Pooling quantitative MRI data: A multi-protocol study of healthy subcortical ageing

Mikhail Zubkov[1&], Kerrin J Pine[2], Pierre-Louis Bazin[3], Puneet Talwar[1], Nasrin Mortazavi[1], Solène Dauby[1], Chloé Geron[1], Elise Beckers[1], Laurent Lamalle[1], Christophe Phillips[1], Fabienne Collette[1], Pierre Maquet[1], Emilie Lommers[4], Anneke Alkemade[5], Nikolaus Weiskopf[2,6,7], Gilles Vandewalle[1*], Evgeniya Kirilina[2*]

[1]Sleep and Health Lab, GIGA-Institute, CRC-Human Imaging Unit, University of Liège (ULiège), Liège, Belgium

[2]Department of Neurophysics, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

[3]Full brain picture Analytics, Leiden, Netherlands

[4]Clinical Neurosciences Lab, GIGA-Institute, CRC-Human Imaging Unit, University of Liège (ULiège), Liège, Belgium

[5]Integrative Model-based Cognitive Neuroscience Research Unit, University of Amsterdam, Amsterdam, Netherlands

[6]Felix Bloch Institute for Solid State Physics, Faculty of Physics and Earth System Sciences, Leipzig University, Leipzig, Germany

[7]Wellcome Centre for Human Neuroimaging, Institute of Neurology, University College London, London, United Kingdom

& corresponding author

* Shared senior authorship

Keywords

Ultra-high field, quantitative imaging, iron, myelin, subcortex, segmentation, ageing

Abstract

Quantitative MRI (qMRI) measures relaxation rates, exchange rates and proton densities that reflect the biophysical properties of tissue and are ideally free from protocol- and scanner-dependence. In practice, qMRI has not yet achieved this level of independence from sequence and hardware choice, and quantitative measurements often differ across sites and acquisition schemes. At the same time pooling data across different sources can be beneficial to statistical power of longitudinal, cross-sectional or case-control studies. Here we investigate how protocol and hardware differences can affect pooling data from difference sources in large ultra high field (UHF) qMRI studies in the context of healthy aging. We combine the openly available ageing UHF qMRI MP2RAGEME-based dataset with two different MPM-based sets of qMRI data. We evaluate how pooling affects age dependence of qMRI parameters and investigate protocol-related biases, with a particular focus on subcortical structures. We focus the analysis, first, on replication and expansion of the reference qMRI dataset on normative aging, second, the examination of the protocol influence on the measured qMRI values, and third, on detecting the protocol effect on the age dependence inferred from the data. We find that the age-related changes for R1, R2* and volume detected by different protocols were of the same order of magnitude supporting the idea that age-related inter-individual variability can be correctly grasped across sites and protocols using qMRI approaches. We further observe larger relative difference across protocols for R1 and

volume, while R2* remains more consistent for most regions. We therefore provide assessment of the pooling effects in the ultra-high resolution UHF qMRI data performed over a relatively large participant cohort, comprising data from different sites collected with different quantitative protocols.

Introduction

Ultra-high field (UHF, i.e. at 7 T and above) quantitative MRI (qMRI) with its ability to provide ultra-high resolution quantitative maps of brain anatomy, is progressively entering both the clinical and research applications, offering promising advances for brain imaging. Recent developments in parallel radio frequency (RF) transmission techniques (pTx), have mitigated biases caused by transmit field inhomogeneity inherent to UHF MRI (1), substantially advancing its applications. Owing to these advances, elusive changes in smaller subcortical nuclei as well as details of cortical laminar structures can now be detected non-invasively (2,3). This enables neuroscientists and clinicians to better characterize their anatomy in healthy population, and to identify subtle alternation associated with brain disorders and neurodegenerative diseases (4,5).
qMRI measures absolute spin metrics (i.e. relaxation rates, exchange rates and proton densities) that reflects the biophysical properties of tissue and are ideally free from protocol- and scanner-dependence (6). By linking these qMRI parameters to tissue microstructure via biophysical models, qMRI provides insights into the tissue microarchitecture beyond the nominal resolution of acquired data. In particular, the longitudinal relaxation rates (R1) and the effective transverse relaxation rate (R2*) have been shown to yield unique information about the tissue macromolecular and iron content (7–10). The combination of UHF and qMRI is therefore very promising for enabling true *in vivo* histology MRI (hMRI) of small brain structures, such as cortical layers and subcortical nuclei. This creates opportunity for a new generation of early-stage, non-invasive MRI biomarkers of neurodegeneration, as small subcortical nuclei are often affected in early pre-symptomatic disease stages. Materializing this potential, however, requires normative populational datasets that capture the full variability of anatomy and microstructure across the healthy adult lifespan. Such studies are pivotal for defining quantitative references values, discriminating between normal and pathologic aging, and setting the basis of clinical diagnostics.

Several approaches for multiparametric qMRI have been developed. Most frequently one of the following two is used. Multi-echo magnetization-prepared rapid gradient echo (MPRAGEME or MP2RAGEME), based on application of inversion pulse followed by rapid gradient echo readouts, yields (11) the quantitative maps of R1 and R2* simultaneously. Alternatively, variable flip angle protocols, such as multiparameter mapping (MPM), prompt to acquire multiple datasets with different sets of excitation angles to provide (12) R1 and R2* maps. Ideally, these qMRI techniques should yield consistent results across protocols and scan sites, while preserving the sensitivity to microstructural and gross anatomical details of individual participants.

In practice, however, UHF qMRI has not yet achieved full independence from protocol- and hardware, and quantitative measurements often differ across sites and acquisition schemes (13,14). As an indirect consequence, only a few studies have used qMRI on large participant cohorts. Those which did, were able to infer the changes of brain microstructure in ageing (3,15–17), in depressive disorder (18), in multiple sclerosis (19), and in COVID-19 (20). Large cohort studies of aging have consistently identified inverted U-shape age dependence of R1, R2*, and brain volume across multiple brain regions, reflecting protracted brain plasticity up to late adulthood followed up by age-related decline. Disease-oriented studies

usually rely on normative values by monitoring the parameter deviations in selected brain structures as indicator of pathology. To create such normative large representative datasets, pooling data across different sites and protocols is highly beneficial, as it increases statistical power of longitudinal, cross-sectional or case-control studies. However, the lack of unified mapping protocols in UHF qMRI and the consequential lack of consensus on normal values of tissue parameters reduces the prognostic utility of the method, prohibiting introduction of cutoff values for disorder diagnostics as inter-individual (3,21) and between protocol differences cannot be properly accounted for.

Encouragingly, recent studies (22–24) have demonstrated that when site differences are taken into consideration and protocol consistency is ensured, the qMRI measures become generalizable at the participant level. In other words, scanning the same person across different sites with the same protocol yields similar outputs. Building on this finding, here, we investigate how protocol and hardware differences can affect pooling data of different individuals in large qMRI studies in the context of healthy aging. We combine the openly available ageing UHF qMRI reference dataset (25) with a heterogeneous (in terms of the protocol in use) locally-acquired data. Specifically, we evaluate how pooling affects age dependence of qMRI parameters and investigate protocol-related biases, with a particular focus on subcortical structures, whose reliable delineation is almost exclusive to UHF MRI due to their small size and segmentation challenges.

Methods

Participant cohort

The sample used in this study comprises three cohorts of healthy control participants with ages covering the human adult lifespan and originates from three different sources, referred here as AHEAD, sTx-MPM and pTx-MPM. The AHEAD study serves as a reference data sourced from a publicly available dataset (25) acquired with the MP2RAGEME sequence on a Philips Achieva 7 T (Philips Healthcare, Eindhoven, Netherlands) scanner at the University of Amsterdam. AHEAD contains 105 participants (60 female, 45 male) aged from 19 to 80 years. The other two cohorts' data sources are the studies conducted locally at the GIGA-CRC Human Imaging of the Université de Liège on a 7 T Siemens Terra (Siemens Healthineers, Erlangen, Germany) scanner and are not yet publicly available. This locally-sourced data consists of two parts making use of different versions of the MPM approach: one labelled as sTx-MPM in

Table 1 and another labelled as pTx-MPM in

Table 1. Both versions employ a multi-echo (ME) magnetization transfer (MT) enabled Fast Low Angle Shot (FLASH)-based magnetic resonance imaging protocol, but with variations in scan parameters and hardware (see Table 2 and a detailed description below). All of the data acquisitions have been approved by the local ethics committee. The exclusion criteria for the participants were history of major neurologic/psychiatric diseases or stroke; recent history of depression/anxiety; sleep disorders; medication affecting the central nervous system; smoking, excessive alcohol (>14 units/week) or caffeine (>5 cups/day) consumption; night shift work in the past 6 months; BMI ≤18 and ≥29 (for older individuals) and ≥25 (for younger individuals). The separate statistics for the studies are presented in

Table 1. The total local MPM sample comprised 135 participants (76 female, 59 male) aged from 18 to 75 years. As described in detail below the data analysis employed the pooled dataset (combining both local and the AHEAD data), producing a combined dataset counting 240 participants with further details provided in

Table 1.

