Multi-contrast Anatomical Subcortical Structures Parcellation

Pierre-Louis Bazin[1,2*], **Anneke Alkemade**[1], **Martijn J. Mulder**[1,3], **Amanda G. Henry**[4], **Birte U. Forstmann**[1]

***For correspondence:** pilou.bazin@uva.nl (PLB)

[1]Integrative Model-based Cognitive Neuroscience research unit, University of Amsterdam, Amsterdam, The Netherlands; [2]Max-Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany; [3]Psychology Department, Utrecht University, Utrecht, The Netherlands; [4]Faculty of Archaeology, Leiden University, Leiden, The Netherlands.

Abstract The human subcortex is comprised of more than 450 individual nuclei which lie deep in the brain. Due to their small size and close proximity, up until now only 7% have been depicted in standard MRI atlases. Thus, the human subcortex can largely be considered as terra incognita. Here we present a new open source parcellation algorithm to automatically map the subcortex. The new algorithm has been tested on 17 prominent subcortical structures based on a large quantitative MRI dataset at 7 Tesla. It has been carefully validated against expert human raters and previous methods, and can easily be extended to other subcortical structures and applied to any quantitative MRI dataset. In sum, we hope this novel parcellation algorithm will facilitate functional and structural neuroimaging research into small subcortical nuclei and help to chart terra incognita.

Introduction

Subcortical brain structures are often neglected in neuroimaging studies due to their small size, limited inter-regional contrast, and weak signal-to-noise ratio in functional imaging (***Forstmann et al., 2017a***; ***Johansen-Berg, 2013***). Yet, these small and diverse structures are prominent nodes in functional networks (***Marquand et al., 2017***; ***Ji et al., 2019***), and they undergo pathological alterations already at early stages of neurodegenerative diseases (***Andersen et al., 2014***; ***Koshiyama et al., 2018***). Deep brain stimulation surgery, originally performed to reduce motor symptoms in essential tremors, is now a promising therapeutic option in later stages of Parkinson's disease and movement disorders, as well as refractory psychiatric illnesses in obsessive-compulsive disorder, anorexia, or depression (***Forstmann et al., 2017b***; ***Mosley et al., 2018***). Evolutionary genetics even uncovered that in modern humans, Neanderthal-inherited alleles were preferentially down-regulated in subcortical and cerebellar regions compared to other brain regions (***McCoy et al., 2017***), suggesting these structures to be essential in making us specifically human.

Despite their importance, these areas are particularly difficult to image. Furthermore, the size, shape and location of these brain regions changes with development and aging (***Fjell et al., 2013***; ***Keuken et al., 2013***; ***Yeatman et al., 2014***; ***Herting et al., 2018***). Experience-based plasticity continuously remodels myelin (***Tardif et al., 2016***; ***Hill et al., 2018***; ***Turner, 2019***), iron and other magnetic substances accumulate with age or pathology (***Andersen et al., 2014***; ***Zhang et al., 2018***), both bringing changes in the magnetic resonance imaging (MRI) appearance of subcortical regions with diverse tissue characteristics (***Draganski et al., 2011***; ***Keuken et al., 2017***).

Thus, mapping the structure and function of the subcortex is a major endeavor as well as a

major challenge for human neuroscience. Extensive work available from animal brain models unfortunately does not translate in a straightforward way to human subcortical anatomy nor does it shed much light on its involvement in human cognition (***Steiner and Tseng, 2017***). Besides serious difficulties in obtaining adequate measures of subcortical neural activity in functional MRI (***de Hollander et al., 2017***; ***Miletić et al., 2020***), atlases and techniques for labeling accurately and reliably individual subcortical structures have also been scarce (***Frazier et al., 2005***; ***Chakravarty et al., 2006***; ***Ahsan et al., 2007***; ***Yelnik et al., 2007***; ***Qiu et al., 2010***; ***Patenaude et al., 2011***), typically labelling the thalamus, striatum (or its subdivision into caudate and putamen), and globus pallidus (internal and external segments combined), sometimes the amygdala. However, recent advances in anatomical MRI, combining multiple contrasts and/or quantitative MRI mapping and utilizing the higher resolution achievable with 7 Tesla (7T) and above have started to reduce the gap, each mapping a few additional structures or sub-structures, primarily the iron-rich substantia nigra, red nucleus and sub-thalamic nucleus (***Keuken et al., 2013***; ***Xiao et al., 2015***; ***Visser et al., 2016a***,b; ***Wang et al., 2016***; ***Makowski et al., 2018***; ***Ewert et al., 2018***; ***Iglesias et al., 2018***; ***Pauli et al., 2018***; ***Sitek et al., 2019***). While these efforts generated valuable atlases, they do not yet enable to identify many subcortical structures in individual subjects. Manual delineation, on the other hand, requires extensive manual labor from highly trained experts which cannot be easily applied to large cohorts or clinical settings.

Here, we propose a new automated parcellation technique to identify and label 17 individual subcortical structures of varying size and composition in individual subjects, based on a large quantitative 7T MRI database (***Alkemade et al., 2020***), using quantitative maps of relaxation rates R1 and R2* (1/T1 and 1/T2* respectively) and quantitative susceptibility maps (QSM) as anatomical contrasts. The algorithm, named Multi-contrast Anatomical Subcortical Structure Parcellation (MASSP), follows a Bayesian multi-object approach similar in essence to previous efforts (***Fischl et al., 2002***; ***Iglesias et al., 2013***; ***Visser et al., 2016a***; ***Garzón et al., 2018***), combining shape priors, intensity distribution models, spatial relationships and global constraints. The main innovation of our approach is to explicitly estimate interfaces between subcortical structures based on a joint model derived from signed distance functions. Modeling interfaces in addition to the structure itself provides a rich basis to encode relationships and anatomical knowledge in shape and intensity priors. A voxel-wise Markovian diffusion regularizes the combined priors for each defined interface, lowering the imaging noise. Finally, the voxel-wise posteriors for the different structures and interfaces are further combined into global anatomical parcels by topology correction and region growing taking into account volumetric priors, which regularizes parcellation results further in smaller nuclei with low or heterogeneous contrast. To validate the results from this new method, in a thorough comparison with expert manual labeling, we show that the proposed method provides results very close from manual raters in many structures and exhibit reasonable bias across the adult lifespan. The method can easily be extended to new structures, can be applied to any quantitative MRI dataset and is available in Open Source as part of Nighres (***Huntenburg et al., 2018***), a neuroimage analysis package aimed at high resolution neuroimaging.

Results

The MASSP parcellation method presented here has been trained to parcellate the following 17 structures: striatum (Str), thalamus (Tha), lateral, 3rd and 4th ventricles (LV, 3V, 4V), amygdala (Amg), globus pallidus internal segment (GPi) and external segment (GPe), SN, STN, red nucleus (RN), ventral tegmental area (VTA), fornix (fx), internal capsule (ic), periaqueductal gray (PAG), pedunculopontine nucleus (PPN), and claustrum (Cl), see Fig. 1. These structures include the most commonly defined subcortical regions (Str, Tha, Amg, LV), the main iron-rich nuclei (GPi, GPe, RN, SN, STN), as well as smaller, less studied areas (VTA, PAG, PPN, Cl), white matter structures (ic, fx) and the central ventricles (3V, 4V).

MASSP uses a data set of ten expert delineations as a basis for its modeling. From the delineations, an atlas of interfaces between structures, shape skeletons, and interface intensity his-

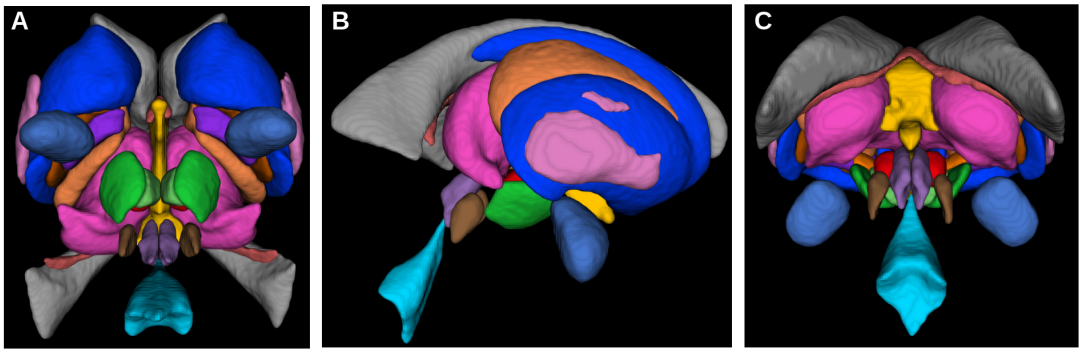

Figure 1. The 17 subcortical structures currently included in the parcellation algorithm in axial (A), sagittal (B) and coronal (C) views.

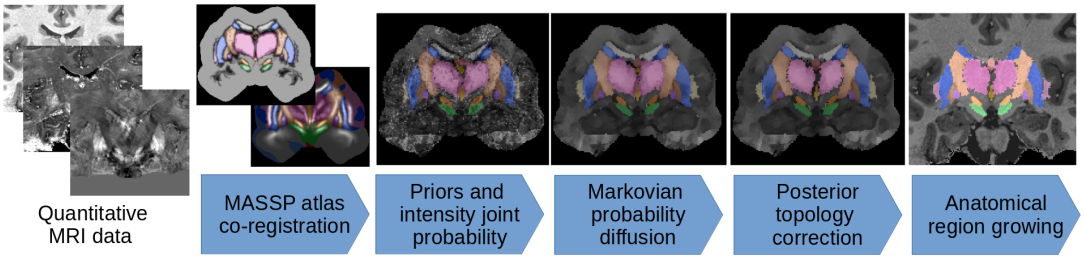

Figure 2. The MASSP parcellation pipeline. Atlas priors for interfaces between structures are combined to the MRI data, regularized via probability diffusion and topology correction, and the final structure posteriors are jointly estimated by region growing.

tograms are generated, and used as prior in a multiple-step non-iterative Bayesian algorithm, see Fig. 2 and Methods.

Validation against manual delineations

In a leave-one-out validation study comparing performance with the manual delineations, MASSP performed above 95% of the level of quality of the raters for Str, Tha, 4V, GPe, SN, RN, VTA, ic in terms of Dice overlap, the most stringent of the quality measures (see Fig. 3 and 4 and Table 1. Several of the smaller structures have lower overlap ratios likely due to their smaller size (GPi, STN, PAG, PPN). Structures with an elongated shape (fx, Cl) remain challenging, due to the fact that small differences in location can substantially reduce overlap (***Bazin et al., 2016***). Despite these challenges, when comparing the dilated Dice scores, all structures were above 75% of overlap, with most reaching over 90% of the manual raters ability. Note that the Dice coefficient is very sensitive to size, as smaller structures will have lower overlap ratios for the same number of mis-classified voxels. The dilated Dice coefficient is more representative of the variability regardless of size, as the smaller structures can reach high levels of overlap, both in manual and automated parcellations (see Table 1). The average surface distance confirms these results, showing values generally between one and two voxels of distance at a resolution of 0.7 mm, except in the cases of Amg, LV, fx, PPN, and Cl. These structures are generally more variable (LV), elongated (fx, Cl), or have a particularly low contrast with neighboring regions (Amg, PPN).

