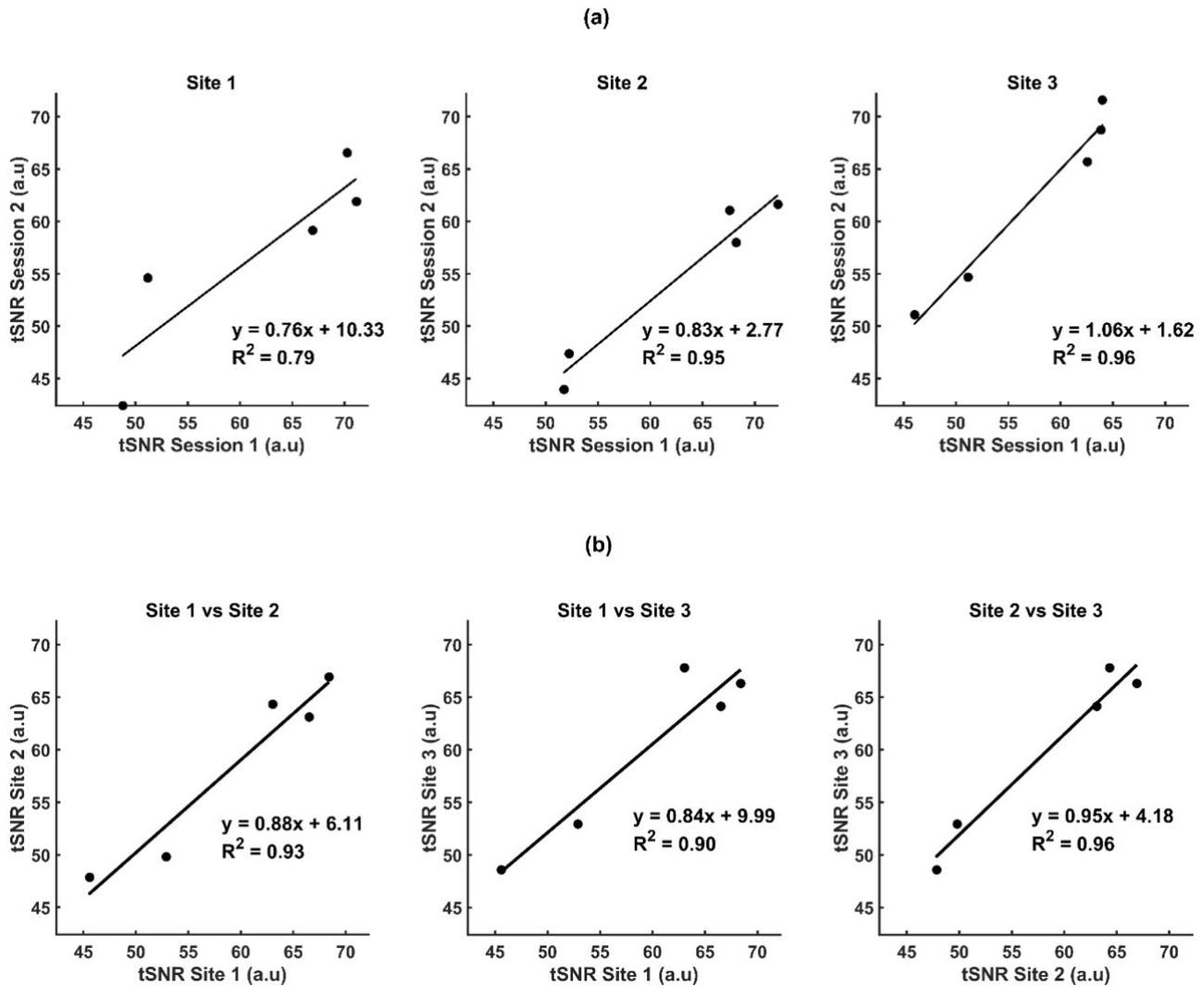


Figure 5. Regression analysis of tSNR a) between-sessions at each site b) between-sites. Each data point represents an ROI included in the analysis, names of each of the ROIs are given in Supplementary Figure 9.

the regression fits remain close to 1 and 0, respectively (Figures 3 and 4; Tables 4 and 5). Therefore, we suggest that the small between-vendor variability is unlikely to be driven by systematic differences between the systems but rather by subtle differences in SNR and acquisition characteristics driven by the differences in hardware.

A crucial property of EPI data, tSNR plays a vital role in ensuring reproducible results for any task performed during an fMRI acquisition. Therefore, to allow future fMRI results across scanners to be comparable, matching tSNR and ensuring it is as high as possible is important to assess during optimisation. Matching tSNR of EPI acquisitions across sites with different scanners is challenging due to several factors. tSNR will be affected by many facets of the MRI acquisition⁶⁶ from MRI hardware, pulse sequence design, to reconstruction methods e.g. parallel imaging techniques such as SENSE⁶⁷ vs GRAPPA⁶⁸ and multiband⁶⁹ implementations. In addition, EPI data will also be affected by physiological noise as well as the thermal noise that affects all MRI acquisition.^{66,70} Therefore, reproducibility of tSNR between-sites is expected to be lower than within-site unless the scanners and sequences, to the level of the RF

pulse designs, have been matched. This level of sequence matching between MR vendors is unfeasible in most studies and certainly not feasible within clinical settings where different hospitals have different 3T MRI scanners and CE/FDA markings must be maintained (i.e., altering pulse sequences and hardware would violate the CE marking).