Table 1 Age and sex distribution of the subsamples comprising the study cohort

Study	Sex (female/male)	Age (median/iqr/min/max), years
AHEAD	60/45	38/37/19/80
sTx-MPM	53/32	53/41/18/75
pTx-MPM	23/27	38/42/20/73
Total (local)	76/59	47/41/18/75
Combined	136/104	45/41/18/80

Table 2 Major features of the reference and local acquisitions. The full parameters for MT-FLASH-ME acquisition sequences, as well as for Spin Echo/Stimulated Echo Echo-Planar Imaging (SE/STE-EPI) based field mapping and Actual Flip Angle (AFI) sequence are provided in the Data acquisition section.

	AHEAD	sTx-MPM	pTx-MPM
Scanner	Philips Achieva 7T	7T Siemens Terra	7T Siemens Terra
Coil	1Tx/32Rx Nova Coil	1Tx/32Rx Nova Coil	8Tx/32Rx Nova Coil
Sequence	MP2RAGEME	MT-FLASH-ME	MT-FLASH-ME
Number of echoes	4	4 and 6	4 and 6
Resolution	<0.7 mm	0.6 mm	0.6 mm
$B1^+$ mapping	none	SE/STE-EPI	AFI
Metrics extracted	R1, R2*, χ	R1, R2*, PD, MTsat	R1, R2*, PD, MTsat
qMRI tool	nighres, tgv-qsm	hMRI-toolbox 0.6.1	hMRI-toolbox 0.6.1

Data acquisition

AHEAD

The full acquisition protocol details of the AHEAD data are described in the original publication (25). Briefly, the study made use of a single transmit 1-Tx/32-Rx RF head coil (Nova Medical, Wilmington, MA, USA) and employed an MP2RAGEME acquisition (11) with two inversion times (TI1/TI2 670/3675.4 ms) and four unequally spaced echoes (echo time (TE) of 3/11.5/19/28.5 ms). The scan covered the Field of View (FoV) of 205×205×164 mm with a 0.64×0.64×0.7 mm nominal resolution. The acquisition was accelerated using the parallel imaging using SENSitivity Encoding (SENSE) technique (26) with the acceleration factor of two in one phase encoding direction. No transmit B1 field inhomogeneity correction was performed.

sTx-MPM

Acquisition of the data labelled sTx-MPM employed scanning with a 1-Tx/32-Rx head coil (Nova Medical, Wilmington, MA, USA) and an MPM protocol (12) comprising three (T1-weighted, PD-weighted, MT-weighted) whole-brain 3D FLASH-based multi-echo sequences covering the FoV of 256×218×173 mm with isotropic resolution of 0.6 mm (repetition time (TR) 19.5 ms, flip angles (FA) PDw/MTw/T1w 5/5/20°, six equispaced echoes with TE ranging from 2.3 to 14.2 ms for PDw and T1w and four equispaced echoes with TE ranging from 2.3 ms to 9.44 ms for MTw (the latter achieved by applying a 4 ms, 140° Gaussian-shaped pulse, 2 kHz off-resonance), parallel imaging using GeneRalized Autocalibrating Partial Parallel Acquisition (GRAPPA) approach (27) was used with acceleration factor of

two in both phase encoding directions) and 3D spin-echo -stimulated-echo echo planar imaging (SE-STE-EPI) based transmit B1 field mapping for transmit field correction (28). The field mapping protocol was acquired at an isotropic 4 mm resolution (TR 500 ms, TE 43.18 ms, and 15 gradually decreasing nominal pulse flip angles form 165° to 60°, GRAPPA acceleration in both phase encoding directions with acceleration factor R=2 in each). A dual-echo gradient echo sequence with 3 mm isotropic resolution (TR 677 ms, FA 39°, TE 5 ms and 6.02 ms) was acquired for B_0 mapping employed in EPI distortion correction.

pTx-MPM

Acquisition of the data labelled pTx-MPM adhered to a protocol making use of the parallel transmission using a 8-Tx/32-Rx RF head coil (Nova Medical, Wilmington, MA, USA) and an MPM protocol comprising three (T1-weighted, PD-weighted, MT-weighted) whole-brain 3D FLASH-based multi-echo sequences covering the FoV of 250×218×173 mm with isotropic resolution of 0.6 mm (TR 22.4 ms, FA PDw/MTw/T1w 7/7/22°, six equispaced echoes with TE ranging from 3 ms to 15.6 ms for PDw and T1w and 4 equispaced echoes with TE ranging from 3 ms to 10.56 ms for MTw (the latter achieved by applying a 4 ms, 130° Gaussian-shaped pulse, 3 kHz off-resonance), Controlled Aliasing In Parallel Imaging Resulting IN Higher Acceleration (CAIPIRINHA) acceleration (29) in two phase encoding directions with acceleration factor R=2 in each, kT-points pulses (30) for RF homogeneity improvement and AFI based B1 mapping for transmit field correction (31). The field mapping protocol was acquired at an isotropic 4 mm resolution (two TR values of 25 ms and 125 ms, TE 2.12 ms, nominal pulse flip angle of 55°, GRAPPA acceleration in one phase encoding direction with an acceleration factor R=2). Two field mappings were performed per protocol to account for different $B1^+$ distributions when using the kT-points excitation pulses and the Gaussian-shaped MT-pulses.

To test if different field mapping procedures in sTx-MPM and pTx-MPM could be a source of bias and error, three participants from the pTx-MPM cohort also had the SE-STE-EPI based field mapping with the parameters matched to the field mapping from the sTx-MPM added to their scan protocols.

Data pipelines

The reference AHEAD data processing pipeline for quantitative parameter extraction is described in the original publication (3). Briefly, it comprises the LCPCA-based denoising (32), standard MP2RAGE calculation of the T1 relaxation time (33), single-exponential fitting of the multi-echo data for the R2* estimation, and TGV-QSM (34) for QSM. The output quantitative maps (R1, R2*, and χ) were then used in the Multi-contrast Anatomical Subcortical Structures Parcellation (MASSP) (35) Bayesian segmentation algorithm based on anatomically-derived priors and contrast-conditional posteriors yielding masks of 14 subcortical structures (left/right averaged).

The locally acquired data from sTx-MPM and pTx-MPM studies was processed through two similar pipelines, which included LCPCA-based denoising of the anatomical magnitude images, hMRI-toolbox (hmri.info) based quantitative parameter maps calculation (36), employing corresponding incomplete spoiling correction, calculated for each pulse sequence, and two-iteration weighted least squares R2* fitting. The final processing step comprised MASSP-based segmentation, as the latter has demonstrated reliable and reproducible segmentation in both AHEAD data and dedicated reproducibility study (22).

The difference in processing the sTx-MPM and the pTx-MPM data was, first, in different models for B1 maps calculation, second, in different inputs for MASSP parcellation, and third in employing different B1 inhomogeneity corrections for the MT-pulses, the latter of least importance due to MTsat maps being unused in this study. The qMRI mapping yielded the PD, R1, R2*, and MTsat maps, with the first three having been used for the MASSP

parcellation (and the MTsat, discarded), that provided segmentation of the same 14 subcortical nuclei (left/right hemisphere averaged, see Figure 1 for details) as in the AHEAD data. Aside from the difference in acquisition protocols, a crucial data-wise difference between AHEAD and local data is that for local scans only magnitude data were exported, thus no QSM data was reconstructed neither for the sTx-MPM nor pTx-MPM protocols. On the other hand, recovering the PD metric allowed to incorporate it into the segmentation procedure in place of QSM.