Comparison to other automated methods

To provide a basis for comparison, we applied other freely available methods for subcortical structure parcellation to the same ten subjects. MASSP performs similarly to or better than Freesurfer, FSL FIRST and a multi-atlas registration using ANTs. Multi-atlas registration provides high accuracy in most structures as well, but is biased toward under-estimating the size of smaller and elongated structures where overlap is systematically reduced across the individual atlas subjects. Multi-atlas

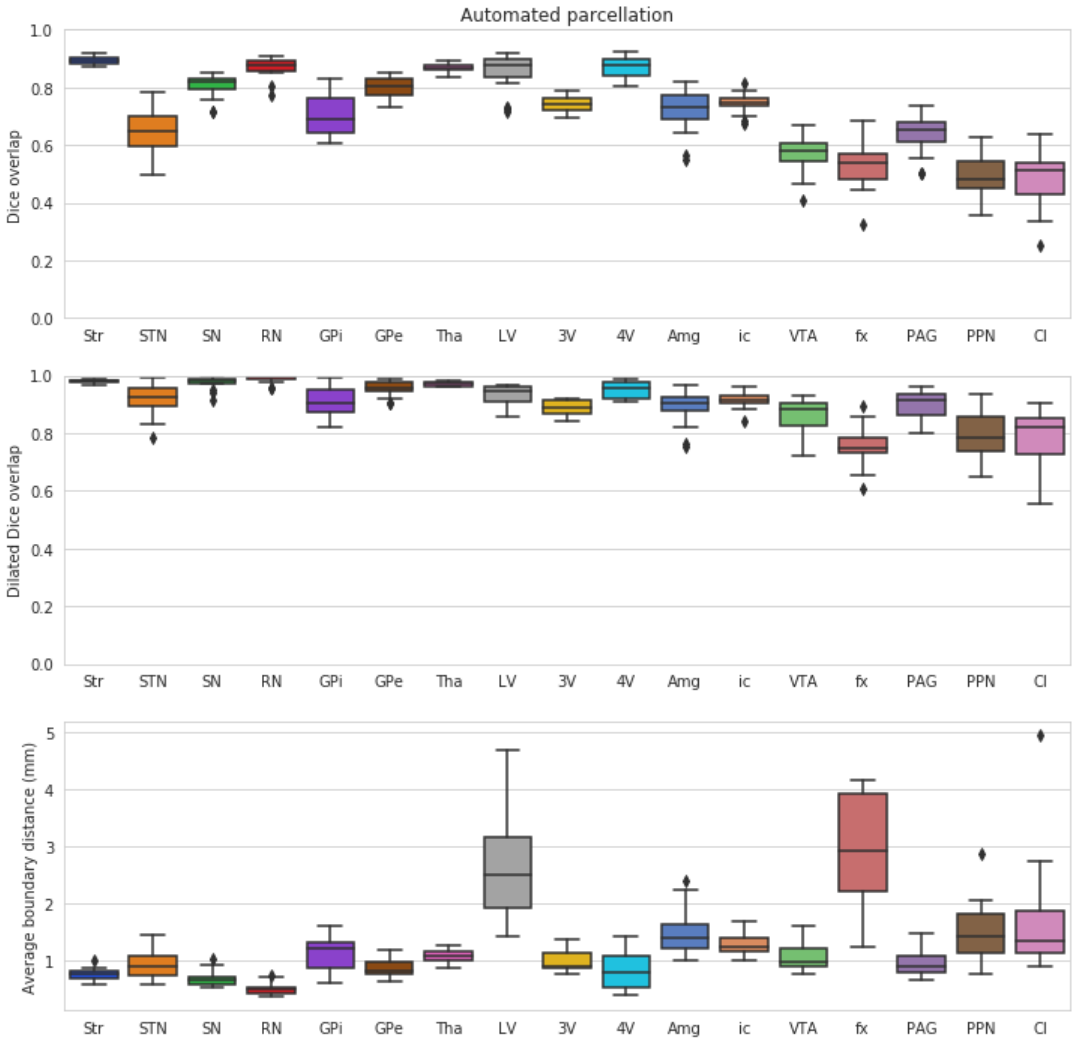

Figure 3. Leave-one-out validation of the structures parcellated by MASSP, compared to the human rater with most neuroanatomical expertise. Scores for the left and right side are computed separately and then combined into box-and-whisker plots.

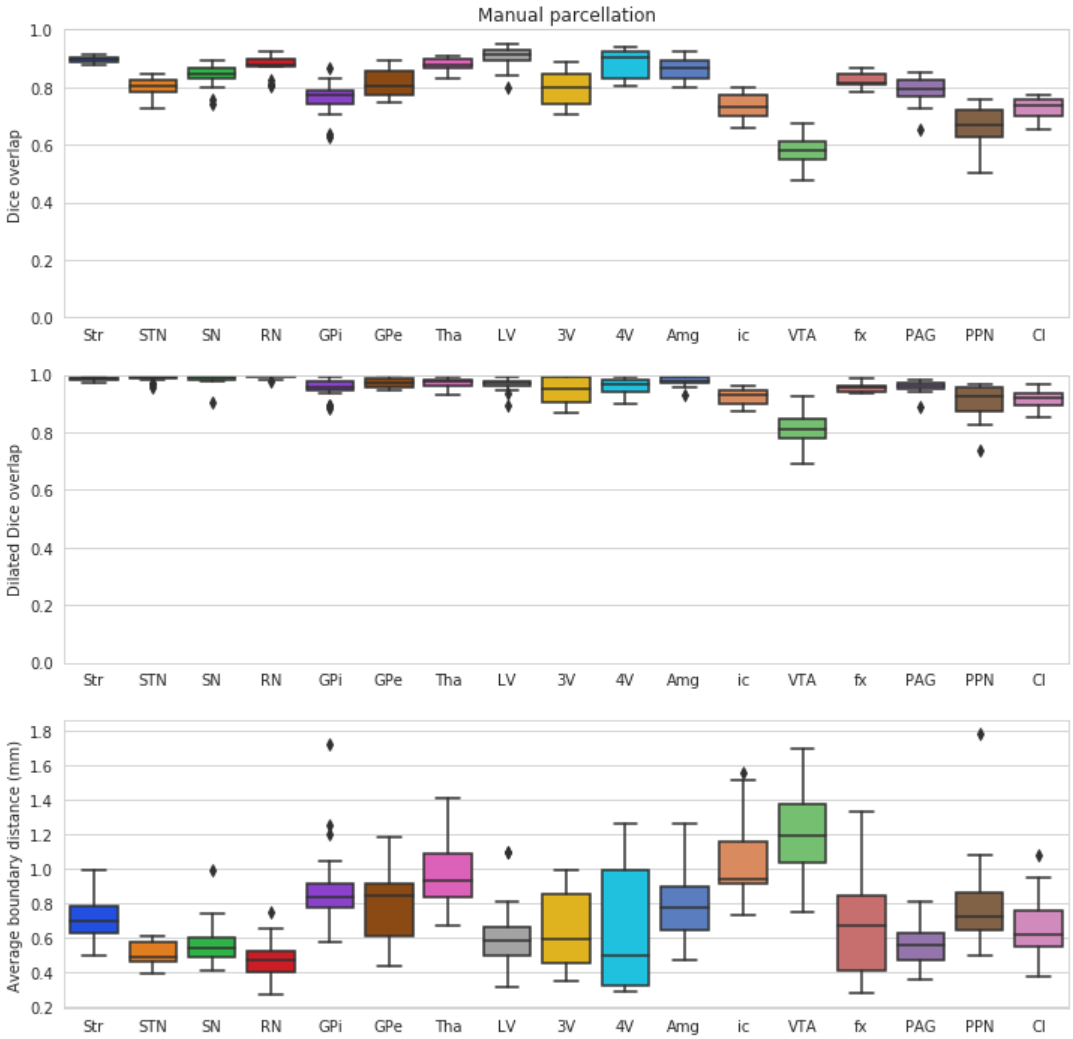

Figure 4. Inter-rater variability for the human expert raters. Scores for the left and right side are computed separately and then combined into box-and-whisker plots.

Table 1. Mean overlap and distance measures for the leave-one-out validation.

	Str	STN	SN	RN	GPi	GPe	Tha	LV	3V	4V	Amg	ic	VTA	fx	PAG	PPN	Cl
Dice overlap																	
MASSP	0.893	0.648	0.805	0.870	0.702	0.800	0.867	0.849	0.741	0.869	0.723	0.745	0.570	0.527	0.641	0.496	0.485
manual	0.897	0.800	0.841	0.875	0.762	0.813	0.877	0.907	0.797	0.882	0.866	0.732	0.574	0.823	0.791	0.665	0.727
ratio	0.995	0.811	0.957	0.996	0.925	0.987	0.989	0.936	0.936	0.988	0.836	1.020	0.994	0.641	0.814	0.754	0.664
Dilated overlap																	
MASSP	0.982	0.919	0.977	0.991	0.909	0.956	0.970	0.929	0.890	0.951	0.891	0.915	0.863	0.756	0.897	0.795	0.789
manual	0.987	0.988	0.985	0.995	0.953	0.972	0.970	0.967	0.944	0.961	0.978	0.924	0.818	0.957	0.960	0.910	0.914
ratio	0.995	0.930	0.992	0.995	0.955	0.984	1.000	0.961	0.946	0.991	0.911	0.992	1.059	0.790	0.935	0.879	0.863
Average surface distance																	
MASSP	0.750	0.911	0.676	0.491	1.140	0.863	1.058	2.690	0.994	0.817	1.476	1.275	1.074	2.950	0.955	1.484	1.685
manual	0.723	0.508	0.571	0.482	0.902	0.804	0.971	0.615	0.637	0.671	0.779	1.045	1.204	0.703	0.555	0.801	0.670
ratio	0.971	0.590	0.852	0.996	0.861	0.943	0.916	0.277	0.662	1.020	0.553	0.834	1.161	0.287	0.610	0.619	0.465

Table 2. Comparison with multi-atlas registration, Freesurfer, and FSL FIRST.

	Str	STN	SN	RN	GPi	GPe	Tha	LV	3V	4V	Amg	ic	VTA	fx	PAG	PPN	Cl
Dice overlap																	
MASSP	0.893	0.648	0.805	0.870	0.702	0.800	0.867	0.849	0.741	0.869	0.723	0.745	0.570	0.527	0.641	0.496	0.485
Multi-atlas	0.855	0.662	0.760	0.820	0.742	0.796	0.859	0.734	0.660	0.691	0.761	0.718	0.626	0.478	0.674	0.539	0.398
Freesurfer	0.876				0.778		0.838	0.858	0.430	0.769	0.692						
FSL FIRST	0.875				0.813		0.839				0.653						
Dilated overlap																	
MASSP	0.982	0.919	0.977	0.991	0.909	0.956	0.970	0.929	0.890	0.951	0.891	0.915	0.863	0.756	0.897	0.795	0.789
Multi-atlas	0.976	0.938	0.968	0.989	0.947	0.968	0.970	0.920	0.920	0.908	0.921	0.939	0.924	0.798	0.943	0.871	0.811
Freesurfer	0.975				0.922		0.946	0.974	0.562	0.911	0.857						
FSL FIRST	0.976				0.946		0.950				0.843						
Average surface distance																	
MASSP	0.750	0.911	0.676	0.491	1.140	0.863	1.058	2.690	0.994	0.817	1.476	1.275	1.074	2.950	0.955	1.484	1.685
Multi-atlas	0.961	0.891	0.858	0.675	0.992	0.882	1.083	1.417	0.932	1.249	1.359	1.129	0.813	1.362	0.794	1.055	1.273
Freesurfer	0.770				1.211		1.405	0.685	4.071	1.361	1.749						
FSL FIRST	0.867				1.143		1.675				1.746						
Volume bias																	
MASSP	0.041	0.017	*0.038	0.007	0.066	0.089	0.040	0.0470	0.121	0.047	0.078	0.183	0.032	*0.016	0.026	0.009	0.023
Multi-atlas	0.020	*0.087	*0.009	0.031	0.009	0.014	0.020	*0.003	*0.007	*0.092	*0.038	0.055	*0.067	*0.264	*0.090	*0.269	*0.376
Freesurfer	0.017				0.087		0.163	0.122	*0.551	0.351	0.468						
FSL FIRST	*0.100				*0.021		0.165				*0.249						

registration is also quite computationally intensive when using multiple contrasts at high resolution. Finally, MASSP provides many more structures than Freesurfer and FSL FIRST, and can be easily applied to new structures based on additional manual delineations.