With the known differences between scanners across sites (Table 1) in our study, we aimed to match tSNR as closely as possible. As outlined in the methods, this was done through an iterative process of altering multiband and image acceleration factors and assessing tSNR on a phantom (data not shown). This enabled us to match tSNR as closely as possible across sites whilst also maintaining matched TE, TR and voxel sizes, all crucial for fMRI studies. The optimised parameters are shown in Table 2, where at Site 1 MB = 2 with SENSE = 2 was used whereas at Sites 2 & 3 MB = 4 with no GRAPPA was used. Between vendor differences in the sequences were also reflected in the spatial smoothness metric: FWHM values were lower for the Philips site than the Siemens sites, mirroring the distinct echo-train durations and reconstruction methods used by each vendor. Previous work has shown that parallel imaging

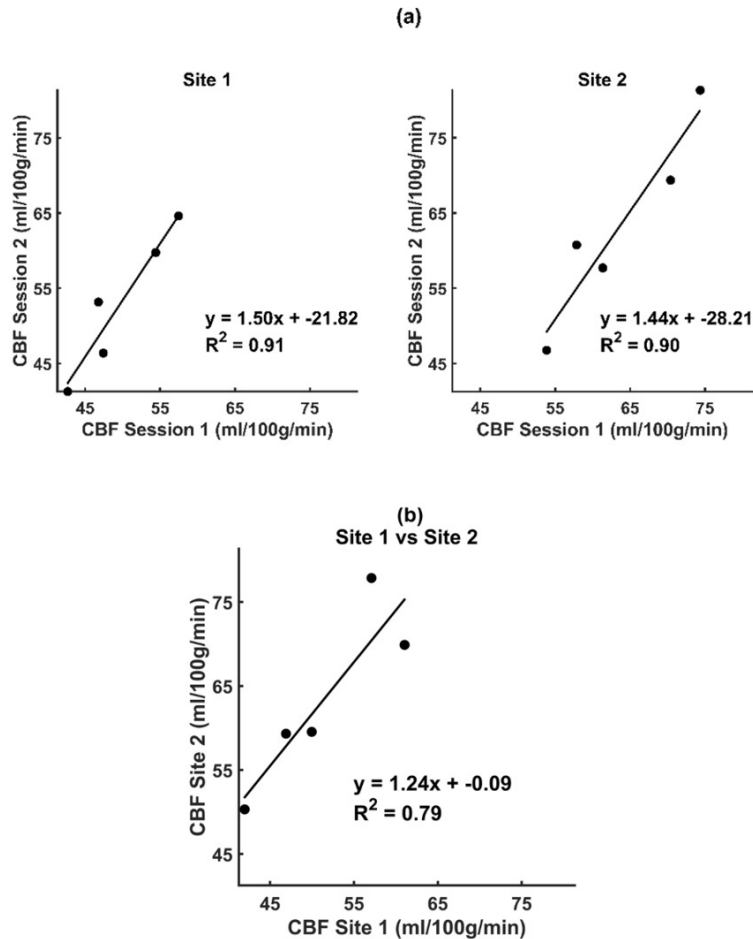


Figure 6. Regression analysis of CBF a) between-sessions at each site b) between-sites. Each data point represents an ROI included in the analysis, names of each of the ROIs are given in Supplementary Figure 11.

shortens the EPI echo train and reduces geometric distortion and T_2 -related blurring, thereby yielding sharper spatial profiles.⁷¹ The differences in point-spread function that were seen may be overcome with post-processing smoothing of data, as previously suggested.⁷²

The wCV% and R2 of tSNR values when comparing between-sites indicate that the optimisation approach was largely successful in harmonising tSNR across scanners. Although small regional differences were present, when considering wCV across the three sites, between-site variability was comparable to, or slightly lower than, within-site variability, suggesting that between vendor optimisation was relatively successful. Importantly, the regression fits of between-site versus within-site tSNR values showed slopes close to unity and minimal intercept offsets, indicating the absence of systematic bias or directional differences between scanners. However, our analysis cannot eliminate the possibility of regional systematic differences. Together, these findings suggest that the optimisation reduced between-site differences in tSNR although some regional variability may still be present.

The optimisation we achieved produced lower variability than the ABCD study,² which scanned a single subject using four MRI scanners, and showed greater differences across sessions and sites than we observed. In our work, the variability (wCV%) was higher for the tSNR measures from the

functional MRI than from the structural measures. This was expected due to the sensitivity of tSNR to physiological noise.⁷³ Large confounding factors that alter physiology such as caffeine⁷⁴ and nicotine⁷⁵ were minimised by the travelling head participant not consuming caffeine for 12 hours before and not smoking. Other factors such as time in the menstrual cycle⁷⁶ when scans were carried out was not controlled for, as this would not be typically controlled for in future work using these sequences. However, despite the physiological noise our results support previous work demonstrating that fMRI data quality results in very good reproducibility (wCV% <10%) and good consistency ($R^2 > 0.75$) globally and therefore is reliable across sessions and sites.^{23,77}

CBF measures using ASL are inherently more difficult to match across sessions and scanners than many other MRI measures due to the low signal to noise ratio of the acquisition, combined with differences in sequences across vendors.⁷⁸ Like BOLD fMRI, subject motion, acquisition techniques, and post-processing errors can also contribute to small differences in CBF reproducibility.⁷⁹ We attempted to mitigate as many sources of variability as we could by employing a multi-delay pCASL sequence with a 2D-EPI read-out across all sites. We wanted to employ a multi-delay sequence since it is possible that changes in microvasculature in mTBI patients may alter transit times in the pa-

tient group we are interested in.⁸⁰ We employed the same analysis pipeline on data from both sites considered.⁸¹ As a result of our approach, the CBF measurements showed a very good level of within-site reproducibility at both Site 1 and Site 2, as indicated by wCV% values <10% (Table 3). The slightly lower R^2 values between-sessions suggest decreased consistency within-sites compared with structural measures (Table 4), likely due to physiological variability.