Two additional analyses were performed in order to estimate the biases resulting from different B_1^+ mapping methods and different B_1^+ field settings employed in sTx and pTx acquisitions. To estimate the effect of different mapping on qMRI parameters, the B_1^+ field mapping data for three participants having both types (AFI and SE-STE-EPI) of field mapping was used to calculate the relative flip angle maps (in percent units of nominal flip angle) and was co-registered to anatomical data using hMRI toolbox. For these participants, joint histogram of relative flip angle values measured by two methods was plotted, and linear model assessing the ratio between the two measurement methods was fit to the data.

To estimate the effect of different B_1^+settings in in sTx and pTx acquisitions in sTx and pTx acquisitions the corresponding relative flip angle maps for all other participants were used in conjunction with MASSP-based segmentations to calculate the mean relative flip angle and the coefficient of variation (CoV) of the relative flip angle across the 14 subcortical structures.

Quality assurance

Potential errors in segmentation and qMRI maps calculation were assessed by plotting the ratio of non-unique R2* values per segmented structure volume for all ROIs and participants. Ideally, due to continuous distribution of R2* values, all voxels should possess unique values, and the number of non-unique values (i.e., values equal at the digital number format precision) should be zero. Systematic errors, e.g., convergence of the fitting procedure to boundary values may produce duplicated values, which can be identified via calculating of the ratio of non-unique values to the ROI volume. Another source of systematic error is the low-voxel count ROIs: these can be generated due to the segmentation errors, and may produce outlier values, while being undetected by the non-unique value count procedure (e.g., a single-voxel ROI has no duplicated values, but is likely to produce outliers). Shannon entropy (37) allows assessing the amount of information the ROI can provide, and favors larger ROIs with non-repeated values. Practically, entropy can be easily thresholded, e.g., excluding values less than 3 bits, thus effectively removing areas less than 8 voxels (particularly, single-voxel segmentations with zero entropy), as well as larger areas with sufficiently frequent value repetition – a common sign of boundary value convergence.

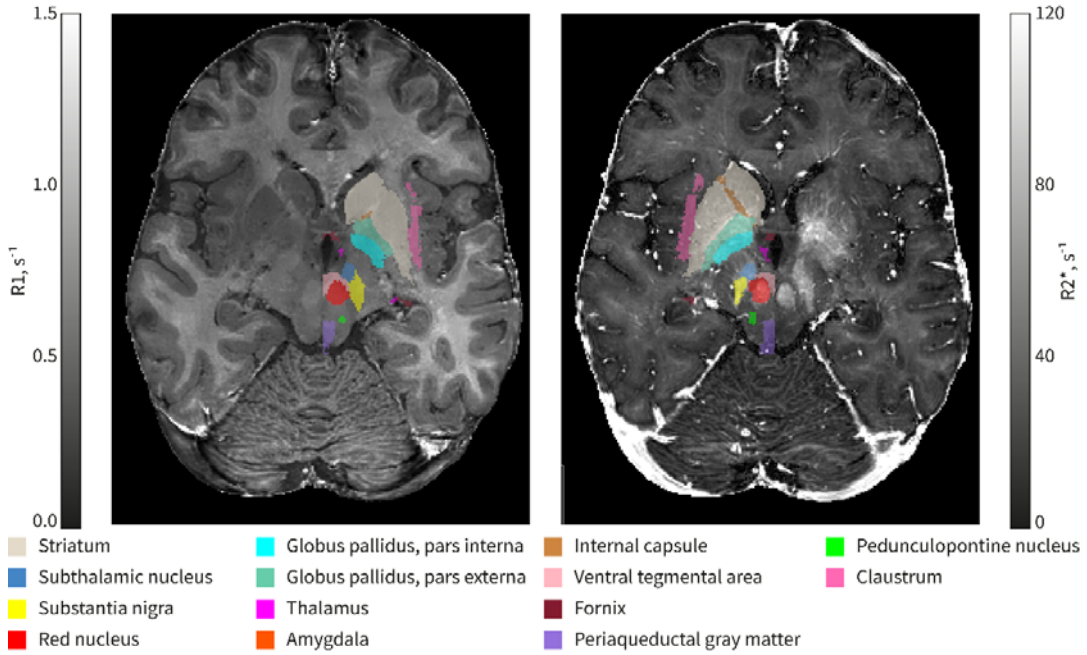

Figure 1 Segmentation example of individual dataset (pTx-MPM, f, 46 years) showing a set of investigated 14 subcortical structures overlaid on the pTx-MPM based R1 parametric map (left) and R2 parametric map (right). Segmentation in only one hemisphere is shown for each map.*

Statistics

Median values for each of the calculated quantitative parameters were extracted for the 14 subcortical structures (average of left and right) with the ventricular system excluded from further analysis. The statistical analysis described below was carried out separately for each of the three different qMRI metrics (R1, R2*, and ROI volume) using the combined cohort comprising 240 healthy participants.

The statistical analysis focused on three goals. First, a replication and expansion of the reference qMRI dataset on normative aging, i.e., the inference of life-span trajectory of subcortical anatomy and microstructure in healthy aging, as measured with UHF qMRI. Second, the examination of the protocol influence on the measured qMRI values (i.e., the detection of protocol-related absolute biases in age-corrected median values of qMRI parameters), and third, the analysis of the protocol effect on the age dependence inferred from the data (i.e., the search for protocol effect on the measured aging trajectory). To achieve these three objectives, a linear model was built in R 4.4.0 for each of the median qMRI metrics with age, age^2, and sex being the regressors (with the second power of age included following the results in (3,15,16)). In order to assess the impact of different protocols on the measured median values, an additional categorical parameter related to the protocol was incorporated into the models with three categories indicating either sTx-MPM, pTx-MPM, or AHEAD data. Furthermore, interaction between the protocol-related term and the age-related variables were considered.

The presence of age and protocol dependence was assessed via various F-tests over a set of models. The set of F-tests to detect age dependence (goal 1) compared models with age, age^2, sex, and protocol type variables against models with only sex, and protocol type variables. The bias detection (goal 2) was conducted via an F-test of the model with age, age^2, sex, and protocol type variables against the model without the and protocol type variable. To detect the influence of protocol selection on age dependence (goal 3) models with age, age^2, sex, and protocol type variables were tested against models with added interactions between age, age^2 and protocol type variables. Each of the qMRI parameters was treated as a distinct

measurement and multiple comparison correction was applied across different subcortical structures separately for each extracted qMRI metric and for each type of test. The effects with F-test p-values lower than 0.05 threshold (Benjamini–Hochberg false discovery rate-corrected (38)) were deemed significant.

To further assess the influence of the protocol on the estimated qMRI parameters, first, the age trajectories were used to calculate the variation of the qMRI parameters with age as the difference of the maximum and minimum values within the studied age range. These variation estimates were then used to assess how different were the protocol-specific quantitative values from the reference AHEAD protocol as compared to the age-related variation of each quantitative parameter estimated using sTx-MPM and pTx-MPM datasets.

Age trajectories of R1 reflecting the processes of myelination in young adulthood and demyelination in older age were used to calculate the age where myelination reached its maximum for each segmented brain region.

Results

Acquisition and segmentation

The combination of the sTx-MPM and pTx-MPM datasets (see Table 1) resulted in a 135-participant dataset with age distribution similar between all three (sTx-MPM, pTx-MPM and AHEAD) cohorts (Figure 2).