Application to new MRI contrasts

Quantitative MRI has only become recently applicable in larger studies, thanks in part to the development of integrated multi-parameter sequences (***Weiskopf et al., 2013***; ***Caan et al., 2019***). Many data sets, including large-scale open databases, use more common T1- and T2-weighted MRI. In order to test the applicability of MASSP to such contrasts, we obtained the test-retest subset of the Human Connectome Project (HCP, ***Van Essen et al.*** (***2013***)) and applied MASSP to the 45 preprocessed and skull-stripped T1- and T2-weighted images from each of the two test and retest sessions. While performing manual delineations on the new contrasts would be preferable, the model is already rich enough to provide stable parcellations. Test-retest reproducibility is similarly high for MASSP and Freesurfer, and are generally in agreement, see Fig. 5 and Table 3.

Biases due to atlas size

A common concern of brain parcellation methods is the risk of biases, as they are typically built from a small number of manual delineations. Our data set is part of a large scale study of the subcortex, for which we obtained manual delineations of the STN, SN, RN, GPe, and GPi on 105 subjects over the adult lifespan (18-80 year old, see ***Alkemade et al.*** (***2020***) for details). First, we investigated the impact of atlas size. We randomly assigned half of the subjects from each decade to two groups, and built atlas priors from subsets of 3, 5, 8, 10, 12, 15, and 18 subjects from the first group. The subjects used in the atlas were taken randomly from each decade (18-30, 31-40,

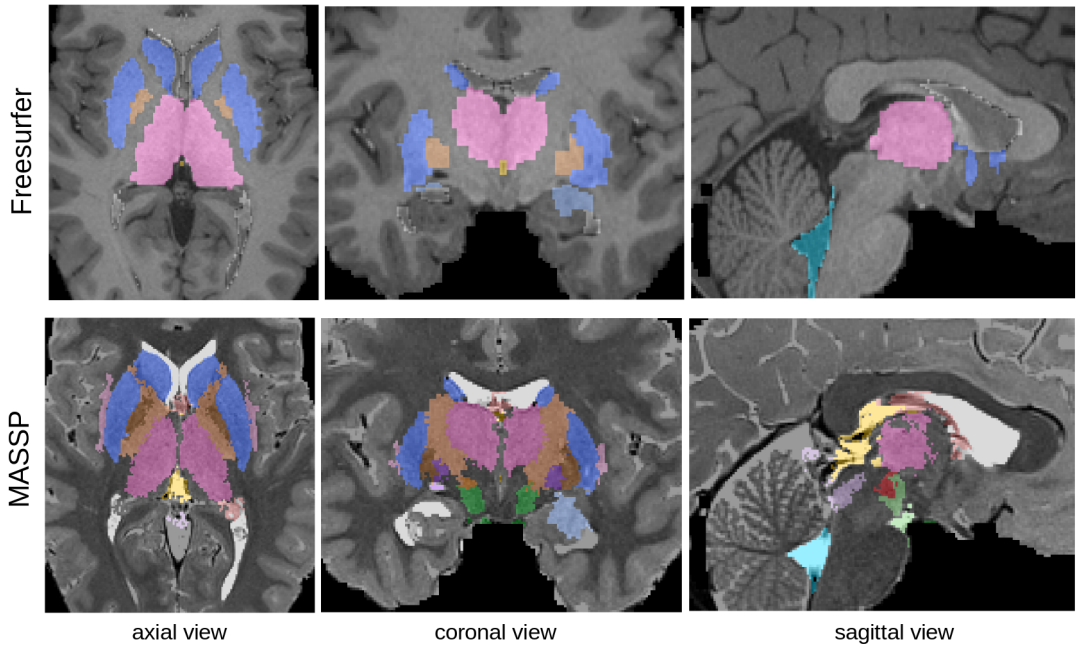

Figure 5. Parcellation with Freesurfer (top, on T1w image) and MASSP (bottom, on T2w image) on Human Connectome Project data. MASSP priors were not derived from the contrasts, but transferred via a spatial mapping of the quantitative MRI intensities from AHEAD subjects.

Table 3. Test-retest comparison with Freesurfer on Human Connectome Project data.

	Str	STN	SN	RN	GPi	GPe	Tha	LV	3V	4V	Amg	ic	VTA	fx	PAG	PPN	Cl
Dice overlap																	
MASSP test-retest	0.914	0.701	0.818	0.829	0.791	0.859	0.928	0.881	0.837	0.870	0.866	0.860	0.738	0.774	0.714	0.713	0.785
Freesurfer test-retest	0.898				0.770	0.919	0.894	0.842	0.849	0.852							
MASSP – Freesurfer	0.876				0.778	0.838	0.858	0.430	0.769	0.692							
Dilated overlap																	
MASSP test-retest	0.987	0.939	0.977	0.978	0.963	0.977	0.990	0.980	0.979	0.986	0.981	0.973	0.969	0.961	0.965	0.972	0.966
Freesurfer test-retest	0.986				0.926		0.986	0.989	0.972	0.975	0.978						
MASSP – Freesurfer	0.954				0.788		0.919	0.934	0.435	0.901	0.866						
Average surface distance																	
MASSP test-retest	0.513	0.528	0.467	0.461	0.532	0.508	0.488	0.509	0.391	0.419	0.533	0.536	0.431	0.464	0.428	0.402	0.433
Freesurfer test-retest	0.876				0.778		0.838	0.858	0.430	0.769	0.692						
MASSP – Freesurfer	0.976				1.673		1.605	1.946	5.699	1.428	1.478						

41-50,51-60, 61-70, 71-80), so as to maximize the age range represented in each atlas. Atlases of increasing size were constructed by adding subjects to previous atlases, so that atlases of increasing complexity include all subjects from simpler atlases. Results applying these atlases to parcellate the second group are given in Fig. 6. As in previous studies (***Iglesias et al., 2013***; ***Bazin and Pham, 2008***), performance quickly stabilized with atlases of more than five subjects (no significant difference in Welch's t-tests between using 18 subjects or any subset of 8 or more for all structures and measures).

Biases due to age differences

To more specifically test the influence of age on parcellation accuracy, we defined again six age groups by decade and randomly selected 10 subjects from each group. Each set of subjects was used as priors for the five structures above, and applied to the other age groups. Results are summarized in Fig. 7. Examining this age bias, we can see a decrease in performance when parcellating subjects in the range of 60 to 80 years of age. The choice of priors seem to have a limited impact, which varies across structures. In particular, using priors from a similar age group is not always beneficial.

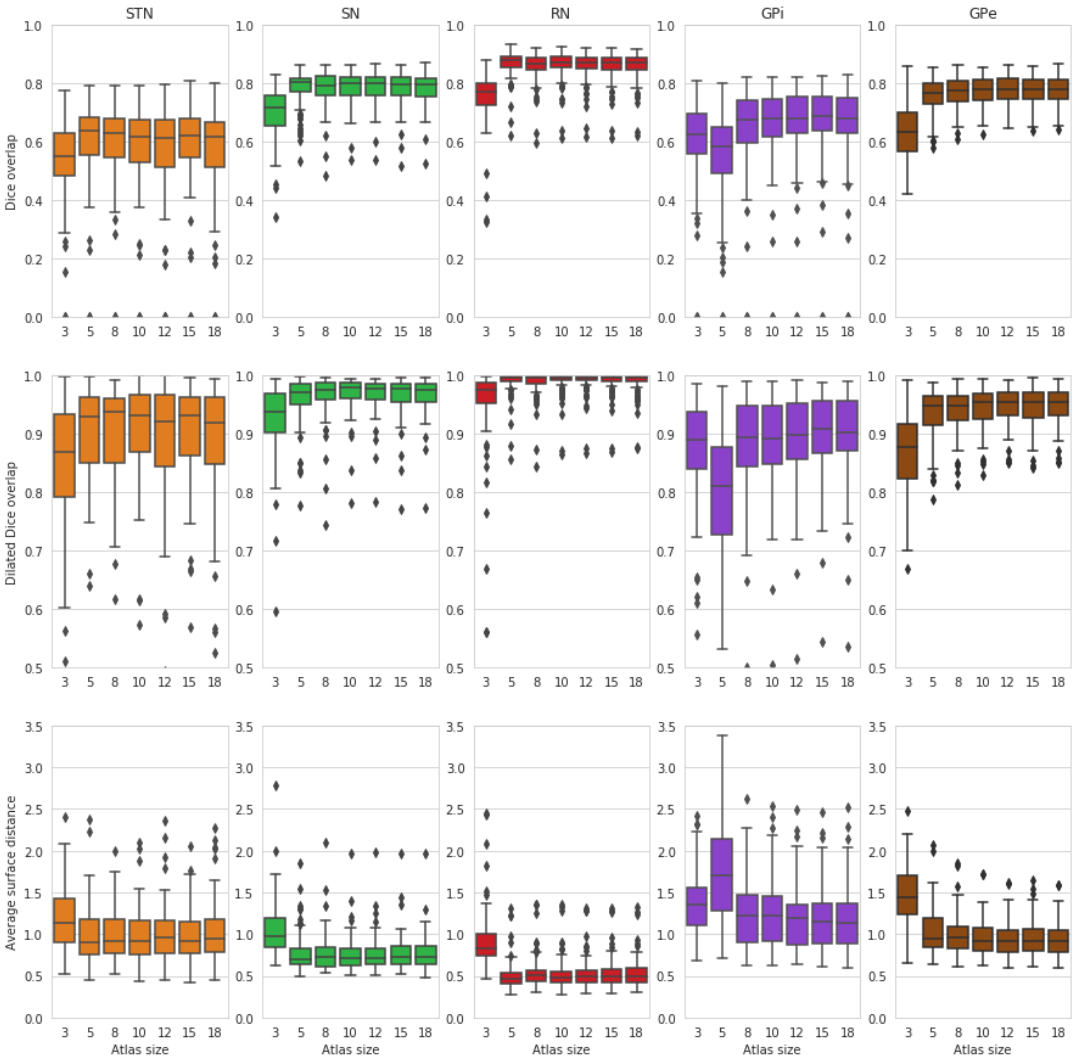

Figure 6. MASSP parcellation scores as a function of increasing number of subjects included in the atlas. Scores for the left and right side are computed separately and then combined into box-and-whisker plots.