When considering between-sites, our reliability was still very good (wCV% values <10%, Table 3) but of all the comparisons these were the lowest between-site agreement. In addition, the consistency when considering the different brain regions was lower with $R^2=0.795$. This reflects the fact that these data were noisier and did not agree across sites as well as other measures. In addition, the estimated FWHM for pCASL images was considerably higher at the Siemens site than at the Philips site reflecting differences in EPI readout implementation and bandwidth contributed to variations in spatial smoothness (Supplementary Table 3). Our findings suggest for these CBF measures a larger change between groups will be needed to be detectable, compared with some of the other measures considered in this work, with inherently higher SNR. Nevertheless, these CBF variability observations agree with previous studies^{78, 82} and changes between groups in CBF measures have been reported.⁸³ The desire to have the multi-delay sequence and to match sequences across scanner platforms prevented the use of the recommended pCASL with 3D-GRASE readout sequence.⁸⁴ However, if this had been available to us at all sites, we would have opted for this sequence which is likely to have improved the reliability of the CBF results obtained, whilst retaining the capability for multi-delay.

Our results suggest the between-site variability is primarily attributed to differences in sequences and vendors across sites, rather than physiological factors which we would expect to observe in the within-site measures as well as the between-site measures. This interpretation is further supported by the minimal head motion observed across sessions (FD < 0.2 mm; RMS < 0.5 mm), indicating that motion-related effects were negligible. One study suggests that even minor changes in ASL sequences can significantly impact CBF measurement. Additionally, small variations in sequence parameters may have a larger effect on ASL reproducibility than hardware or software differences between vendors, even when the same labelling and readouts are used.⁸¹ The main differences in pCASL between the two sites are due to vendor-specific implementations. A more comprehensive comparison between vendors could be performed with additional subjects, given the slight differences in sequences. Moreover, factors like B_0 field homogeneity and signal-to-noise ratio, which may not impact anatomical imaging, can significantly affect ASL sequences.⁸⁵ These findings highlight the inherent challenges of achieving consistent ASL CBF quantification across multiple MRI platforms. To improve reproducibility in future studies, additional steps such as vendor-specific calibration,⁸⁶ partial volume correction,⁸⁷ and post-acquisition harmonisation⁸⁸ may help to minimise systematic differences and enhance cross-site comparability. However,

caution must be exercised when performing post-acquisition harmonisation as between group differences may be obscured if the region chosen to perform the harmonisation is unknowingly altered between the groups. For mTBI specifically this warrants further investigation in the future.

LIMITATIONS

Here we aimed to set up sequences for a subsequent assessment of variability of potential biomarkers of mTBI across a larger cohort of healthy controls and patients. A key limitation of this study is the use of a single healthy participant, which constrains both the choice of statistical analysis and statistical power. However, the primary aim was to examine the reproducibility of key imaging sequences across sessions and sites and to assess data quality prior to collecting data on a larger cohort involving both healthy and mTBI patients. This was a vital first step before progressing to a larger-scale variability study of 20 healthy participants and 20 with mTBI. This study focused specifically on acquisition-level protocol optimisation across scanners rather than applying post-acquisition statistical harmonisation methods. Our intention was to first evaluate the extent to which careful acquisition optimisation alone could minimise inter-scanner differences, thereby establishing a robust foundation for future harmonisation work. No post-acquisition pipelines were applied, and this should be considered when interpreting cross-site variability. This approach was chosen because improving similarity at the acquisition stage reduces the burden on post-acquisition harmonisation methods, which may not fully correct for vendor or sequence-dependent differences.²²

Additionally, the participant was a healthy adult female, whereas previous research suggests that the reproducibility of certain measures may be influenced by factors such as age and sex.^{89,90} The study also did not control for the time of day at which the scans were acquired. However, the factors which cause variability due to time of day, point in menstrual cycle and sex is likely to be present and confounds in any study of clinical populations too.

CONCLUSION

This study assessed the reproducibility and consistency of multi-site imaging measurements, focussing on within and between-site differences using a single travelling head. We employed commonly used neuroimaging sequences in MRI, including T1, DTI, fMRI, and ASL, evaluating our ability to produce consistent results. Structural metrics exhibited excellent reproducibility across sessions and sites, while functional and physiological measures showed very good reproducibility between-sessions and sites, within the limits expected for measures from the human brain.⁴⁹ Overall, this study reports the reliability and reproducibility of neuroimaging measurements, demonstrating their value for evaluating biomarkers in mTBI and investigating disease status and treatment protocols.

.....

DATA AND CODE AVAILABILITY

The imaging data used in this study are formatted according to the Brain Imaging Data Structure and stored on secure institutional servers. Although not publicly available at this time, the data will be made accessible via the OpenNeuro platform upon acceptance of the manuscript. The custom scripts used in the analysis are openly available on GitHub: <https://github.com/MRIRW/mri-sequence-harmonisation>.

FUNDING SOURCES

The work was funded by the Ministry of Defence, United Kingdom, through the mTBI Predict Consortium. The study was performed with the support of the Birmingham Clinical Trials Unit.

Ruwan Wannan Arachchige, Yidian Gao, Iman Idrees, Jessikah Fildes, Aliza Finch, and Waheeda Hawa, were all funded by the mTBI Predict Consortium at the time of data acquisition and analysis for this work.

Prof. Alexandra J. Sinclair is funded by a Sir Jules Thorn Award for Biomedical Science.

CONFLICTS OF INTEREST

Professor Sinclair reports personal fees from Invex therapeutics in her role as Director with stock holdings, during the conduct of the study; other from Allergan, Novartis, Cheisi and Amgen outside the submitted work.

ACKNOWLEDGEMENTS

We acknowledge the input from the United Kingdom mTBI Predict Consortium.

This study has been delivered through the National Institute for Health and Care Research (NIHR) Birmingham Biomedical Research Centre (BRC). The views expressed are those of the author(s) and not necessarily those of the Ministry of Defence, United Kingdom, the NIHR or the Department of Health and Social Care.