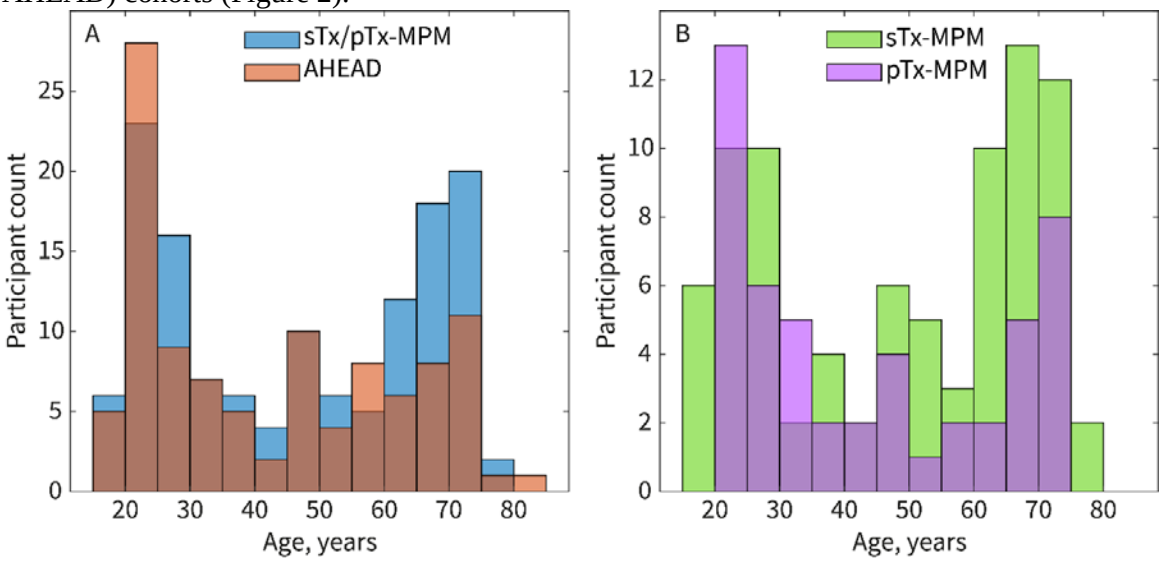

Figure 2 A. Age distribution in the AHEAD and local (sTx-MPM and pTx-MPM)) datasets. The similarity of the age profiles justifies employing similar approaches to data analysis as in (3) and allows comparing age dependence. B. Age distribution in the two parts of the local dataset. Again, the distribution similarity allows pooling without risk of causing age biases by particular cohort subset.

High quality, 0.6 mm resolution, co-aligned maps of qMRI parameters R2, PD, R1 were obtained in all participants of the combined studies. The resolution of the parametric maps and the contrast in cortical and subcortical anatomical structures allowed MASSP parcellations (35) of the subcortical structures highlighted in Figure 1 in each participant. The quality*

assurance procedure has shown all the qMRI calculations and segmentation to perform well (

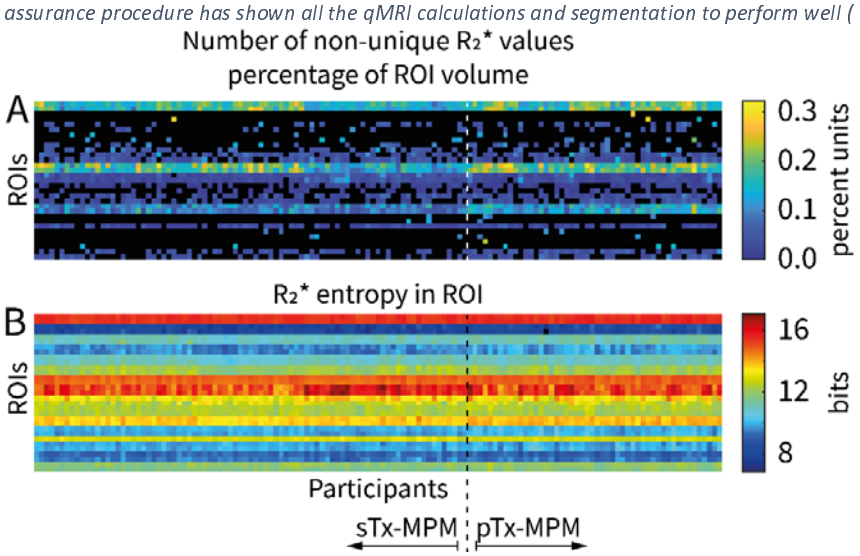

Figure 3). Particularly, no regions exhibited higher than 0.3% of non-unique values per ROI volume. No regions with Shannon entropy (37) values less than 6 bits were observed, thus allowing to include all of the segmented regions into analysis and suggesting substituting χ with PD measures did not dramatically degrade MASSP segmentation quality.

Data extracted from a number of representative ROIs (Striatum, Globus Pallidus and Thalamus) are provided in Figure 4. Data for all the 14 parcellated regions are provided in supplementary Figures S1 to S6.

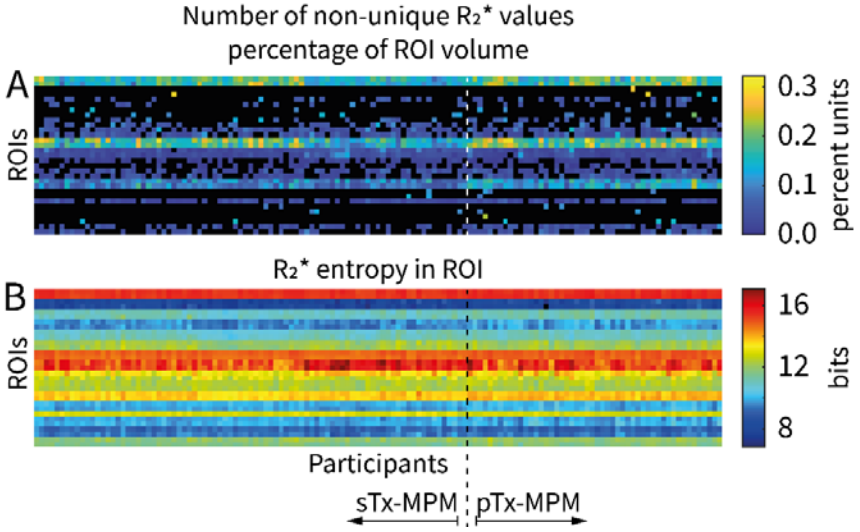

Figure 3 Quality assurance results for the locally acquired data. Note, that the number of ROI is higher than the one described in the Methods section due to the quality metrics being extracted prior to averaging left and right structures. A. The overall low percentage (well below 1%) of non-unique values in all ROIs indicates lack of convergence of qMRI metric calculation to boundary values. B. High entropy (above 6 bits) in all the ROIs indicates lack of low-voxel count ROIs and confirms lack of convergence to boundary values.

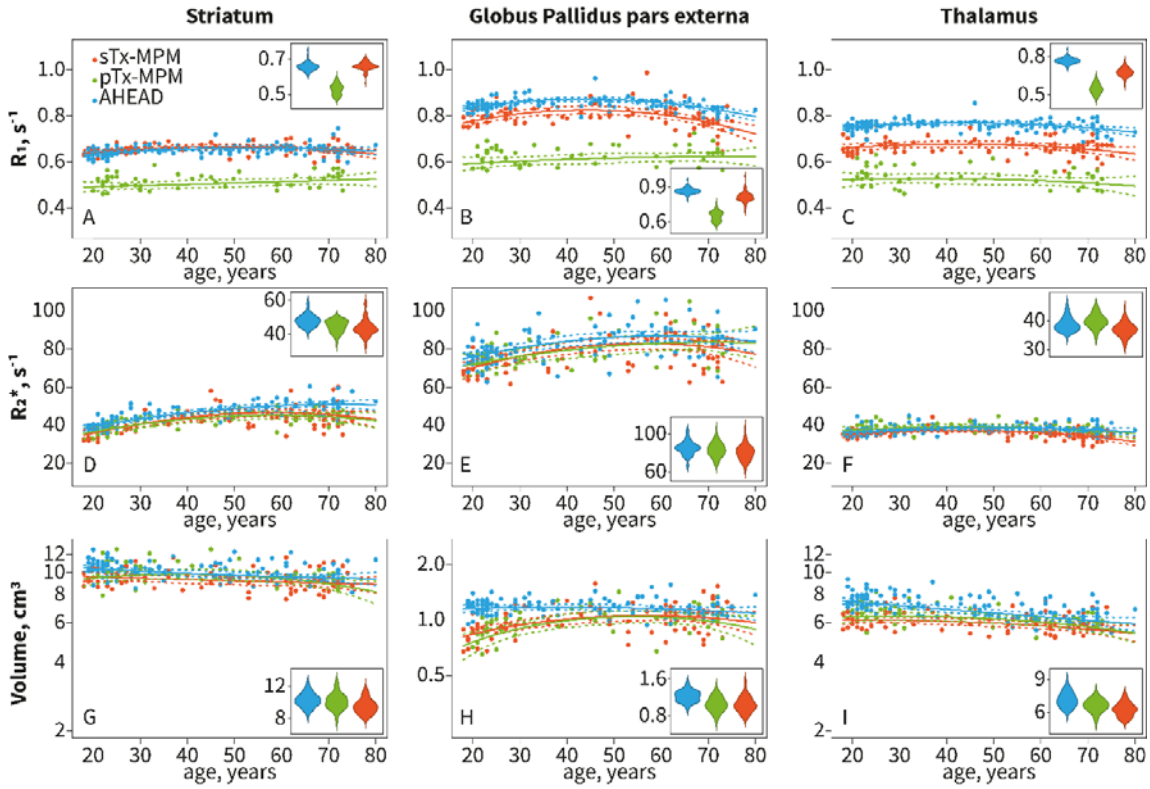

Figure 4 Examples of R1 (A, B, C), R2 (D, E, F) and ROI volume (G, H, I) age dependence in Striatum (A, D, G), Globus Pallidus pars externa (B, E, H), and Thalamus (C, F, I). Each graph also shows models that were fit for individual protocols (not used in statistical analysis). Insets show age-adjusted data distribution across different protocols, where all participant data has been adjusted to median cohort age (45 years). Note the volume plots use logarithmical scale and different scales across different ROIs.*