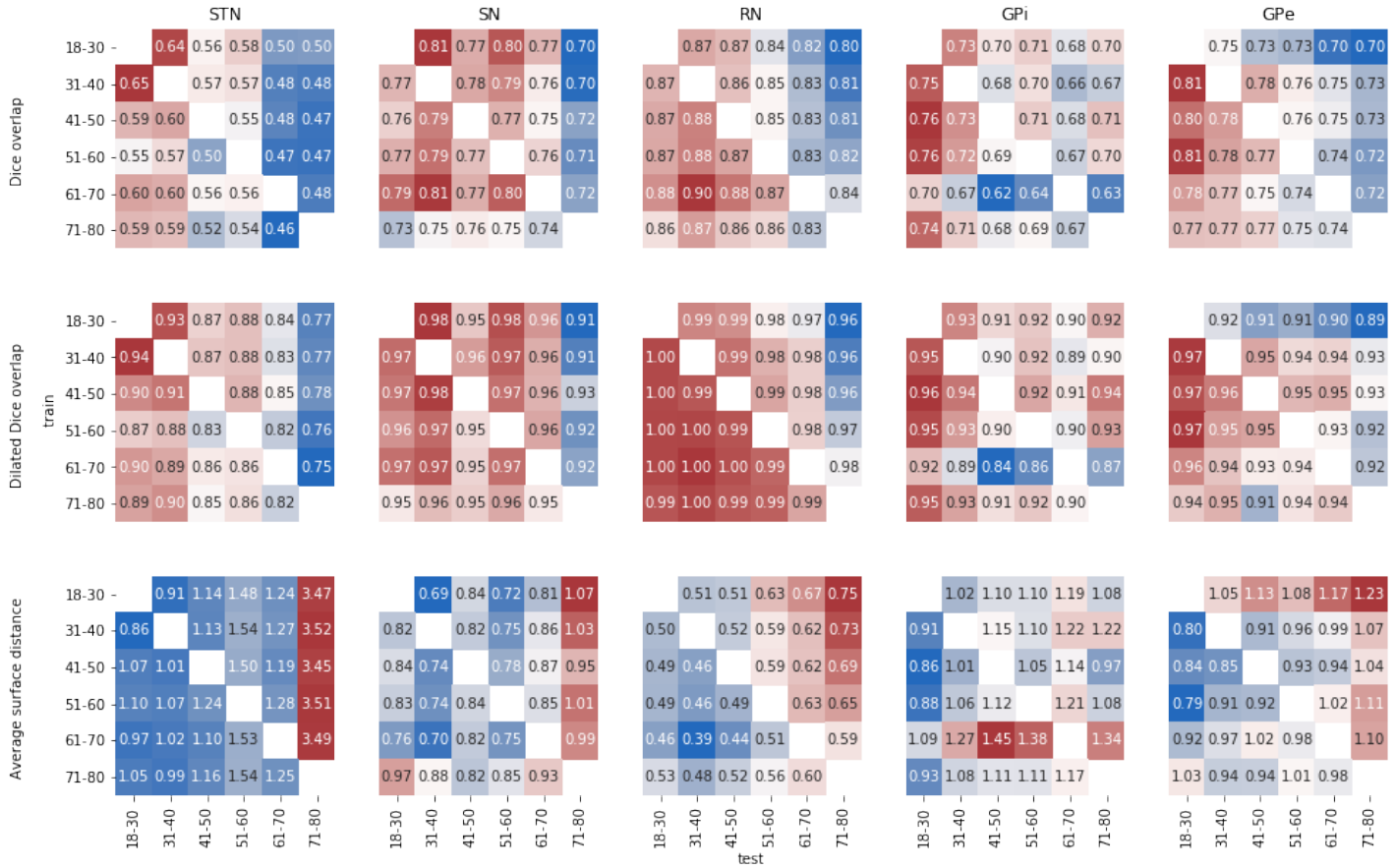

Figure 7. MASSP parcellation scores over the lifespan. Each matrix show the average Dice overlap (top), dilated Dice overlap (middle), and average surface distance (bottom) for using one age group as prior ('train') to parcellate another age group ('test').

Table 4. Mean volume and quantitative MRI values for each age group.

	Str	STN	SN	RN	GPi	GPe	Tha	LV	3V	4V	Amg	ic	VTA	fx	PAG	PPN	Cl
Volume (mm^3)																	
18-40	10656	118	566	253	567	1366	7112	7524	1895	1391	1315	4335	254	1632	250	193	843
41-60	10572	124	583	256	586	1403	7492	8850	2024	1408	1363	4495	264	1808	255	195	830
61-80	10734	130	584	260	586	1397	7463	9142	2023	1407	1321	4407	272	1910	259	192	829
Thickness (mm)																	
18-40	5.94	1.89	2.55	4.64	3.09	3.56	8.31	4.27	2.77	4.03	4.81	4.06	1.69	2.06	1.78	1.92	1.79
41-60	5.47	1.86	2.66	4.58	2.96	3.41	8.28	5.08	2.95	3.89	4.85	4.19	1.76	1.96	1.84	1.86	1.80
61-80	5.22	1.83	2.60	4.11	2.92	3.22	8.28	4.90	3.18	4.06	4.73	4.19	1.80	1.97	1.90	1.95	1.82
qR1 (Hz)																	
18-40	0.647	0.949	0.857	0.928	0.868	0.850	0.761	0.332	0.346	0.274	0.546	0.906	0.819	0.714	0.654	0.779	0.650
41-60	0.662	0.968	0.893	0.939	0.879	0.856	0.758	0.278	0.315	0.269	0.559	0.904	0.833	0.671	0.653	0.771	0.664
61-80	0.648	0.952	0.882	0.903	0.860	0.830	0.743	0.273	0.300	0.270	0.552	0.884	0.814	0.638	0.647	0.764	0.669
qR2* (Hz)																	
18-40	43.8	67.1	67.8	63.2	75.2	79.6	38.1	14.7	18.9	9.0	25.5	36.8	39.2	37.4	25.9	32.7	32.6
41-60	50.4	74.1	74.1	77.1	80.2	87.9	40.3	8.4	12.4	11.7	28.1	38.7	42.8	37.4	28.0	33.4	36.9
61-80	51.8	77.0	72.5	73.8	77.8	87.0	40.1	8.5	10.2	12.0	30.1	39.6	52.6	35.7	28.4	34.2	35.4
QSM (ppm)																	
18-40	0.0329	0.0609	0.0738	0.0717	0.1015	0.1150	0.0138	0.0130	0.0100	0.0279	0.0036	-0.0234	0.0241	0.0079	0.0119	0.0135	-0.0122
41-60	0.0400	0.0647	0.0713	0.0829	0.0984	0.1241	0.0134	0.0115	0.0025	0.0234	0.0085	-0.0226	0.0201	0.0079	0.0089	0.0099	-0.0110
61-80	0.0411	0.0705	0.0610	0.0738	0.0925	0.1249	0.0064	0.0089	-0.0034	0.0236	0.0061	-0.0243	0.0177	0.0100	0.0039	0.0096	-0.0091

Bias on individual measures

Finally, we investigated the impact of this decrease in performance in the estimation of anatomical quantities, see Fig. 8. The bias did affect the morphometric measures of structure volume and thickness, but the effects on the local measure of thickness was reduced compared to the global measure of volume. Quantitative MRI averages were very stable even when age biases are present in the parcellations.

For reference, we report structure volumes, thickness, R1, R2* and QSM values estimated from the entire AHEAD cohort for different age groups, extending our previous work based on manual delineations on a different data set (***Keuken et al., 2017***; ***Forstmann et al., 2014***). Results are given in Table 4, describing average volumes, thickness, and quantitative MRI parameters for young, middle-aged, and older subjects for the 17 subcortical structures.

Discussion

Our goal with the MASSP algorithm was to provide a fully automated method to delineate as many subcortical structures as possible on high resolution structural MRI now available on 7T scanners. We modeled 17 distinct structures, taking into account location, shape, volume, and quantitative MRI contrasts to provide individual subject parcellations. Based on our results, we can be confident that the automated parcellation technique performs comparably to human experts, providing delineations within one or two voxels of the structure boundaries (dilated Dice overlap over 75% for all structures, including in aging groups). Results were nearly indistinguishable from expert delineations for eight major structures (Str, Tha, 4V, GPe, SN, RN, VTA, ic), and smaller structures retain high levels of overlap, comparable to trained human raters. This parcellation includes the most commonly defined structures (Str, Tha, SN, RN, STN) with overlap scores comparable to those previously reported (***Garzón et al., 2018***; ***Visser et al., 2016a***,b; ***Iglesias et al., 2013***; ***Chakravarty et al., 2013***; ***Patenaude et al., 2011***). More importantly, it also includes structures seldom or never before considered in MRI atlases and parcellation methods, such as GPe, GPi, VTA, 3V, 4V, ic, fx, PAG, PPN, Cl. The technique handles structures of varying sizes well, as indicated by dilated overlap and boundary distance. Additional structures can be added, if they can be reliably delineated by expert raters on single-subject MRI at achievable resolutions. Some enhancement techniques such as building a multi-subject template (***Pauli et al., 2018***) or adding a denoising step (***Bazin et al., 2019***) may be beneficial. Co-registration to a high-precision atlas as in (***Ewert et al., 2018***) may also improve the initial alignment over the MASSP group average template.

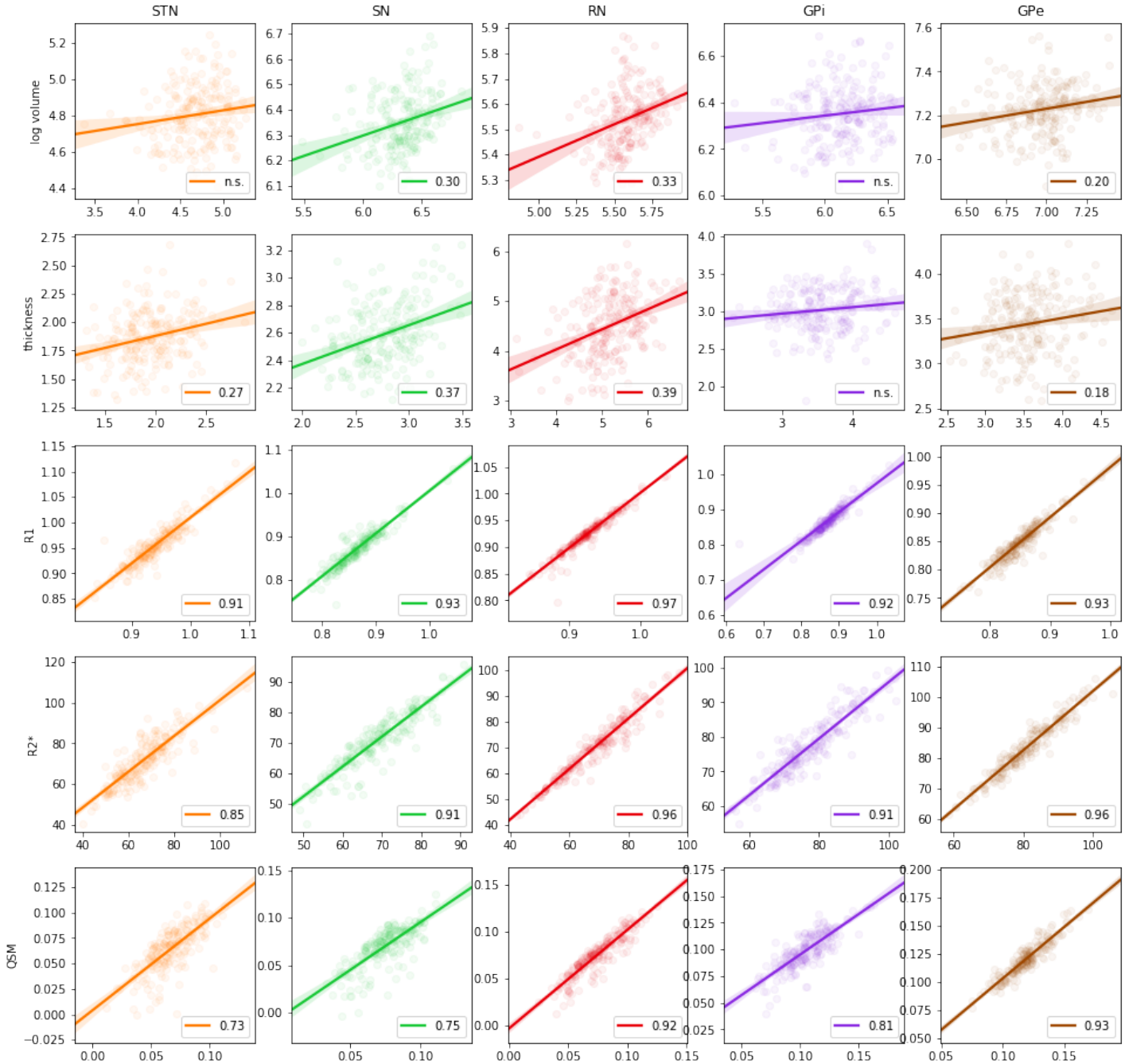

Figure 8. Regression of volume (log scale), structure thickness, R1, R2*, and QSM MRI parameters estimated using manual delineations versus MASSP automated parcellations. Circles show individual data points, linear regression is indicated by a straight line, and 95% confidence interval is given as the shaded area. Pearson correlation coefficients are indicated when significant (p-value<0.01).