Submitted: July 26, 2025 CDT. Accepted: February 09, 2026 CDT. Published: March 27, 2026 CDT.



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

REFERENCES

1. Bento M, Fantini I, Park J, Rittner L, Frayne R. Deep Learning in Large and Multi-Site Structural Brain MR Imaging Datasets. *Front Neuroinform.* 2022;15. doi:[10.3389/fninf.2021.805669](https://doi.org/10.3389/fninf.2021.805669)
2. Bano W, Pulli E, Cantonas L, et al. Implementing ABCD study® MRI sequences for multi-site cohort studies: Practical guide to necessary steps, preprocessing methods, and challenges. *MethodsX.* 2024;12. doi:[10.1016/j.mex.2024.102789](https://doi.org/10.1016/j.mex.2024.102789)
3. Van Horn JD, Toga AW. Multisite neuroimaging trials. *Curr Opin Neurol.* 2009;22(4):370-378. doi:[10.1097/WCO.0b013e32832d92de](https://doi.org/10.1097/WCO.0b013e32832d92de)
4. Olsen A, Babikian T, Bigler ED, et al. Toward a global and reproducible science for brain imaging in neurotrauma: the ENIGMA adult moderate/severe traumatic brain injury working group. *Brain Imaging Behav.* 2021;15(2):526-554. doi:[10.1007/s11682-020-00313-7](https://doi.org/10.1007/s11682-020-00313-7)
5. McInnes K, Friesen CL, MacKenzie DE, Westwood DA, Boe SG. Mild Traumatic Brain Injury (mTBI) and chronic cognitive impairment: A scoping review. *PLoS One.* 2017;12(4). doi:[10.1371/journal.pone.0174847](https://doi.org/10.1371/journal.pone.0174847)
6. Cassidy JD, Carroll LJ, Peloso PM, et al. Incidence, risk factors and prevention of mild traumatic brain injury: Results of the WHO Collaborating Centre Task Force on Mild Traumatic Brain Injury. *J Rehabil Med Suppl.* 2004;(43):28-60. doi:[10.1080/16501960410023732](https://doi.org/10.1080/16501960410023732)
7. Hollis SJ, Stevenson MR, McIntosh AS, Shores A, Collins MW, Taylor CB. Incidence, risk, and protective factors of mild traumatic brain injury in a cohort of Australian nonprofessional male rugby players. *American Journal of Sports Medicine.* 2009;37(12):2328-2333. doi:[10.1177/0363546509341032](https://doi.org/10.1177/0363546509341032)
8. McAllister T, McCrea M. Long-Term cognitive and neuropsychiatric consequences of repetitive concussion and head-impact exposure. *J Athl Train.* 2017;52(3):309-317. doi:[10.4085/1062-6050-52.1.14](https://doi.org/10.4085/1062-6050-52.1.14)
9. Sassani M, Ghafari T, Wannan Arachchige PR, et al. Current and prospective roles of magnetic resonance imaging in mild traumatic brain injury. *Brain Commun.* 2025;7(2):fcf120. doi:[10.1093/braincomms/fcaf120](https://doi.org/10.1093/braincomms/fcaf120)
10. Yue JK, Vassar MJ, Lingsma HF, et al. Transforming research and clinical knowledge in traumatic brain injury pilot: Multicenter implementation of the common data elements for traumatic brain injury. *J Neurotrauma.* 2013;30(22):1831-1844. doi:[10.1089/neu.2013.2970](https://doi.org/10.1089/neu.2013.2970)
11. Bigler ED, Abildskov TJ, Goodrich-Hunsaker NJ, et al. Structural Neuroimaging Findings in Mild Traumatic Brain Injury. *Sports Med Arthrosc Rev.* 2016;24(3):e42-e52. doi:[10.1097/ISA.0000000000000119](https://doi.org/10.1097/ISA.0000000000000119)
12. Wilde EA, Dennis EL, Tate DF. The ENIGMA Brain Injury working group: approach, challenges, and potential benefits. *Brain Imaging Behav.* 2021;15(2):465-474. doi:[10.1007/s11682-021-00450-7](https://doi.org/10.1007/s11682-021-00450-7)
13. Maas AIR, Menon DK, Steyerberg EW, et al. Collaborative European neurotrauma effectiveness research in traumatic brain injury (CENTER-TBI): A prospective longitudinal observational study. *Neurosurgery.* 2015;76(1):67-80. doi:[10.1227/NEU.0000000000000575](https://doi.org/10.1227/NEU.0000000000000575)
14. Broglio SP, McCrea M, McAllister T, et al. A National Study on the Effects of Concussion in Collegiate Athletes and US Military Service Academy Members: The NCAA-DoD Concussion Assessment, Research and Education (CARE) Consortium Structure and Methods. *Sports Medicine.* 2017;47(7):1437-1451. doi:[10.1007/s40279-017-0707-1](https://doi.org/10.1007/s40279-017-0707-1)
15. Tosetti P, Hicks RR, Theriault E, Phillips A, Koroshetz W, Draghia-Akli R. Toward an international initiative for traumatic brain injury research. *J Neurotrauma.* 2013;30(14):1211-1222. doi:[10.1089/neu.2013.2896](https://doi.org/10.1089/neu.2013.2896)
16. Cifu DX. Clinical research findings from the long-term impact of military-relevant brain injury consortium-Chronic Effects of Neurotrauma Consortium (LIMBIC-CENC) 2013-2021. *Brain Inj.* 2022;36(5):587-597. doi:[10.1080/02699052.2022.2033843](https://doi.org/10.1080/02699052.2022.2033843)
17. Jack CR, Bernstein MA, Fox NC, et al. The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. *Journal of Magnetic Resonance Imaging.* 2008;27(4):685-691. doi:[10.1002/jmri.21049](https://doi.org/10.1002/jmri.21049)