Statistics for the combined dataset

We start with the description of the age-related changes in the qMRI metrics. Figure 5 provides a representation of effect size for the effect of interest in all 14 ROI. The generalized linear models showed significant ($p<0.05$, FDR-corrected) combined age and age^2 dependence in R_1, R2* and ROI volume for all of the segmented nuclei, except for the Periaqueductal Gray volume (see significance indicated via * in Figure 5).

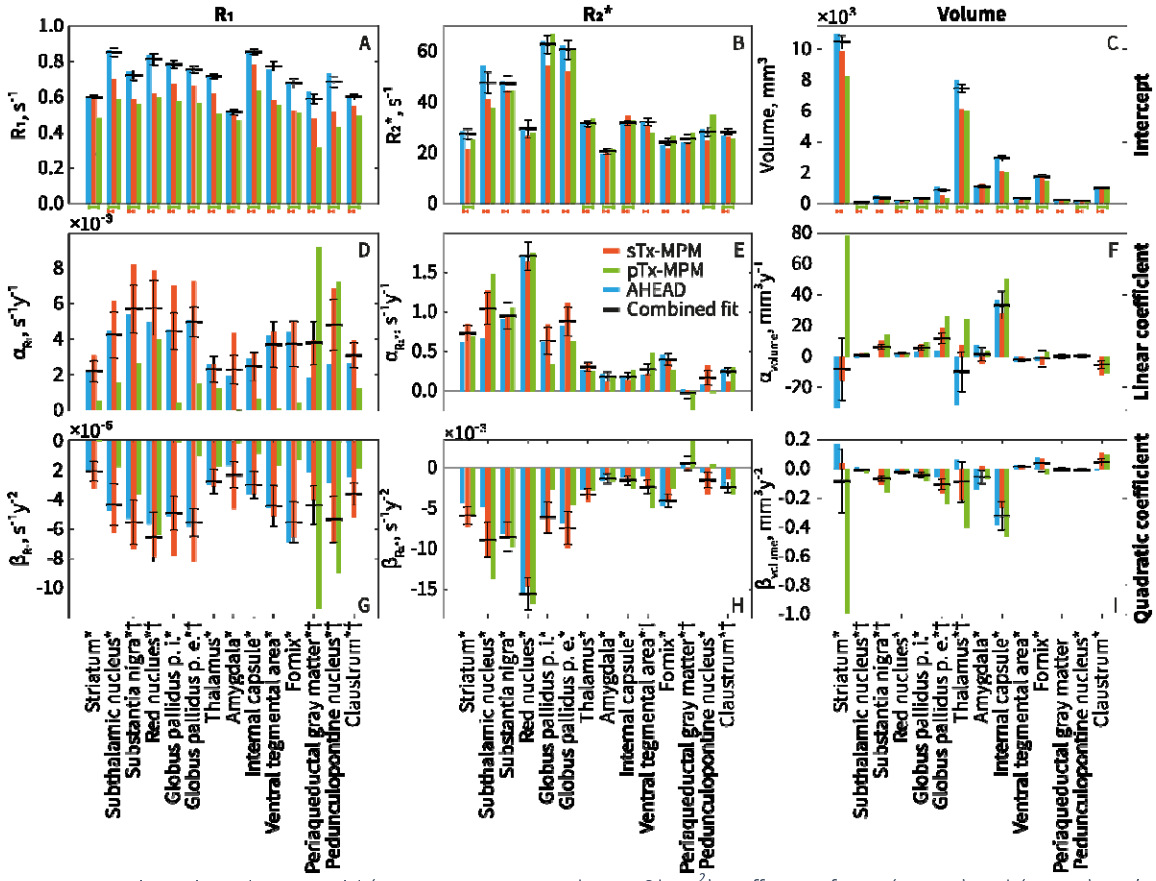

*Figure 5 Age-dependence linear model (metric ~ intercept + α*age + β*age^2) coefficients for R1 (A, D, G), R2* (B, E, H), and ROI volume (C, F, I). Colored columns show effect size for per-dataset fits (blue for AHEAD data, red for sTx-MPM, green for pTx-MPM), black bar indicates effect size for a combined dataset fit. Top row (A, B, C) shows the model intercept, middle row (D, E, F) shows the model linear coefficient, bottom row (G, H, I) shows the model quadratic coefficient. Asterisk (*) following the ROI name indicates statistical significance of the age dependence in the combined dataset. Red bar under intercept graphs (A, B, C) indicates statistically significant bias of the sTx-MPM data against AHEAD data, green bar under intercept graphs indicates statistically significant bias of the pTx-MPM data against AHEAD data. Dagger (†) indicates statistically significant interaction between protocol-related categorical variable and age-related variables.*

Next, we focus on the differences between protocols (see color-coded significance across Figure 5). When comparing the locally acquired data with the AHEAD data, the models showed significant differences of R1 values measured with the pTx-MPM protocol in all the ROIs, and with the sTx-MPM protocol in all ROIs except Striatum (see Figure 5 A for the effect sizes). Significant differences against the AHEAD protocol were detected in R2* measured by the pTx-MPM protocol in Striatum, Thalamus, Amygdala, Internal Capsule Pedunculopontine Nucleus, and Claustrum. Measured R2* differences for the sTx-MPM protocol were found in all ROIs except Pedunculopontine Nucleus (see Figure 5 B for the effect sizes). Furthermore, differences in ROI volume measured with the pTx-MPM protocol, compared to the ones calculated in the AHEAD data were found in all ROIs except Striatum, Substantia Nigra, Amygdala, and Periaqueductal Gray, and with the sTx-MPM protocol – in all ROIs (see Figure 5 C for the effect sizes).

Finally, we consider the influence of the protocol on the age dependence (see significance indicated via † in Figure 5). Significant interactions between age-related variables and protocol were detected for R1 in Substantia Nigra, Red Nucleus, Globus Pallidus (pars externa) Periaqueductal Gray, Pedunculopontine Nucleus, and Claustrum; interactions for R2* measurements were detected in Ventral Tegmental Area, Periaqueductal Gray, and

Claustrum; and interactions in volume measurements were detected in Subthalamic Nucleus, Substantia Nigra, Globus Pallidus pars externa, and Thalamus.

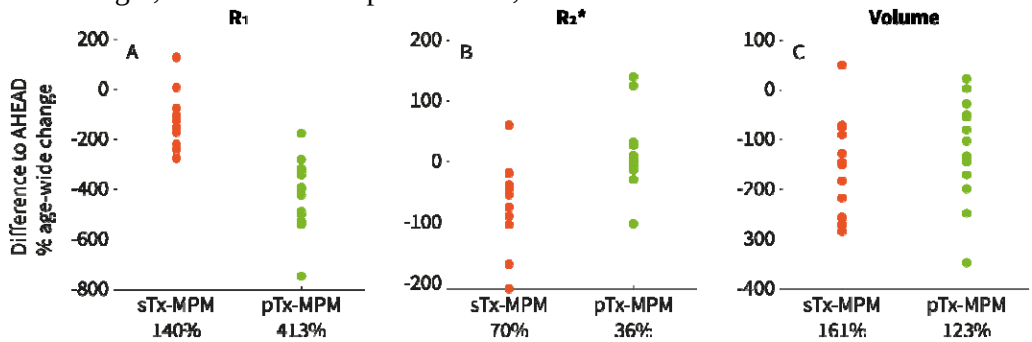

Figure 6 The effect of protocol selection on the measured protocol differences in quantitative value normalized to the age-related change in the metric, shown for 14 subcortical structures (dots). A – pTx-MPM (green) exhibits lower R_1 values with mean absolute difference of 413%, while sTx-MPM (red) demonstrates a difference of 140% of the expected age-related change (significant with $p<10^{-5}$). B – both sTx-MPM and pTx-MPM exhibit average difference in R_2^* lower than the expected age-related change, 70% and 36% correspondingly (difference significant with $p<0.001$). C – both sTx-MPM and pTx-MPM show comparable differences against AHEAD data (161% and 123% of the expected age-related change, significant difference only with $p<0.05$).