Age biases are present both in expert manual delineations and automated parcellation techniques. Age trajectories in volume and quantitative MR parameters indicate systematic shifts in contrast intensities and an increasing variability with age, associated with changing myelination, iron deposition, and brain atrophy (***Draganski et al., 2011***; ***Daugherty and Raz, 2013***; ***Fjell et al., 2013***; ***Keuken et al., 2017***). These changes seem only to impact the parcellation accuracy for age groups beyond 60 years and age-matched priors did not provide specific improvements, thus indicating that an explicit modeling of age effects may be required to further improve parcellation quality in elderly populations. These results also point to exercising caution when applying automated parcellation methods to study morphometry in elderly or diseased populations, where measured differences may include biases. They also point out that while global volume and local thickness are indeed affected by such biases, quantitative MRI measures are much more robust. Note that this bias is likely present is many automated methods, although they have not been systematically investigated due to the extensive manual labor required. Interestingly, biases also exist in expert delineations: when the size or shape of a structure is refined in neuroanatomical studies, experts may become more or less conservative in their delineations. Automated methods provide a more objective measure in such case, as the source of their bias is explicitly encoded in the atlas prior delineations and computational model. Important applications of subcortical parcellation also include deep-brain stimulation surgery (***Ewert et al., 2018***), where the number of structures parcellated by MASSP can help neurosurgeons orient themselves more easily, although precise targeting will still require manual refinements, especially in neurodegenerative diseases.

We observed that dilated overlap, i.e. the overlap of structures up to one voxel, provided a measure of accuracy largely independent of size, for automated or manual delineations. Imprecision in the range of one voxel in the boundary is to be expected due to partial voluming which impacts Dice overlap. The dilated overlap measure is a better representative of performance, and indicates that conservative or inclusive versions of the subcortical regions can be obtained by eroding or dilating the estimated boundary by a single voxel. Such masks may be useful when separating functional MRI signals between neighboring nuclei or when locating smaller features inside a structure. Additionally, the Bayesian estimation framework provides voxel-wise probability values, which can also be used to further weight the contribution of each voxel within a region in subsequent analyses.

In summary, our method provides fast and accurate parcellation for subcortical structures of varying size, taking advantage of the high resolution offered by 7T and the specificity of quantitative MRI. The algorithm is based on an explicit model of structures given in a Bayesian framework, and is free of tuning parameters. Given a different set of regions of interest or different populations, new priors can be automatically generated and used as the basis for the algorithm. If more MRI contrasts are available, the method can also be augmented to take them into account. The main requirement for the technique is a set of manual delineations of all the structures of interest in a small group of representative subjects. Performance may further improve with the number of included structures, as the number of distinct interfaces increases, refining in particular the intensity priors. In future works, we plan to include more structures or sub-structures and model the effects of age on the priors. We hope that the method, available in open source, will help neuroscience researchers to include more subcortical regions in their structural and functional imaging studies.

Materials and Methods

Data acquisition

Our parcellation method has been developed for the MP2RAGEME sequence (***Caan et al., 2019***). Briefly, the MP2RAGEME consists of two interleaved MPRAGEs with different inversions and four echoes in the second inversion. Based on these images, one can estimate quantitative MR parameters of R1, R2* and QSM. In this work, we used the following sequence parameters: inversion times TI1,2 = 670 ms, 3675.4 ms; echo times TE1= 3 ms, TE2,1-4 = 3, 11.5, 19, 28.5 ms; flip angles FA1,2 = 4°, 4°; TRGRE1,2 = 6.2 ms, 31 ms; bandwidth = 404.9 MHz; TRMP2RAGE = 6778 ms; SENSE

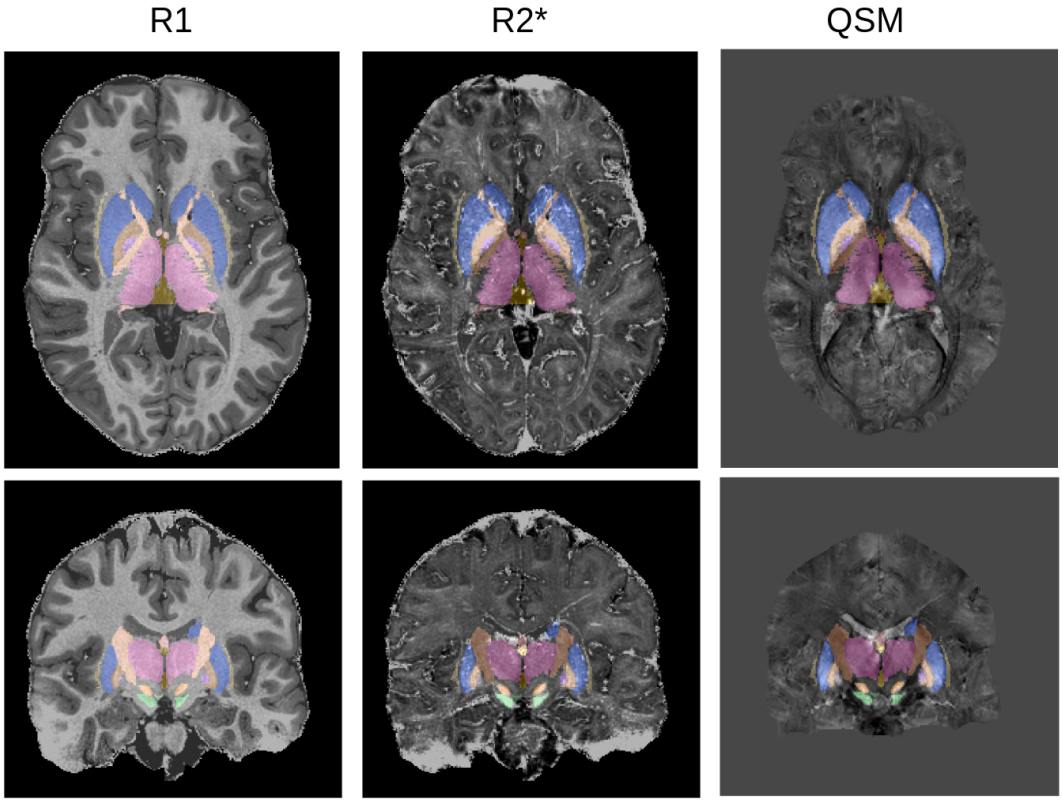

Figure 9. MP2RAGEME maps and delineations: quantitative R1 (left), quantitative R2* (middle), QSM (right). Manual delineations for the 17 structures of interest are overlaid on all images.

acceleration factor = 2; FOV = 205 x 205 x 164mm; acquired voxel size = 0.70 x 0.7 x 0.7 mm; acquisition matrix was 292 x 290; reconstructed voxel size = 0.64 x 0.64 x 0.7 mm; turbo factor (TFE) = 150 resulting in 176 shots; total acquisition time = 19.53 min.

T1-maps were computed using a look-up table (***Marques et al., 2010***). T2*-maps were computed by least-squares fitting of the exponential signal decay over the multi-echo images of the second inversion. R1 and R2* maps were obtained as the inverse of T1 and T2*. For QSM, phase maps were pre-processed using iHARPERELLA (integrated phase unwrapping and background phase removal using the Laplacian) of which the QSM images were computed using LSQR (***Li et al., 2014***). Skull information was removed through creation of a binary mask using FSL's brain extraction tool on the reconstructed uniform T1-weighted image and then applied to the quantitative contrasts (***Smith, 2002***). As all images were acquired as part of a single sequence, no co-registration of the quantitative maps was required (see Fig. 9).

Anatomical structure delineations

Manual delineations of subcortical structures were performed by two raters trained by an expert anatomist, according to protocols optimized to use the better contrast or combination of contrasts for each structure and to ensure a consistent approach across raters. The following 17 structures were defined on a group of ten subjects (average age 24.4, 8 female): striatum (Str), thalamus (Tha), lateral, 3rd and 4th ventricles (LV, 3V, 4V), amygdala (Amg), globus pallidus internal segment (GPi) and external segment (GPe), SN, STN, red nucleus (RN), ventral tegmental area (VTA), fornix (fx), internal capsule (ic), periaqueductal gray (PAG), pedunculopontine nucleus (PPN), and claustrum (Cl). Separate masks for left and right hemisphere were delineated except for 3V, 4V, and fx. In the following the algorithm treats each side separately, resulting in a total of 31 distinct structures (see Fig. 1).

Anatomical interface priors

In order to inform the algorithm, we built a series of priors derived from the manual delineations. Each subject was first co-registered to a MP2RAGEME anatomical template built from 105 subjects co-aligned with the MNI2009b atlas (***Fonov et al., 2011***) with the SyN algorithm of ANTs (***Avants et al., 2008***) using successively rigid, affine, and non-linear transformations, high levels of regularisation as recommended for the subcortex (***Ewert et al., 2019***) and mutual information as cost function.

The first computed prior is a prior of anatomical interfaces, recording the most likely location of boundaries between the different structures, defined as follows. Given two delineated structures i, j, let φ_i, φ_j be the signed distance functions to their respective boundary, i.e. $\varphi_i.x/$ is the Euclidean distance of any given voxel to the boundary of i, with a negative sign inside the structure. Then we define the interface B_{ij} with the distance function d_{ij}:

$$d_{ij}.x/ = \min\left(\varphi_i.x/, \varphi_j.x/ * \delta, 0/\right. \tag{1}$$

where δ is a scale parameter for the thickness of the interface. These interfaces functions are not symmetrical, as the intensity inside i next to j is generally different from the intensity inside j next to i. Based on this definition, the prior for a given interface based on N manual delineations is given by:

$$\begin{gathered} P.x \text{ Ë } B_{ij}/ \text{ í } \frac{1}{\sqrt{2\pi\sigma_{ij}^2.x/}} \exp * \frac{1}{2} \frac{\mu_{ij}^2.x/}{\sigma_{ij}^2.x/} \\ \mu_{ij}.x/ = \frac{1}{N} {}_{n\text{ËN}} d_{ij,n}.x/,\ \sigma_{ij}.x/ = \sqrt{\frac{1}{N} {}_{n\text{ËN}} \left(d_{ij,n}.x/ * \mu_{ij}.x//\right.^2 + \delta} \end{gathered} \tag{2}$$

These probability functions are calculated for all possible configurations including ii, which represent the inside of each structure. We thus have a total of N^2 functions, but only a few are non-zero at a given voxel x, and we may keep only the 16 largest values to account for any number of interfaces in 3D (***Bazin et al., 2007***). Finally, we need to scale the prior to be globally consistent with the priors below by assuming that the 95th percentile of the highest kept $P.x$ Ë $B_{ij}/$ values have a probability of 0.95. The scale parameter δ is set to 1 voxel, representing the expected amount of partial voluming. The resulting interface prior is shown in Fig. 10A.