18. Walker WC, Carne W, Franke LM, et al. The Chronic Effects of Neurotrauma Consortium (CENC) multi-centre observational study: Description of study and characteristics of early participants. *Brain Inj.* 2016;30(12):1469-1480. doi:[10.1080/02699052.2016.1219061](https://doi.org/10.1080/02699052.2016.1219061)
19. Nencka AS, Meier TB, Wang Y, et al. Stability of MRI metrics in the advanced research core of the NCAA-DoD concussion assessment, research and education (CARE) consortium. *Brain Imaging Behav.* 2018;12(4):1121-1140. doi:[10.1007/s11682-017-9775-y](https://doi.org/10.1007/s11682-017-9775-y)
20. Fortin JP, Cullen N, Sheline YI, et al. Harmonization of cortical thickness measurements across scanners and sites. *Neuroimage.* 2018;167:104-120. doi:[10.1016/j.neuroimage.2017.11.024](https://doi.org/10.1016/j.neuroimage.2017.11.024)
21. Warrington S, Ntata A, Mougin O, et al. A resource for development and comparison of multimodal brain 3 T MRI harmonisation approaches. *Imaging Neuroscience.* 2023;1:1-27. doi:[10.1162/imag_a_00042](https://doi.org/10.1162/imag_a_00042)
22. Fortin JP, Parker D, Tunç B, et al. Harmonization of multi-site diffusion tensor imaging data. *Neuroimage.* 2017;161:149-170. doi:[10.1016/j.neuroimage.2017.08.047](https://doi.org/10.1016/j.neuroimage.2017.08.047)
23. Friedman L, Stern H, Brown GG, et al. Test-retest and between-site reliability in a multicenter fMRI study. *Hum Brain Mapp.* 2008;29(8):958-972. doi:[10.1002/hbm.20440](https://doi.org/10.1002/hbm.20440)
24. Richter S, Winzeck S, Correia MM, et al. Predicting recovery in patients with mild traumatic brain injury and a normal CT using serum biomarkers and diffusion tensor imaging (CENTER-TBI): an observational cohort study. *EClinicalMedicine.* 2024;75. doi:[10.1016/j.eclinm.2024.102751](https://doi.org/10.1016/j.eclinm.2024.102751)
25. Hähnel S, Herweh C. MRI biomarkers in mild traumatic brain injury. *Neurology.* 2015;84(6):554-555. doi:[10.1212/WNL.0000000000001244](https://doi.org/10.1212/WNL.0000000000001244)
26. Li F, Lu L, Li H, et al. Disrupted resting-state functional connectivity and network topology in mild traumatic brain injury: An arterial spin labelling study. *Brain Commun.* 2023;5(5). doi:[10.1093/braincomms/fcad254](https://doi.org/10.1093/braincomms/fcad254)
27. Wright DK, Trezise J, Kamnaksh A, et al. Behavioral, blood, and magnetic resonance imaging biomarkers of experimental mild traumatic brain injury. *Sci Rep.* 2016;6. doi:[10.1038/srep28713](https://doi.org/10.1038/srep28713)
28. Eierud C, Craddock RC, Fletcher S, et al. Neuroimaging after mild traumatic brain injury: Review and meta-analysis. *Neuroimage Clin.* 2014;4:283-294. doi:[10.1016/j.nicl.2013.12.009](https://doi.org/10.1016/j.nicl.2013.12.009)
29. Littlejohns TJ, Holliday J, Gibson LM, et al. The UK Biobank imaging enhancement of 100,000 participants: rationale, data collection, management and future directions. *Nat Commun.* 2020;11:2624. doi:[10.1038/s41467-020-15948-9](https://doi.org/10.1038/s41467-020-15948-9)
30. Harms MP, Somerville LH, Ances BM, et al. Extending the Human Connectome Project across ages: Imaging protocols for the Lifespan Development and Aging projects. *Neuroimage.* 2018;183:972-984. doi:[10.1016/j.neuroimage.2018.09.060](https://doi.org/10.1016/j.neuroimage.2018.09.060)
31. Kurtin DL, Violante IR, Zimmerman K, et al. Investigating the interaction between white matter and brain state on tDCS-induced changes in brain network activity. *Brain Stimul.* 2021;14(5):1261-1270. doi:[10.1016/j.brs.2021.08.004](https://doi.org/10.1016/j.brs.2021.08.004)
32. van der Horn HJ, Vergara VM, Espinoza FA, Calhoun VD, Mayer AR, van der Naalt J. Functional outcome is tied to dynamic brain states after mild to moderate traumatic brain injury. *Hum Brain Mapp.* 2020;41(3):617-631. doi:[10.1002/hbm.24827](https://doi.org/10.1002/hbm.24827)
33. Burley CV, Francis ST, Thomas KN, Whittaker AC, Lucas SJE, Mullinger KJ. Contrasting Measures of Cerebrovascular Reactivity Between MRI and Doppler: A Cross-Sectional Study of Younger and Older Healthy Individuals. *Front Physiol.* 2021;12. doi:[10.3389/fphys.2021.656746](https://doi.org/10.3389/fphys.2021.656746)
34. Li X, Morgan PS, Ashburner J, Smith J, Rorden C. The first step for neuroimaging data analysis: DICOM to NIfTI conversion. *J Neurosci Methods.* 2016;264:47-56. doi:[10.1016/j.jneumeth.2016.03.001](https://doi.org/10.1016/j.jneumeth.2016.03.001)
35. Fischl B. FreeSurfer. *Neuroimage.* 2012;62(2):774-781. doi:[10.1016/j.neuroimage.2012.01.021](https://doi.org/10.1016/j.neuroimage.2012.01.021)
36. Han X, Jovicich J, Salat D, et al. Reliability of MRI-derived measurements of human cerebral cortical thickness: The effects of field strength, scanner upgrade and manufacturer. *Neuroimage.* 2006;32(1):180-194. doi:[10.1016/j.neuroimage.2006.02.051](https://doi.org/10.1016/j.neuroimage.2006.02.051)
37. McGuire SA, Wijtenburg SA, Sherman PM, et al. Reproducibility of quantitative structural and physiological MRI measurements. *Brain Behav.* 2017;7(9). doi:[10.1002/brb3.759](https://doi.org/10.1002/brb3.759)