The expected age-related change in quantitative metrics varied across different ROI between 3.8% (0.025 s^{-1}) and 16.8% (0.119 s^{-1}) of the mean value for R_1 (0.65 s^{-1} and 0.71 s^{-1} respectively), between 6.4% (2.06 s^{-1}) and 29.7% (20.96 s^{-1}) for the R_2^* (mean values of 32.29 s^{-1} and 70.55 s^{-1} respectively) and between 4.9% (10.7 mm^3) and 26.5% (268.1 mm^3) for the region volume (mean values of 220.8 mm^3 and 1010.2 mm^3 respectively). At the same time (Figure 6) the mean absolute protocol-related difference from the AHEAD data for the sTx-MPM was 140% of the expected age-wide variation for R1, 70% for R2*, and 161% for the volume. For the pTx-MPM the values were 413% of the expected age-wide variation for R1, 36% for R2*, and 123% for the volume. When referenced to the expected mean value of the quantitative metric (instead of the expected age-wide variation), the sTx-MPM protocols exhibited on average a 9.5% lower R_1 values, a 6.6% lower R_2^*, and 17% lower volume. For the pTx-MPM the R_1 values were on average 27% lower than in the AHEAD data, the R_2^* values were 0.5% higher, and regional volume was 12.7% lower. The ages of maximal R1 calculated from the age trajectories are presented in the Supplementary Figure S7.

B_1^+ field measurements

The analysis of potential bias resulting from different B_1^+ mapping procedure and different B_1^+ settings in sTx and pTx protocols are shown in Figure 7. Comparison of different measurement B_1^+ mapping techniques (Figure 7, C), indicated that in three subjects where both AFI and SE-STE-EPI based field measurements were performed, the resulting field maps matched closely. Particularly, linear models with intercept fixed at zero (i.e., assuming both methods showing zero field when no field is present), demonstrated a slope of 1.002 with adjusted r-square of 0.9657.

Exploration of the measured B_1^+ field in sTx-MPM and pTx-MPM protocols demonstrated (Figure 7, A) an average 14% higher B_1^+ field in subcortical nuclei and a more consistent field settings across subjects in the pTx-MPM protocol (across subject CoV for mean field values in sTx-MPM if 0.062 versus 0.049 in pTx). In addition (Figure 7, B), pTx-MPM protocol showed more homogeneous fields across subcortical structures, demonstrating lower CoV across subcortical structures of 0.093 versus 0.118 in sTx-MPM (in group average terms).

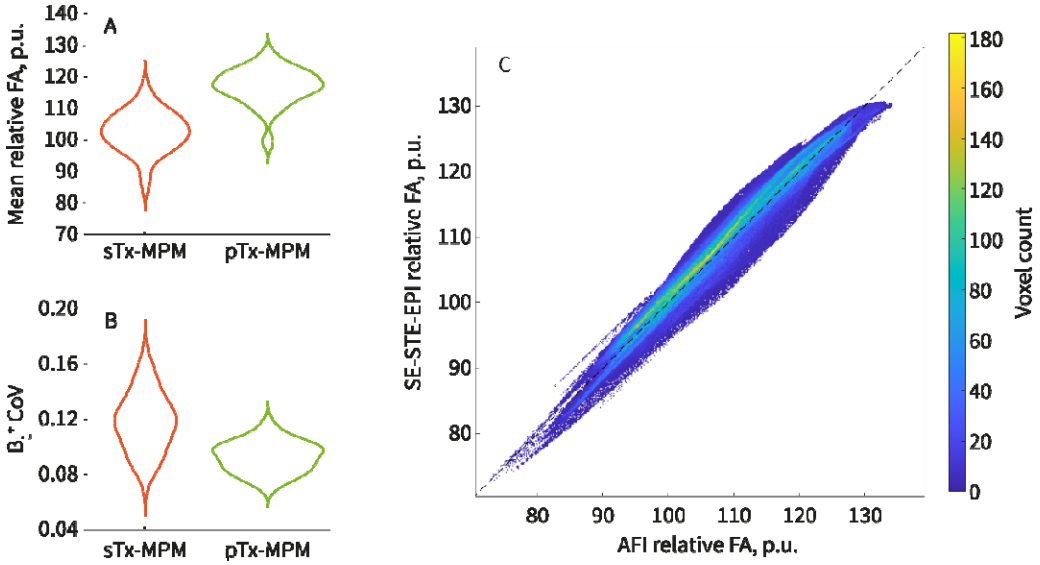

Figure 7 A – Measured B_1^+ values (in percent units relative to value corresponding to the nominal flip angle) averaged across all the segmented subcortical structures. B - Coefficient of variation (CoV) of the measured B_1^+ values across all subcortical structures. The graph A demonstrates the pTx-MPM protocol to produce higher B_1^+ values, which also vary less from subject to subject. Graph B demonstrates the pTx-MPM protocol to produce more uniform B_1^+ distribution with less homogeneity variation across subjects. C – joint histogram of B_1^+ measured consecutively in the same subjects with SE-STE-EPI and AFI methods. The histogram shows almost perfect match of the field measured with two methods across subcortical structures.

Discussion

UHF qMRI promises to provide reproducible values of tissue microstructural parameters measured in small brain structures across sites and protocols. Here we explore how these promises hold up to pooling datasets acquired at different sites and employing different acquisition techniques. We used age-related changes in qMRI parameters to identify protocol dependence of two qMRI metrics (R1 and R2*), and volume of 14 subcortical structures segmented based on multiparametric qMRI data with the same automatic approach. While the subcortical structures benefit from the increased image quality in UFH MRI their imaging might be affected by the lower reception sensitivity compared to the cortical parts of the brain. We found that the age-related changes for R1, R2* and volume detected by different protocols were of the same order of magnitude across all ROI, thus supporting the idea that age-related inter-individual variability can be correctly grasped across sites and protocols using qMRI approaches. We further observe larger relative difference across protocols for R1 and volume, while R2* remains more consistent for most regions, both in terms of difference against the reference AHEAD data (Figure 6), and in terms of displaying less interaction terms and lower amount of protocol-significant offsets.

Overall, the analysis of age-related changes demonstrates all the qMRI metrics to exhibit the inverted U-shape dependence, with the linear term in the myelin-related qMRI metrics (i.e., R1) and the iron-related metrics (i.e., R2*) showing increase with age. This is in line with the reference dataset analysis (3) and probably arises from known changes in brain tissue structure (e.g., myelin loss, and iron accumulation)(39). The low number of identified

interaction terms between age-related variables and protocol category suggests that the age dependence matches across protocols. This is particularly true for R2* and volume measurements, having the least number of significant interaction-containing models, and less so for the R1 measurements comprising the majority of detected interactions. Nevertheless, the age distribution used to derive age dependence both in local and AHEAD samples is uneven with a few participants in the middle age range (30 to 50 years) suggesting the age effects might have been modelled suboptimally.

Examining the variation of qMRI metrics between MP2RAGE and variable flip angle approaches have indicated (Figure 6, A) higher R1 in the reference AHEAD data compared to the sTx-MPM and pTx-MPM data. The difference is extremely striking when compared to the expected adult lifespan variation of the R1, showing that the measured R1 is more affected by the protocol choice rather than participant differences. It could be speculated that due to R1 measurements in MPM protocol being strongly depend on the knowledge of the flip angles, which is usually addressed by incorporating the field mapping (36), it is the field mappings (40) that could have caused the biases. Nevertheless, the field measurements (Figure 7, C) show that the different B1 mapping methods used in sTx-MPM and pTx-MPM produce consistent estimates of the $B1^+$ field. The measured $B1^+$ field (Figure 7, A) is nevertheless different in sTx-MPM and pTx-MPM configuration, suggesting it could serve as the source of the error in the R1 estimation, but not via the biases in the field mapping procedure, but rather due to incomplete correction of the B_1^+ effect during R1 calculation. This suggests further investigations into the assumptions and approximations used in the qMRI parameter extraction tools are required.