Anatomical skeleton priors

Next, we defined priors for the skeleton of each structure, representing their essential shape regardless of exact boundaries (Blum, 1973). As we are mostly interested in the most likely components of the skeleton or medial axis S_i, we follow a simple method to estimate its location:

$$S_i = \left| x, (\varphi_i.x/ < \frac{1}{2} \right| \tag{3}$$

We define as $s_i.x/$ the signed distance function of this discrete skeleton, and define prior probabilities as above:

$$\begin{gathered} P.x \text{ Ë } S_i/ \text{ í } \frac{1}{\sqrt{2\pi\sigma_i^2.x/}} \exp * \frac{1}{2} \frac{\mu_i^2.x/}{\sigma_i^2.x/} \\ \mu_i.x/ = \frac{1}{N} {}_{n\text{ËN}} s_{i,n}.x/\ \sigma_i.x/ = \sqrt{\frac{1}{N} {}_{n\text{ËN}} \left(s_{i,n}.x/ * \mu_i.x//\right.^2 + \delta} \end{gathered} \tag{4}$$

The skeletons are defined inside each structure, which implies $P.x$ Ë $S_i/ \leq P.x$ Ë $B_{ii}/$. To respect this relationship, we scale $P.x$ Ë $S_i/$ with the same factor as $P.x$ Ë $B_{ii}/$ but use $\sqrt{P.x \text{ Ë } S_i/}$ when combining probabilities during the estimation stage. The obtained anatomical skeleton priors are given on Fig. 10B.

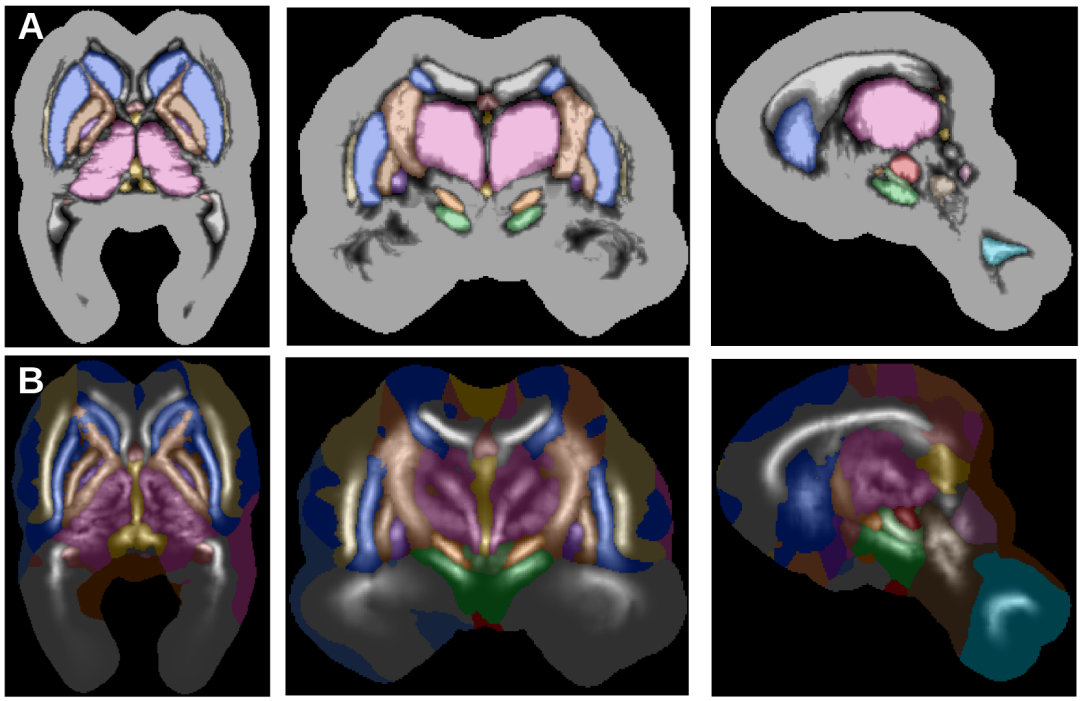

Figure 10. Anatomical interface (A) and skeleton (B) priors derived from the 10 manually delineated subjects.

Interface intensity priors

While anatomical priors already provide rich information, they are largely independent of the underlying MRI. From the co-aligned quantitative MRI maps and manual delineations, we defined intensity priors for every interface ij, in the form of intensity histograms to ensure a flexible representation of intensity distributions. Given a quantitative contrast $R_n.x/$, we built a histogram $H_{ij,n}$ for each subject n and interface ij. Histograms have 200 bins covering the entire intensity range within a radius of $10mm$ from any of the delineated structures. To obtain an average histogram, we combine each histogram with a weighting function $w_n.x/$ giving the likelihood of the subject's intensity measurement compared to the group:

$$w_{ij,n}.x/ = P.x \text{ Ë } B_{ij}/\frac{1}{\sqrt{2\pi\sigma_R^2}} \exp * \frac{1}{2} \frac{.R_n.x/ * \mu_R.x//^2}{\sigma_R.x/^2} \tag{5}$$

where $\mu_R.x/$ is the median of the $R_n.x/$ values at x, and $\sigma_R.x/$ is 1.349 times the inter-quartile range of $R_n.x/$. These are robust estimators of the mean and standard deviation, used here to avoid biases by intensity outliers. To further combine the R1, R2*, and QSM contrasts we take the geometric mean of the histogram probabilities: $H_{ij}.x/ = {}_R\, H_{ij}.R.x//^{1\cdot 3}$.

Global volume priors

The last type of priors extracted from manual delineations are volume priors for each of the structure. Here we assume a log-normal distribution for the volumes V_i and simply estimate the mean $\mu_{V,i}$ and standard deviation $\sigma_{V,i}$ of $\log V_{i,n}$ over the subjects.

Voxel-wise posterior probabilities

When parcellating a new subject, we first co-register its R1, R2*, and QSM maps jointly to the template and use the inverse transformation to deform the anatomical priors into subject space. Then

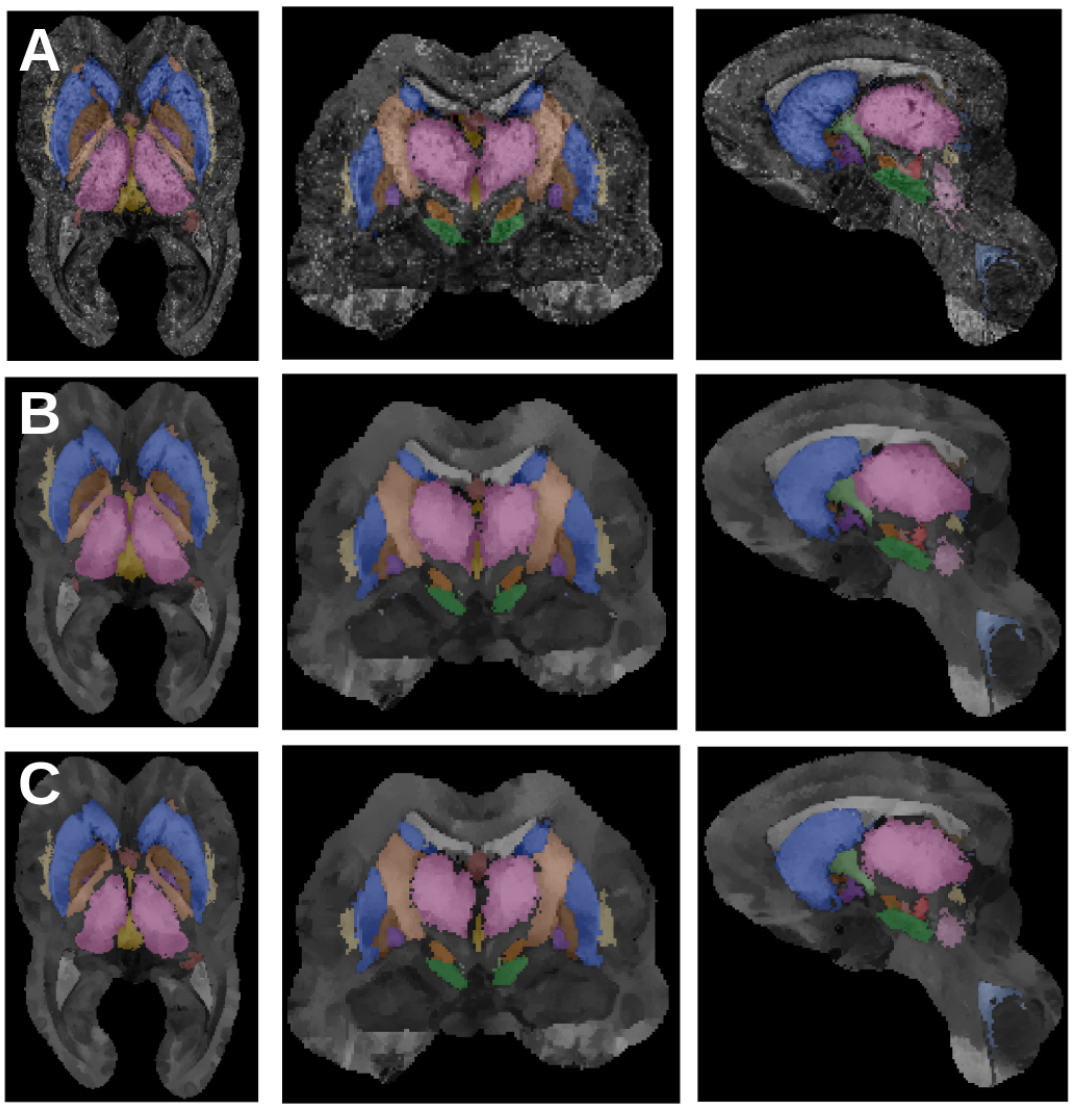

Figure 11. Successive parcellation results: A) voxel-wise posteriors and parcellation, B) diffused posteriors and parcellation, C) topology-corrected posteriors and final region-growing parcellation.

we derive voxel-wise posteriors as follows:

$$P.x \text{ Ë } B_{ij}R.x// \text{ í } P.x \text{ Ë } B_{ij}/H_{ij}.x/ \text{ if } i \neq j$$
$$\text{and} \tag{6}$$
$$P.x \text{ Ë } B_{ii}S_i.x/, R.x// \text{ í } \max.P.x \text{ Ë } B_{ii}/, P.x \text{ Ë } S_i/^{1^{\cdot}2}/H_{ii}.x/$$

Once again we should compute all possible combinations, but due to the multiplication of the priors we can restrict ourselves to the 16 highest probabilities previously estimated. To balance the contribution of the anatomical priors and the intensity histograms, we also need to normalize the intensity priors sampled on the subject's intensities. We use the same approach, namely assuming that the 95th percentile of the highest kept $H_{ij}.x/$ values have a probability of 0.95, separately for each contrast. The voxel-wise parcellation and posteriors obtained are shown in Fig. 11A.

Markovian diffusion

The voxel-wise posteriors are independent from each other, and do not reflect the continuous nature of the structures. The next step is to combine information from neighboring voxels. We

define a sparse Markov Random Field model for the posteriors:

$$P.x \text{ Ë } B_{ij}R, S, C/ = {}_{y\text{Ë}C.x/} P.y \text{ í } xR/P.y \text{ Ë } B_{ij}R, S, C/ \tag{7}$$

with $P.y$ í $xR/ = {}_R \exp{}^*.R.y/ * R.x//^{2\cdot}2\sigma_R^2$, where σ_R is the median of the standard deviations $\sigma_{ij,R}$ of the contrast histograms $H_{ij}.R.x//$. The neighborhood $C.x/$ is defined as x itself and the four 26-connected neighboring voxels with highest probability $P.y$ í $xR/$, thus representing the neighbors most likely to be connected to x. The model is similar to a diffusion process and can be estimated with an iterated conditional modes (ICM) approach, updating sequentially the probabilities (Li, 2009):

$$P.x \text{ Ë } B_{ij}R, S, C/ \leftarrow {}_{y\text{Ë}C.x/} P.y \text{ í } xR/P.y \text{ Ë } B_{ij}R, S, C/ \tag{8}$$

from the initial voxel-wise posteriors until the ratio of changed parcellation labels decreases below 0.1, typically within 50-80 iterations. The diffused probabilities and parcellation are shown in Fig. 11B.