38. Andersson JLR, Sotiropoulos SN. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *Neuroimage*. 2016;125:1063-1078. doi:[10.1016/j.neuroimage.2015.10.019](https://doi.org/10.1016/j.neuroimage.2015.10.019)
39. Andersson JLR, Skare S, Ashburner J. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *Neuroimage*. 2003;20(2):870-888. doi:[10.1016/S1053-8119\(03\)00336-7](https://doi.org/10.1016/S1053-8119(03)00336-7)
40. Smith SM, Jenkinson M, Johansen-Berg H, et al. Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. *Neuroimage*. 2006;31(4):1487-1505. doi:[10.1016/j.neuroimage.2006.02.024](https://doi.org/10.1016/j.neuroimage.2006.02.024)
41. Jenkinson M, Bannister P, Brady M, Smith S. Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *Neuroimage*. 2002;17(2):825-841. doi:[10.1006/NIMG.2002.1132](https://doi.org/10.1006/NIMG.2002.1132)
42. Jenkinson M, Smith S. A global optimisation method for robust affine registration of brain images. *Med Image Anal*. 2001;5(2):143-156. doi:[10.1016/S1361-8415\(01\)00036-6](https://doi.org/10.1016/S1361-8415(01)00036-6)
43. Andersson JLR, Jenkinson M, Smith S. *Non-Linear Registration Aka Spatial Normalisation FMRIB Technical Report TR07JA2*.; 2007.
44. Chappell MA, Groves AR, Whitcher B, Woolrich MW. Variational Bayesian inference for a nonlinear forward model. *IEEE Transactions on Signal Processing*. 2009;57(1):223-236. doi:[10.1109/TSP.2008.2005752](https://doi.org/10.1109/TSP.2008.2005752)
45. Groves AR, Chappell MA, Woolrich MW. Combined spatial and non-spatial prior for inference on MRI time-series. *Neuroimage*. 2009;45(3):795-809. doi:[10.1016/j.neuroimage.2008.12.027](https://doi.org/10.1016/j.neuroimage.2008.12.027)
46. Woolrich MW, Chiarelli P, Gallichan D, Perthen J, Liu TT. Bayesian inference of hemodynamic changes in functional arterial spin labeling data. *Magn Reson Med*. 2006;56(4):891-906. doi:[10.1002/mrm.21039](https://doi.org/10.1002/mrm.21039)
47. Chappell MA, Groves AR, MacIntosh BJ, Donahue MJ, Jezzard P, Woolrich MW. Partial volume correction of multiple inversion time arterial spin labeling MRI data. *Magn Reson Med*. 2011;65(4):1173-1183. doi:[10.1002/mrm.22641](https://doi.org/10.1002/mrm.22641)
48. Stejskal EO, Tanner JE. Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient. *J Chem Phys*. 1965;42(1):288-292. doi:[10.1063/1.1695690](https://doi.org/10.1063/1.1695690)
49. Marengo S, Rawlings R, Rohde GK, et al. Regional distribution of measurement error in diffusion tensor imaging. *Psychiatry Res Neuroimaging*. 2006;147(1):69-78. doi:[10.1016/j.psychresns.2006.01.008](https://doi.org/10.1016/j.psychresns.2006.01.008)
50. Aronhime S, Calcagno C, Jajamovich GH, et al. DCE-MRI of the liver: Effect of linear and nonlinear conversions on hepatic perfusion quantification and reproducibility. *Journal of Magnetic Resonance Imaging*. 2014;40(1):90-98. doi:[10.1002/jmri.24341](https://doi.org/10.1002/jmri.24341)
51. Luque Laguna PA, Combes AJE, Streffer J, et al. Reproducibility, reliability and variability of FA and MD in the older healthy population: A test-retest multiparametric analysis. *Neuroimage Clin*. 2020;26. doi:[10.1016/j.nicl.2020.102168](https://doi.org/10.1016/j.nicl.2020.102168)
52. Basukala D, Mikheev A, Sevlimedu V, et al. Multisite MRI Intravoxel Incoherent Motion Repeatability and Reproducibility across 3 T Scanners in a Breast Diffusion Phantom: A BREast Intravoxel Incoherent Motion Multisite (BRIMM) Study. *Journal of Magnetic Resonance Imaging*. 2024;59(6):2226-2237. doi:[10.1002/jmri.29008](https://doi.org/10.1002/jmri.29008)
53. Mukaka MM. Statistics Corner: A guide to appropriate use of Correlation coefficient in medical research. *Malawi Med J*. 2012;24(3):69-71.
54. Hui SCN, Mikkelsen M, Zöllner HJ, et al. Frequency drift in MR spectroscopy at 3T. *Neuroimage*. 2021;241:118430. doi:[10.1016/j.NEUROIMAGE.2021.118430](https://doi.org/10.1016/j.NEUROIMAGE.2021.118430)
55. Yendiki A, Koldewyn K, Kakunoori S, Kanwisher N, Fischl B. Spurious group differences due to head motion in a diffusion MRI study. *Neuroimage*. 2014;88:79-90. doi:[10.1016/j.neuroimage.2013.11.027](https://doi.org/10.1016/j.neuroimage.2013.11.027)
56. Andrews DS, Lee JK, Solomon M, Rogers SJ, Amaral DG, Nordahl CW. A diffusion-weighted imaging tract-based spatial statistics study of autism spectrum disorder in preschool-Aged children. *J Neurodev Disord*. 2019;11(1). doi:[10.1186/s11689-019-9291-z](https://doi.org/10.1186/s11689-019-9291-z)
57. Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage*. 2012;59(3):2142-2154. doi:[10.1016/j.neuroimage.2011.10.018](https://doi.org/10.1016/j.neuroimage.2011.10.018)
58. Satterthwaite TD, Elliott MA, Gerraty RT, et al. An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *Neuroimage*. 2013;64(1):240-256. doi:[10.1016/j.neuroimage.2012.08.052](https://doi.org/10.1016/j.neuroimage.2012.08.052)