Alternatively, the discrepancy between AHEAD (MPRAGE-like acquisition) and MPMs may be induced by inherent simplification used in R1 calculation, assuming one universal R1 parameter to be able to characterize complex brain tissue, comprising multiple cellular and subcellular compartments, all exhibiting potentially different microstructural properties and macromolecular composition. More realistic, four pool models may lift these simplifications, but usually results in explosion of parameter space, therefore practically difficult to use.

The R2* data exhibits smaller biases (Figure 6, B), which are smaller than the estimated R2* change over the adult lifespan. This is potentially due to the less pronounced dependence of R2* on the B1+ field , particularly in areas with low myelin content, that results in the least amount of biases and interactions in the R2* part of the pooled dataset. The differences in measured R2* between protocols may partly result from the model being simplified by assuming single R2* values for complex biological tissue with different compartments (41). This results in measured R2* depending on sequence parameters (e.g. flip angle in MPM acquisition). While these effects are expected to be most prominent in white matter, they may also affect subcortical nuclei, which are partially myelinated.

Finally, volume estimation is systematically higher in the AHEAD data (Figure 5 C), which may be related to the absence of χ maps in the locally acquired MPM data so that segmentation procedure could not be matched across all the cohorts. Notably, the difference in estimated volume between sTx-MPM and pTx-MPM protocols (Figure 6, C) is hardly significant, supporting the hypothesis of the lack of QSM data input into the segmentation algorithm being the reason for differences in the measured volume, and emphasizing the need in maximum harmonization of processing pipelines. Note, that the systematic differences in segmentation procedure may have also contributed to the observed differences in qMRI values, as slightly different regions were compared across protocols.

To our knowledge this work demonstrated the first assessment of the pooling effects in the ultra-high resolution UHF qMRI data performed over a relatively large participant cohort, comprising data from different sites collected with different quantitative protocols. This also included initial comparison of the impact introduction of the parallel transmission technology

can have on the measured quantitative parameters. Moreover, this study demonstrated the diversity of age-related and scanner-related effects across different structures by concentrating on numerous subcortical structures, while previous research was mostly focused on larger tissue classes, such as white and gray matter (15,16,18,19).

Examining the UHF qMRI measurements across a number of protocols and measurement sites helped identify the issues arising when attempting to merge heterogeneous datasets: the lack of consistency in absolute values of qMRI metrics, as well as the discrepancies in segmentation relying on different input data. Overall, our results suggest that while inferring age-related changes (i.e., forward inference) in e.g. iron accumulation (based on R2* values) can be reliable using pooled heterogenous qMRI datasets, inferences on the brain status, e.g., for calculating the brain age from the qMRI values (42), or for calculating the pathology cutoff values (i.e., reverse inference (43)), may be obscured by the protocol-related biases and must be made with caution. Thus, pooling across sites and protocol requires tailored harmonization procedures (44,45), using traveling heads studies or dedicated phantoms, potentially in the framework of MRI metrology (46). R2* and ROI volume are seemingly more suited for such reverse inference or prediction in large and heterogenous multicentric studies due to their inter-protocol stability. This encourages the application of UHF qMRI to the disorders correlated with the change of the R2* in the brain tissues, such as Alzheimer's disease, Parkinson's disease as well as multiple sclerosis (47).

Acknowledgments

This work was conducted at the GIGA-In Vivo Imaging platform of ULiège, Belgium. We thank Nikita Beliy, Sophie Laloux, Annick Claes, Christian Degueldre, Brigitte Herbillon, Gregory Hammad, Benjamin Lauricella, Alexia Lesoinne, Pierre Maquet, John Read and Aubry Robert for their help in the different stages of the study. This work was supported by Fonds National de la Recherche Scientifique (FRS-FNRS, T.0242.19, T.0238.23 and J. 0222.20), the EU Joint Programme Neurodegenerative Disease Research (JPND) IRONSLEEP/HISTOPARK and SCAIFIELD projects, respectively – (FNRS references: PINT-MULTI R.8011.21/R.8002.25 & 8006.20); Action de Recherche Concertée – Fédération Wallonie-Bruxelles (ARC SLEEPDEM 17/27-09), Fondation Recherche Alzheimer (SAO-FRA 2019/0025 & 2022/0014), Fondation Léon Frédéricq, ULiège, European Regional Development Fund (Radiomed, Biomed-Hub, WALBIOIMAGING).

EB was supported by the Maastricht University - ULiège Imaging Valley. NM, CP, MZ, EL, SD, CG, FB and GV are/were supported by the FRS-FNRS. PT and LL are/were supported by the EU Joint Programme Neurodegenerative Disease Research (JPND) IRONSLEEP/HISTOPARK and SCAIFIELD projects, respectively – FNRS references: PINT-MULTI R.8011.21/R.8002.25 & 8006.20. LL is supported by the European Regional Development Fund (WALBIOIMAGING).

References

1. Webb AG, Collins CM. Parallel transmit and receive technology in high-field magnetic resonance neuroimaging. Int J Imaging Syst Technol. 2010;20(1):2–13.

2. Trampel R, Bazin PL, Pine K, Weiskopf N. In-vivo magnetic resonance imaging (MRI) of laminae in the human cortex. NeuroImage. 2019 Aug 15;197:707–15.

3. Miletić S, Bazin PL, Isherwood SJS, Keuken MC, Alkemade A, Forstmann BU. Charting human subcortical maturation across the adult lifespan with in vivo 7 T MRI. NeuroImage. 2022 Apr 1;249:118872.

4. Forstmann BU, de Hollander G, van Maanen L, Alkemade A, Keuken MC. Towards a mechanistic understanding of the human subcortex. Nat Rev Neurosci. 2017 Jan;18(1):57–65.

5. Keuken MC, Isaacs BR, Trampel R, van der Zwaag W, Forstmann BU. Visualizing the Human Subcortex Using Ultra-high Field Magnetic Resonance Imaging. Brain Topogr. 2018 July 1;31(4):513–45.

6. Weiskopf N, Edwards LJ, Helms G, Mohammadi S, Kirilina E. Quantitative magnetic resonance imaging of brain anatomy and in vivo histology. Nat Rev Phys. 2021 Aug;3(8):570–88.

7. Stüber C, Morawski M, Schäfer A, Labadie C, Wähnert M, Leuze C, et al. Myelin and iron concentration in the human brain: A quantitative study of MRI contrast. NeuroImage. 2014 June 1;93:95–106.

8. Metere R, Möller HE. Linear Models for Estimating Myelin and Iron Content in the Brain. In: Proceedings of the ISMRM 24th Annual Meeting & Exhibition. Singapore; 2016.

9. Mackay A, Whittall K, Adler J, Li D, Paty D, Graeb D. *In vivo* visualization of myelin water in brain by magnetic resonance. Magn Reson Med. 1994 June;31(6):673–7.

10. Sereno MI, Lutti A, Weiskopf N, Dick F. Mapping the Human Cortical Surface by Combining Quantitative T1 with Retinotopy†. Cereb Cortex. 2013 Sept;23(9):2261–8.

11. Caan MWA, Bazin PL, Marques JP, de Hollander G, Dumoulin SO, van der Zwaag W. MP2RAGEME: T1, T2*, and QSM mapping in one sequence at 7 tesla. Hum Brain Mapp. 2019;40(6):1786–98.

12. Weiskopf N, Suckling J, Williams G, Correia M, Inkster B, Tait R, et al. Quantitative multi-parameter mapping of R1, PD*, MT, and R2* at 3T: a multi-center validation. Front Neurosci [Internet]. 2013 [cited 2023 Nov 30];7. Available from: https://www.frontiersin.org/articles/10.3389/fnins.2013.00095

13. Lee Y, Callaghan MF, Acosta-Cabronero J, Lutti A, Nagy Z. Establishing intra- and inter-vendor reproducibility of T_1 relaxation time measurements with 3T MRI. Magn Reson Med. 2019 Jan;81(1):454–65.