Topology correction

The final step of the parcellation algorithm takes a global view of the individual structures, growing from the highest posterior values inside toward the boundaries. This region growing approach makes the implicit assumption that posterior maps should be monotonically decreasing from inside to outside, which is not necessarily the case. Therefore, we perform first a topology correction step on the individual structure posteriors $P.x \text{ Ë } iR, B, S, C/ = \max_{ij} P.x \text{ Ë } B_{ij}R, S, C/$ with a fast marching algorithm (***Bazin and Pham, 2007***). While the corrected posterior is very similar to the original one (see Fig. 11C), it ensures that all regions obtained by growing to a threshold have spherical object topology.

Anatomical region growing

Last, we turn the posteriors into optimized parcellations, by growing them concurrently (to avoid overlaps) until the target volume for each structure is reached. Given the volume $V_i.R, B, S, C/$ of the parcellation of the diffused and topology-corrected posteriors, we define the following target volume:

$$V_i = P.V_i\mu_{V,i}, \sigma_{V,i}/V_i + .1 * P.V_i\mu_{V,i}, \sigma_{V,i}// \exp \mu_{V,i} \tag{9}$$

taking a weighted average of the volume estimated from the data and the prior volume. This approach ensures that even in extreme cases where some structures have low posteriors, they are still able to grow to a plausible size. The region growing algorithm is driven from the most likely voxels, defined as $P.x \text{ Ë } iR, B, S, C/ * \max_{j\neq i} P.x \text{ Ë } jR, B, S, C/$, and further modulated to follow isocontours of the skeleton prior:

$$\begin{aligned} P.x \leftarrow y/ \text{ í } \quad & P.y \text{ Ë } iR, B, S, C/ * \max_{j\neq i} P.y \text{ Ë } jR, B, S, C/ \\ & *P.y \text{ Ë } S_i/ * P.x \text{ Ë } S_i/ \end{aligned} \tag{10}$$

Directionality of internal structures is a useful tool for understanding mechanical function in bones (***Maquer et al., 2015***). Here, we adapt this concept by using the skeleton isocontours as a representation of internal directionality, maintaining the intrinsic shape of structures. Thus, voxels with highest probability compared to the other structures and with similar distance to the internal skeleton are preferentially selected. The final parcellation is given in Fig. 11C.

Validation metrics

To validate the method against manual expert delineations, we compared the MASSP results and the expert delineations with the following three measures:

1. the Dice overlap coefficient (Dice, 1945) $D.A, B/ = \frac{2A\text{ã}B}{A+B}$, which measures the strict overlap between voxels in both delineation;

2. the dilated overlap coefficient $dD.A, B/ = \frac{A\ddot{a}d.B/+B\ddot{a}d.A/}{A+B}$, where $d../$ is a dilation of the delineation by one voxel, which measures the overlap between delineations allowing for one voxel of uncertainty;
3. the average surface distance $asd.A, B/$, measuring the average distance between voxels on the surface boundary of the first delineation to the other one and reciprocally, which measures the distance between both delineations.

We computed all three measures for the manual delineations from the two independent raters, as well as the ratio of overlaps (automated over manual) and distances (manual over automated) to compare both performances, as detailed in the Results section.

Comparisons with other automated methods

To assess the performance of MASSP compared to existing parcellation tools, we ran Freesurfer (***Fischl et al., 2002***), FSL FIRST (***Patenaude et al., 2011***) and a multi-atlas registration approach (co-registering 9 of the 10 manually delineated subjects on the remaining one with ANTs (***Avants et al., 2008***) and labelling each structure by majority voting, similarly to the MAGeT Brain approach of (***Chakravarty et al., 2013***)). Freesurfer and FIRST were run on the skull-stripped R1 map, while the multi-atlas approach used all three R1, R2* and QSM contrasts. All methods were compared in terms of Dice overlap, dilated overlap and average surface distance. We also assessed the presence of a systematic volume bias, defined as the average of the signed difference of the estimated structure volume to the manually delineated volume, normalized by the manually delineated volume.

Application to new MRI contrasts

Before applying MASSP to unseen contrasts, we need to convert its intensity prior histograms $H_{ij,R}$ to the new intensities. In order to perform this mapping, we first created a groupwise median of the HCP subjects, by co-registering every subject to the MASSP template using ANTs with non-linear registration and both T1w, T2w contrasts matched to the template's R1 and R2* maps. The histogram bins are then updated as follows:

$$H_{\text{bin},ij,R} \equiv_{xR.x/\ddot{\text{E}}\text{bin}} P.x\ \ddot{\text{E}}\ B_{ij}/H_{ij,R1}H_{ij,R2\text{i}}H_{ij,QSM} \tag{11}$$

adding the joint probability of the quantitative contrasts weighted by their importance for each interface to define the new intensity histograms. This model is essentially projecting the joint likelihood of the MASSP contrasts onto the new contrasts, assuming that the co-registration between the two is accurate enough. With these new histograms, we compared the test-retest reliability and overall agreement of MASSP with Freesurfer parcellations included in the HCP pre-processed data set.

Measurement of structure thickness

Finally, when comparing derived measures obtained over the lifespan with MASSP compared to manual delineations, we explored the utility of a shape thickness metric, based on the medial representation. Given the signed distance function φ_i of the structure boundary and s_i of the structure skeleton, the thickness is given by:

$$th_i.x/ = 2\left(s_i.x/ * \varphi_i.x//\right. \tag{12}$$

Like in cortical morphometry, thickness is a local measure, defined everywhere inside the structure, and expected to provide additional information about anatomical variations. Indeed, a similar measure of shape thickness has recently been able to highlight subtle anatomical changes in depression (***Ho et al., 2020***).

Software implementation

The proposed method, Multi-contrast Anatomical Subcortical Structure Parcellation (MASSP), has been implemented as part of the Nighres toolbox (***Huntenburg et al., 2018***), using Python and Java for optimized processing. The software is available in open source from (release-1.3.0) and . A complete parcellation pipeline is included with the Nighres examples. Computations take under 30 minutes per subject on a modern workstation.

Acknowledgments

We would like to thank Josephine Groot, Nikita Berendonk, Nicky Lute for their help collecting the AHEAD database, and Wietske van der Zwaag and Matthan Caan for their help in setting up the MP2RAGEME sequence. We also thank Steven Miletić and Dagmar Timmann for stimulating discussions around this topic, and all undergraduate students who contributed to the manual delineations. This work was supported by a NWO Vici grant (BF) and a NWO STW grant (AA, BF). HCP data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

References

Ahsan RL, Allom R, Gousias IS, Habib H, Turkheimer FE, Free S, Lemieux L, Myers R, Duncan JS, Brooks DJ. Volumes, spatial extents and a probabilistic atlas of the human basal ganglia and thalamus. NeuroImage. 2007 Nov; 38(2):261–270. https://linkinghub.elsevier.com/retrieve/pii/S1053811907005307, doi: 10.1016/j.neuroimage.2007.06.004.

Alkemade A, Mulder MJ, Groot JM, Isaacs BR, van Berendonk N, Lute N, Isherwood SJ, Bazin PL, Forstmann BU. 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. https://linkinghub.elsevier.com/retrieve/pii/S1053811920306868, doi: 10.1016/j.neuroimage.2020.117200.

Andersen HH, Johnsen KB, Moos T. Iron deposits in the chronically inflamed central nervous system and contributes to neurodegeneration. Cellular and Molecular Life Sciences. 2014 May; 71(9):1607–1622. http://link.springer.com/10.1007/s00018-013-1509-8, doi: 10.1007/s00018-013-1509-8.

Avants B, Epstein C, Grossman M, Gee J. Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. Medical Image Analysis. 2008 Feb; 12(1):26–41. https://linkinghub.elsevier.com/retrieve/pii/S1361841507000606, doi: 10.1016/j.media.2007.06.004.

Bazin PL, Alkemade A, van der Zwaag W, Caan M, Mulder M, Forstmann BU. Denoising High-Field Multi-Dimensional MRI With Local Complex PCA. Frontiers in Neuroscience. 2019 Oct; 13. https://www.frontiersin.org/article/10.3389/fnins.2019.01066/full, doi: 10.3389/fnins.2019.01066.

Bazin PL, Ellingsen LM, Pham DL. Digital Homeomorphisms in Deformable Registration. In: Karssemeijer N, Lelieveldt B, editors. *Information Processing in Medical Imaging*, vol. 4584 Berlin, Heidelberg: Springer Berlin Heidelberg; 2007.p. 211–222. http://link.springer.com/10.1007/978-3-540-73273-0_18, doi: 10.1007/978-3-540-73273-0_18.

Bazin PL, Pham DL. Topology correction of segmented medical images using a fast marching algorithm. Computer Methods and Programs in Biomedicine. 2007 Nov; 88(2):182–190. https://linkinghub.elsevier.com/retrieve/pii/S0169260707002106, doi: 10.1016/j.cmpb.2007.08.006.

Bazin PL, Pham DL. Homeomorphic brain image segmentation with topological and statistical atlases. Medical Image Analysis. 2008 Oct; 12(5):616–625. https://linkinghub.elsevier.com/retrieve/pii/S1361841508000601, doi: 10.1016/j.media.2008.06.008.

Bazin PL, Plessis V, Fan AP, Villringer A, Gauthier CJ. Vessel segmentation from quantitative susceptibility maps for local oxygenation venography. In: *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)* Prague, Czech Republic: IEEE; 2016. p. 1135–1138. http://ieeexplore.ieee.org/document/7493466/, doi: 10.1109/ISBI.2016.7493466.

Caan MWA, Bazin P, Marques JP, Hollander G, Dumoulin SO, Zwaag W. MP2RAGEME: T_1, T_2^*, and QSM mapping in one sequence at 7 tesla. Human Brain Mapping. 2019 Apr; 40(6):1786–1798. https://onlinelibrary.wiley.com/doi/abs/10.1002/hbm.24490, doi: 10.1002/hbm.24490.

Chakravarty MM, Bertrand G, Hodge CP, Sadikot AF, Collins DL. The creation of a brain atlas for image guided neurosurgery using serial histological data. NeuroImage. 2006 Apr; 30(2):359–376. https://linkinghub.elsevier.com/retrieve/pii/S1053811905006993, doi: 10.1016/j.neuroimage.2005.09.041.

Chakravarty MM, Steadman P, van Eede MC, Calcott RD, Gu V, Shaw P, Raznahan A, Collins DL, Lerch JP. Performing label-fusion-based segmentation using multiple automatically generated templates: MAGeT Brain: Label Fusion Segmentation Using Automatically Generated Templates. Human Brain Mapping. 2013 Oct; 34(10):2635–2654. http://doi.wiley.com/10.1002/hbm.22092, doi: 10.1002/hbm.22092.

Daugherty A, Raz N. Age-related differences in iron content of subcortical nuclei observed in vivo: A meta-analysis. NeuroImage. 2013 Apr; 70:113–121. https://linkinghub.elsevier.com/retrieve/pii/S1053811912012268, doi: 10.1016/j.neuroimage.2012.12.040.

Draganski B, Ashburner J, Hutton C, Kherif F, Frackowiak RSJ, Helms G, Weiskopf N. Regional specificity of MRI contrast parameter changes in normal ageing revealed by voxel-based quantification (VBQ). NeuroImage. 2011 Apr; 55(4):1423–1434. https://linkinghub.elsevier.com/retrieve/pii/S1053811911000887, doi: 10.1016/j.neuroimage.2011.01.052.