59. Ye S, Lim JY, Huang W. Statistical considerations for repeatability and reproducibility of quantitative imaging biomarkers. *BJR|Open*. 2022;4(1). doi:[10.1259/bjro.20210083](https://doi.org/10.1259/bjro.20210083)
60. Fujita S, Hagiwara A, Hori M, et al. 3D quantitative synthetic MRI-derived cortical thickness and subcortical brain volumes: Scan-rescan repeatability and comparison with conventional T1-weighted images. *Journal of Magnetic Resonance Imaging*. 2019;50(6):1834-1842. doi:[10.1002/jmri.26744](https://doi.org/10.1002/jmri.26744)
61. Vollmar C, O'Muircheartaigh J, Barker GJ, et al. Identical, but not the same: Intra-site and inter-site reproducibility of fractional anisotropy measures on two 3.0T scanners. *Neuroimage*. 2010;51(4):1384-1394. doi:[10.1016/j.neuroimage.2010.03.046](https://doi.org/10.1016/j.neuroimage.2010.03.046)
62. Grech-Sollars M, Hales PW, Miyazaki K, et al. Multi-centre reproducibility of diffusion MRI parameters for clinical sequences in the brain. *NMR Biomed*. 2015;28(4):468-485. doi:[10.1002/nbm.3269](https://doi.org/10.1002/nbm.3269)
63. Veenith TV, Mada M, Carter E, et al. Comparison of inter subject variability and reproducibility of whole brain proton spectroscopy. *PLoS One*. 2014;9(12). doi:[10.1371/journal.pone.0115304](https://doi.org/10.1371/journal.pone.0115304)
64. Jones DK, Alexander DC, Bowtell R, et al. Microstructural imaging of the human brain with a "super-scanner": 10 key advantages of ultra-strong gradients for diffusion MRI. *Neuroimage*. 2018;182:8-38. doi:[10.1016/j.neuroimage.2018.05.047](https://doi.org/10.1016/j.neuroimage.2018.05.047)
65. Mukherjee P, Chung SW, Berman JI, Hess CP, Henry RG. Diffusion tensor MR imaging and fiber tractography: Technical considerations. *American Journal of Neuroradiology*. 2008;29:843-852. doi:[10.3174/ajnr.A1052](https://doi.org/10.3174/ajnr.A1052)
66. Triantafyllou C, Polimeni JR, Wald LL. Physiological noise and signal-to-noise ratio in fMRI with multi-channel array coils. *Neuroimage*. 2011;55(2):597-606. doi:[10.1016/j.neuroimage.2010.11.084](https://doi.org/10.1016/j.neuroimage.2010.11.084)
67. Pruessmann KP, Weiger M, Scheidegger MB, Boesiger P. SENSE: Sensitivity encoding for fast MRI. *Magn Reson Med*. 1999;42(5):952-962. doi:[10.1002/\(SICI\)1522-2594\(199911\)42:5](https://doi.org/10.1002/(SICI)1522-2594(199911)42:5)
68. Griswold MA, Jakob PM, Heidemann RM, et al. Generalized Autocalibrating Partially Parallel Acquisitions (GRAPPA). *Magn Reson Med*. 2002;47(6):1202-1210. doi:[10.1002/mrm.10171](https://doi.org/10.1002/mrm.10171)
69. Feinberg DA, Setsompop K. Ultra-Fast MRI of the Human Brain with Simultaneous Multi-Slice Imaging. *Journal of Magnetic Resonance*. 2013;229:90-100. doi:[10.1016/j.jmr.2013.02.002](https://doi.org/10.1016/j.jmr.2013.02.002)
70. Jovicich J, Minati L, Marizzoni M, et al. Longitudinal reproducibility of default-mode network connectivity in healthy elderly participants: A multicentric resting-state fMRI study. *Neuroimage*. 2016;124:442-454. doi:[10.1016/j.neuroimage.2015.07.010](https://doi.org/10.1016/j.neuroimage.2015.07.010)
71. Schmiedeskamp H, Newbould RD, Pisani LJ, et al. Improvements in parallel imaging accelerated functional MRI using multiecho echo-planar imaging. *Magn Reson Med*. 2010;63(4):959-969. doi:[10.1002/mrm.22222](https://doi.org/10.1002/mrm.22222)
72. Kayvanrad A, Arnott SR, Churchill N, et al. Resting state fMRI scanner instabilities revealed by longitudinal phantom scans in a multi-center study. *Neuroimage*. 2021;237:118197. doi:[10.1016/j.NEUROIMAGE.2021.118197](https://doi.org/10.1016/j.NEUROIMAGE.2021.118197)
73. Kruger G, Glover GH. Physiological Noise in Oxygenation-Sensitive Magnetic Resonance Imaging. *Magn Reson Med*. 2001;46(4):631-637. doi:[10.1002/mrm.1240](https://doi.org/10.1002/mrm.1240)
74. Yang HC, Liang Z, Yao J, Shen X, Frederick BD, Tong Y. Vascular effects of caffeine found in BOLD fMRI. *J Neurosci Res*. 2019;97(4):456-466. doi:[10.1002/jnr.24360](https://doi.org/10.1002/jnr.24360)
75. Brody AL. Functional brain imaging of tobacco use and dependence. *J Psychiatr Res*. 2006;40(5):404-418. doi:[10.1016/j.jpsychires.2005.04.012](https://doi.org/10.1016/j.jpsychires.2005.04.012)
76. Dubol M, Epperson CN, Sacher J, et al. Neuroimaging the menstrual cycle: A multimodal systematic review. *Front Neuroendocrinol*. 2021;60(January). doi:[10.1016/j.yfrne.2020.100878](https://doi.org/10.1016/j.yfrne.2020.100878)
77. Gountouna VE, Job DE, McIntosh AM, et al. Functional Magnetic Resonance Imaging (fMRI) reproducibility and variance components across visits and scanning sites with a finger tapping task. *Neuroimage*. 2010;49(1):552-560. doi:[10.1016/j.neuroimage.2009.07.026](https://doi.org/10.1016/j.neuroimage.2009.07.026)
78. Gevers S, Van Osch MJ, Bokkers RPH, et al. Intra- and multicenter reproducibility of pulsed, continuous and pseudo-continuous arterial spin labeling methods for measuring cerebral perfusion. *Journal of Cerebral Blood Flow and Metabolism*. 2011;31(8):1706-1715. doi:[10.1038/jcbfm.2011.10](https://doi.org/10.1038/jcbfm.2011.10)