14. Bauer CM, Jara H, Killiany R. Whole brain quantitative T2 MRI across multiple scanners with dual echo FSE: Applications to AD, MCI, and normal aging. NeuroImage. 2010 Aug 15;52(2):508–14.

15. Carradus AJ, Mougin O, Hunt BAE, Tewarie PK, Geades N, Morris PG, et al. Age-related differences in myeloarchitecture measured at 7 T. Neurobiol Aging. 2020 Dec 1;96:246–54.

16. Piredda GF, Caneschi S, Hilbert T, Bonanno G, Joseph A, Egger K, et al. Submillimeter T1 atlas for subject-specific abnormality detection at 7T. Magn Reson Med. 2023;89(4):1601–16.

17. Keuken MC, Bazin PL, Backhouse K, Beekhuizen S, Himmer L, Kandola A, et al. Effects of aging on T1, T2*, and QSM MRI values in the subcortex. Brain Struct Funct. 2017 Aug 1;222(6):2487–505.

18. Heij J, van der Zwaag W, Knapen T, Caan MWA, Forstman B, Veltman DJ, et al. Quantitative MRI at 7-Tesla reveals novel frontocortical myeloarchitecture anomalies in major depressive disorder. Transl Psychiatry. 2024 June 20;14(1):1–9.

19. Al-Radaideh A, Mougin OE, Lim SY, Chou IJ, Constantinescu CS, Gowland P. Histogram analysis of quantitative T1 and MT maps from ultrahigh field MRI in clinically isolated syndrome and relapsing–remitting multiple sclerosis. NMR Biomed. 2015;28(11):1374–82.

20. Rua C, Raman B, Rodgers CT, Newcombe VFJ, Manktelow A, Chatfield DA, et al. Quantitative susceptibility mapping at 7 T in COVID-19: brainstem effects and outcome associations. Brain. 2024 Dec 1;147(12):4121–30.

21. Callaghan MF, Freund P, Draganski B, Anderson E, Cappelletti M, Chowdhury R, et al. Widespread age-related differences in the human brain microstructure revealed by quantitative magnetic resonance imaging. Neurobiol Aging. 2014 Aug 1;35(8):1862–72.

22. Pine K, Zubkov M, Bazin PL, Perlaki G, Edwards L, Alkemade A, et al. Quantitative multi-parametric mapping of human subcortex at ultrahigh field. In: Proceedings of the Annual Meeting of ISMRM. Singapore; 2024. p. 3714.

23. Rua C, Clarke WT, Driver ID, Mougin O, Morgan AT, Clare S, et al. Multi-centre, multi-vendor reproducibility of 7T QSM and R2* in the human brain: Results from the UK7T study. NeuroImage. 2020 Dec 1;223:117358.

24. Leutritz T, Seif M, Helms G, Samson RS, Curt A, Freund P, et al. Multiparameter mapping of relaxation (R1, R2*), proton density and magnetization transfer saturation at 3 T: A multicenter dual-vendor reproducibility and repeatability study. Hum Brain Mapp. 2020;41(15):4232–47.

25. Alkemade A, Mulder MJ, Groot JM, Isaacs BR, Van Berendonk N, Lute N, et al. The Amsterdam Ultra-high field adult lifespan database (AHEAD): A freely available multimodal 7 Tesla submillimeter magnetic resonance imaging database. NeuroImage. 2020 Nov;221:117200.

26. Pruessmann KP, Weiger M, Scheidegger MB, Boesiger P. SENSE: Sensitivity encoding for fast MRI. Magn Reson Med. 1999 Nov;42(5):952–62.

27. Griswold MA, Jakob PM, Heidemann RM, Nittka M, Jellus V, Wang J, et al. Generalized autocalibrating partially parallel acquisitions (GRAPPA). Magn Reson Med. 2002 June;47(6):1202–10.

28. Jiru F, Klose U. Fast 3D radiofrequency field mapping using echo-planar imaging. Magn Reson Med. 2006;56(6):1375–9.

29. Breuer FA, Blaimer M, Mueller MF, Seiberlich N, Heidemann RM, Griswold MA, et al. Controlled aliasing in volumetric parallel imaging (2D CAIPIRINHA). Magn Reson Med. 2006 Mar;55(3):549–56.

30. Pine K, Groß-Weege N, Edwards L, Leutritz T, Freund P, Weiskopf N. Parallel transmit (pTx) kT-points pulses improve 500 µm resolution quantitative multi- parameter mapping (MPM) at 7T. In: Proceedings of the Annual Meeting of ISMRM. Toronto, Canada; 2023. p. 2163.

31. Yarnykh VL. Actual flip-angle imaging in the pulsed steady state: A method for rapid three-dimensional mapping of the transmitted radiofrequency field. Magn Reson Med. 2007;57(1):192–200.

32. Bazin PL, Alkemade A, van der Zwaag W, Caan M, Mulder M, Forstmann BU. Denoising High-Field Multi-Dimensional MRI With Local Complex PCA. Front Neurosci [Internet]. 2019 Oct 9 [cited

2024 Nov 6];13. Available from: https://www.frontiersin.org/journals/neuroscience/articles/10.3389/fnins.2019.01066/full

33. Marques JP, Kober T, Krueger G, van der Zwaag W, Van de Moortele PF, Gruetter R. MP2RAGE, a self bias-field corrected sequence for improved segmentation and T_1-mapping at high field. NeuroImage. 2010 Jan 15;49(2):1271–81.

34. Langkammer C, Bredies K, Poser BA, Barth M, Reishofer G, Fan AP, et al. Fast quantitative susceptibility mapping using 3D EPI and total generalized variation. NeuroImage. 2015 May 1;111:622–30.

35. Bazin PL, Alkemade A, Mulder MJ, Henry AG, Forstmann BU. Multi-contrast anatomical subcortical structures parcellation. Verstynen T, Frank MJ, Verstynen T, Neumann WJ, editors. eLife. 2020 Dec 16;9:e59430.

36. Tabelow K, Balteau E, Ashburner J, Callaghan MF, Draganski B, Helms G, et al. hMRI – A toolbox for quantitative MRI in neuroscience and clinical research. NeuroImage. 2019 July 1;194:191–210.

37. Shannon CE. A mathematical theory of communication. Bell Syst Tech J. 1948 July;27(3):379–423.

38. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B Methodol. 1995 Jan 1;57(1):289–300.

39. Bartzokis G. Alzheimer's disease as homeostatic responses to age-related myelin breakdown. Neurobiol Aging. 2011 Aug 1;32(8):1341–71.

40. Lutti A, Stadler J, Josephs O, Windischberger C, Speck O, Bernarding J, et al. Robust and Fast Whole Brain Mapping of the RF Transmit Field B1 at 7T. Zhan W, editor. PLoS ONE. 2012 Mar 12;7(3):e32379.

41. Milotta G, Corbin N, Lambert C, Lutti A, Mohammadi S, Callaghan MF. Mitigating the impact of flip angle and orientation dependence in single compartment R2 * estimates via 2-pool modeling. Magn Reson Med. 2023 Jan;89(1):128–43.

42. Bacas E, Kahhalé I, Raamana PR, Pablo JB, Anand AS, Hanson JL. Probing multiple algorithms to calculate brain age: Examining reliability, relations with demographics, and predictive power. Hum Brain Mapp. 2023;44(9):3481–92.

43. Bzdok D, Ioannidis JPA. Exploration, Inference, and Prediction in Neuroscience and Biomedicine. Trends Neurosci. 2019 Apr 1;42(4):251–62.

44. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. 2007 Jan 1;8(1):118–27.

45. Pomponio R, Erus G, Habes M, Doshi J, Srinivasan D, Mamourian E, et al. Harmonization of large MRI datasets for the analysis of brain imaging patterns throughout the lifespan. NeuroImage. 2020 Mar 1;208:116450.

46. Hall MG, Cashmore M, Cho HM, Ittermann B, Keenan KE, Kolbitsch C, et al. Metrology for MRI: the field you've never heard of. Magn Reson Mater Phys Biol Med. 2025 July 1;38(3):387–412.

47. Zecca L, Youdim MBH, Riederer P, Connor JR, Crichton RR. Iron, brain ageing and neurodegenerative disorders. Nat Rev Neurosci. 2004 Nov;5(11):863–73.