79. Petersen ET, Mouridsen K, Golay X. The QUASAR reproducibility study, Part II: Results from a multi-center Arterial Spin Labeling test-retest study. *Neuroimage*. 2010;49(1):104-113. doi:[10.1016/j.neuroimage.2009.07.068](https://doi.org/10.1016/j.neuroimage.2009.07.068)
80. Van Der Thiel M, Rodriguez C, Giannakopoulos P, et al. Brain perfusion measurements using multidelay arterial spin-labeling are systematically biased by the number of delays. *American Journal of Neuroradiology*. 2018;39(8):1432-1438. doi:[10.3174/ajnr.A5717](https://doi.org/10.3174/ajnr.A5717)
81. Mutsaerts HJMM, van Osch MJP, Zelaya FO, et al. Multi-vendor reliability of arterial spin labeling perfusion MRI using a near-identical sequence: Implications for multi-center studies. *Neuroimage*. 2015;113:143-152. doi:[10.1016/j.neuroimage.2015.03.043](https://doi.org/10.1016/j.neuroimage.2015.03.043)
82. Morgan CA, Roberts RP, Chaffey T, et al. Reproducibility and repeatability of magnetic resonance imaging in dementia. *Physica Medica*. 2022;101(July):8-17. doi:[10.1016/j.ejmp.2022.06.012](https://doi.org/10.1016/j.ejmp.2022.06.012)
83. Gaggi NL, Ware JB, Dolui S, et al. Temporal dynamics of cerebral blood flow during the first year after moderate-severe traumatic brain injury: A longitudinal perfusion MRI study. *Neuroimage Clin*. 2023;37. doi:[10.1016/j.nicl.2023.103344](https://doi.org/10.1016/j.nicl.2023.103344)
84. Boland M, Stirnberg R, Pracht ED, et al. Accelerated 3D-GRASE imaging improves quantitative multiple post labeling delay arterial spin labeling. *Magn Reson Med*. 2018;80(6):2475-2484. doi:[10.1002/mrm.27226](https://doi.org/10.1002/mrm.27226)
85. Jahanian H, Noll DC, Hernandez-Garcia L. B 0 field inhomogeneity considerations in pseudo-continuous arterial spin labeling (pCASL): Effects on tagging efficiency and correction strategy. *NMR Biomed*. 2011;24(10):1202-1209. doi:[10.1002/nbm.1675](https://doi.org/10.1002/nbm.1675)
86. Pinto J, Chappell MA, Okell TW, et al. Calibration of arterial spin labeling data—potential pitfalls in post-processing. *Magn Reson Med*. 2019;83(4):1222-1234. doi:[10.1002/mrm.28000](https://doi.org/10.1002/mrm.28000)
87. Jovicich J, Czanner S, Greve D, et al. Reliability in multi-site structural MRI studies: Effects of gradient non-linearity correction on phantom and human data. *Neuroimage*. 2006;30(2):436-443. doi:[10.1016/j.neuroimage.2005.09.046](https://doi.org/10.1016/j.neuroimage.2005.09.046)
88. Dijsselhof MBJ, Moore C, Amiri S, et al. Multi-cohort, multi-sequence harmonisation for cerebrovascular brain age. *Imaging Neuroscience*. 2025;3:IMAG.a.964. doi:[10.1162/IMAG.a.964](https://doi.org/10.1162/IMAG.a.964)
89. Song J, Desphande AS, Meier TB, et al. Age-Related Differences in Test-Retest Reliability in Resting-State Brain Functional Connectivity. *PLoS One*. 2012;7(12):e49847. doi:[10.1371/journal.pone.0049847](https://doi.org/10.1371/journal.pone.0049847)
90. Cieri F, Zhuang X, Cordes D, Kaplan N, Cummings J, Caldwell J. Relationship of sex differences in cortical thickness and memory among cognitively healthy subjects and individuals with mild cognitive impairment and Alzheimer disease. *Alzheimers Res Ther*. 2022;14(1). doi:[10.1186/s13195-022-00973-1](https://doi.org/10.1186/s13195-022-00973-1)

SUPPLEMENTARY MATERIALS

Supplementary Material

Download: <https://apertureneuro.org/article/157692-assessing-optimisation-of-3t-mri-protocol-across-multi-site-multi-vendor-settings-for-studying-mild-traumatic-brain-injury-mtbi-predict-consortium-s/attachment/330834.docx>