Ewert S, Horn A, Finkel F, Li N, Kühn AA, Herrington TM. Optimization and comparative evaluation of nonlinear deformation algorithms for atlas-based segmentation of DBS target nuclei. NeuroImage. 2019 Jan; 184:586–598. https://linkinghub.elsevier.com/retrieve/pii/S1053811918318664, doi: 10.1016/j.neuroimage.2018.09.061.

Ewert S, Plettig P, Li N, Chakravarty MM, Collins DL, Herrington TM, Kühn AA, Horn A. Toward defining deep brain stimulation targets in MNI space: A subcortical atlas based on multimodal MRI, histology and structural connectivity. NeuroImage. 2018 Apr; 170:271–282. https://linkinghub.elsevier.com/retrieve/pii/S105381191730407X, doi: 10.1016/j.neuroimage.2017.05.015.

Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, van der Kouwe A, Killiany R, Kennedy D, Klaveness S, Montillo A, Makris N, Rosen B, Dale AM. Whole Brain Segmentation. Neuron. 2002 Jan; 33(3):341–355. https://linkinghub.elsevier.com/retrieve/pii/S089662730200569X, doi: 10.1016/S0896-6273(02)00569-X.

Fjell AM, Westlye LT, Grydeland H, Amlien I, Espeseth T, Reinvang I, Raz N, Holland D, Dale AM, Walhovd KB. Critical ages in the life course of the adult brain: nonlinear subcortical aging. Neurobiology of Aging. 2013 Oct; 34(10):2239–2247. https://linkinghub.elsevier.com/retrieve/pii/S0197458013001590, doi: 10.1016/j.neurobiolaging.2013.04.006.

Fonov V, Evans AC, Botteron K, Almli CR, McKinstry RC, Collins DL. Unbiased average age-appropriate atlases for pediatric studies. NeuroImage. 2011 Jan; 54(1):313–327. https://linkinghub.elsevier.com/retrieve/pii/S1053811910010062, doi: 10.1016/j.neuroimage.2010.07.033.

Forstmann BU, de Hollander G, van Maanen L, Alkemade A, Keuken MC. Towards a mechanistic understanding of the human subcortex. Nature Reviews Neuroscience. 2017 Jan; 18(1):57–65. http://www.nature.com/articles/nrn.2016.163, doi: 10.1038/nrn.2016.163.

Forstmann BU, Isaacs BR, Temel Y. Ultra High Field MRI-Guided Deep Brain Stimulation. Trends in Biotechnology. 2017 Oct; 35(10):904–907. https://linkinghub.elsevier.com/retrieve/pii/S0167779917301592, doi: 10.1016/j.tibtech.2017.06.010.

Forstmann BU, Keuken MC, Schafer A, Bazin PL, Alkemade A, Turner R. Multi-modal ultra-high resolution structural 7-Tesla MRI data repository. Scientific Data. 2014 Dec; 1(1). http://www.nature.com/articles/sdata201450, doi: 10.1038/sdata.2014.50.

Frazier JA, Chiu S, Breeze JL, Makris N, Lange N, Kennedy DN, Herbert MR, Bent EK, Koneru VK, Dieterich ME, Hodge SM, Rauch SL, Grant PE, Cohen BM, Seidman LJ, Caviness VS, Biederman J. Structural Brain Magnetic Resonance Imaging of Limbic and Thalamic Volumes in Pediatric Bipolar Disorder. American Journal of Psychiatry. 2005 Jul; 162(7):1256–1265. http://psychiatryonline.org/doi/abs/10.1176/appi.ajp.162.7.1256, doi: 10.1176/appi.ajp.162.7.1256.

Garzón B, Sitnikov R, Bäckman L, Kalpouzos G. Automated segmentation of midbrain structures with high iron content. NeuroImage. 2018 Apr; 170:199–209. https://linkinghub.elsevier.com/retrieve/pii/S105381191730486X, doi: 10.1016/j.neuroimage.2017.06.016.

Herting MM, Johnson C, Mills KL, Vijayakumar N, Dennison M, Liu C, Goddings AL, Dahl RE, Sowell ER, Whittle S, Allen NB, Tamnes CK. Development of subcortical volumes across adolescence in males and females: A multisample study of longitudinal changes. NeuroImage. 2018 May; 172:194–205. https://linkinghub.elsevier.com/retrieve/pii/S1053811918300211, doi: 10.1016/j.neuroimage.2018.01.020.

Hill RA, Li AM, Grutzendler J. Lifelong cortical myelin plasticity and age-related degeneration in the live mammalian brain. Nature Neuroscience. 2018 May; 21(5):683–695. http://www.nature.com/articles/s41593-018-0120-6, doi: 10.1038/s41593-018-0120-6.

Ho TC, Gutman B, Pozzi E, Grabe HJ, Hosten N, Wittfeld K, Völzke H, Baune B, Dannlowski U, Förster K, Grotegerd D, Redlich R, Jansen A, Kircher T, Krug A, Meinert S, Nenadic I, Opel N, Dinga R, Veltman DJ, et al. Subcortical shape alterations in major depressive disorder: Findings from the ENIGMA major depressive disorder working group. Human Brain Mapping. 2020 Mar; https://onlinelibrary.wiley.com/doi/10.1002/hbm.24988, doi: 10.1002/hbm.24988.

de Hollander G, Keuken MC, van der Zwaag W, Forstmann BU, Trampel R. Comparing functional MRI protocols for small, iron-rich basal ganglia nuclei such as the subthalamic nucleus at 7 T and 3 T: Optimizing fMRI for Subcortical Nuclei. Human Brain Mapping. 2017 Jun; 38(6):3226–3248. http://doi.wiley.com/10.1002/hbm.23586, doi: 10.1002/hbm.23586.

Huntenburg JM, Steele CJ, Bazin PL. Nighres: processing tools for high-resolution neuroimaging. GigaScience. 2018 Jul; 7(7). https://academic.oup.com/gigascience/article/doi/10.1093/gigascience/giy082/5049008, doi: 10.1093/gigascience/giy082.

Iglesias JE, Insausti R, Lerma-Usabiaga G, Bocchetta M, Van Leemput K, Greve DN, van der Kouwe A, Fischl B, Caballero-Gaudes C, Paz-Alonso PM. A probabilistic atlas of the human thalamic nuclei combining ex vivo MRI and histology. NeuroImage. 2018 Dec; 183:314–326. https://linkinghub.elsevier.com/retrieve/pii/S1053811918307109, doi: 10.1016/j.neuroimage.2018.08.012.

Iglesias JE, Sabuncu MR, Van Leemput K. A unified framework for cross-modality multi-atlas segmentation of brain MRI. Medical Image Analysis. 2013 Dec; 17(8):1181–1191. https://linkinghub.elsevier.com/retrieve/pii/S136184151300114X, doi: 10.1016/j.media.2013.08.001.

Ji JL, Spronk M, Kulkarni K, Repovš G, Anticevic A, Cole MW. Mapping the human brain's cortical-subcortical functional network organization. NeuroImage. 2019 Jan; 185:35–57. https://linkinghub.elsevier.com/retrieve/pii/S1053811918319657, doi: 10.1016/j.neuroimage.2018.10.006.

Johansen-Berg H. Human connectomics — What will the future demand? NeuroImage. 2013 Oct; 80:541–544. https://linkinghub.elsevier.com/retrieve/pii/S105381191300582X, doi: 10.1016/j.neuroimage.2013.05.082.

Keuken MC, Bazin PL, Backhouse K, Beekhuizen S, Himmer L, Kandola A, Lafeber JJ, Prochazkova L, Trutti A, Schäfer A, Turner R, Forstmann BU. Effects of aging on T_1, $T_2{}^*$, and QSM MRI values in the subcortex. Brain Structure and Function. 2017 Aug; 222(6):2487–2505. http://link.springer.com/10.1007/s00429-016-1352-4, doi: 10.1007/s00429-016-1352-4.

Keuken MC, Bazin PL, Schafer A, Neumann J, Turner R, Forstmann BU. Ultra-High 7T MRI of Structural Age-Related Changes of the Subthalamic Nucleus. Journal of Neuroscience. 2013 Mar; 33(11):4896–4900. http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.3241-12.2013, doi: 10.1523/JNEUROSCI.3241-12.2013.

Koshiyama D, Fukunaga M, Okada N, Yamashita F, Yamamori H, Yasuda Y, Fujimoto M, Ohi K, Fujino H, Watanabe Y, Kasai K, Hashimoto R. Role of subcortical structures on cognitive and social function in schizophrenia. Scientific Reports. 2018 Dec; 8(1). http://www.nature.com/articles/s41598-017-18950-2, doi: 10.1038/s41598-017-18950-2.

Li W, Avram AV, Wu B, Xiao X, Liu C. Integrated Laplacian-based phase unwrapping and background phase removal for quantitative susceptibility mapping: HARPERELLA PHASE PROCESSING. NMR in Biomedicine. 2014 Feb; 27(2):219–227. http://doi.wiley.com/10.1002/nbm.3056, doi: 10.1002/nbm.3056.

Makowski C, Béland S, Kostopoulos P, Bhagwat N, Devenyi GA, Malla AK, Joober R, Lepage M, Chakravarty MM. Evaluating accuracy of striatal, pallidal, and thalamic segmentation methods: Comparing automated approaches to manual delineation. NeuroImage. 2018 Apr; 170:182–198. https://linkinghub.elsevier.com/retrieve/pii/S1053811917301805, doi: 10.1016/j.neuroimage.2017.02.069.

Maquer G, Musy SN, Wandel J, Gross T, Zysset PK. Bone Volume Fraction and Fabric Anisotropy Are Better Determinants of Trabecular Bone Stiffness Than Other Morphological Variables: THE BEST DETERMINANTS OF TRABECULAR BONE STIFFNESS. Journal of Bone and Mineral Research. 2015 Jun; 30(6):1000–1008. http://doi.wiley.com/10.1002/jbmr.2437, doi: 10.1002/jbmr.2437.

Marquand AF, Haak KV, Beckmann CF. Functional corticostriatal connection topographies predict goal-directed behaviour in humans. Nature Human Behaviour. 2017 Aug; 1(8). http://www.nature.com/articles/s41562-017-0146, doi: 10.1038/s41562-017-0146.

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 T1-mapping at high field. NeuroImage. 2010 Jan; 49(2):1271–1281. https://linkinghub.elsevier.com/retrieve/pii/S1053811909010738, doi: 10.1016/j.neuroimage.2009.10.002.

McCoy RC, Wakefield J, Akey JM. Impacts of Neanderthal-Introgressed Sequences on the Landscape of Human Gene Expression. Cell. 2017 Feb; 168(5):916–927.e12. https://linkinghub.elsevier.com/retrieve/pii/S0092867417301289, doi: 10.1016/j.cell.2017.01.038.

Miletić S, Bazin PL, Weiskopf N, van der Zwaag W, Forstmann BU, Trampel R. fMRI protocol optimization for simultaneously studying small subcortical and cortical areas at 7T. NeuroImage. 2020 Oct; 219:116992. https://linkinghub.elsevier.com/retrieve/pii/S105381192030478X, doi: 10.1016/j.neuroimage.2020.116992.

Mosley PE, Smith D, Coyne T, Silburn P, Breakspear M, Perry A. The site of stimulation moderates neuropsychiatric symptoms after subthalamic deep brain stimulation for Parkinson's disease. NeuroImage: Clinical. 2018; 18:996–1006. https://linkinghub.elsevier.com/retrieve/pii/S2213158218300792, doi: 10.1016/j.nicl.2018.03.009.

Patenaude B, Smith SM, Kennedy DN, Jenkinson M. A Bayesian model of shape and appearance for subcortical brain segmentation. NeuroImage. 2011 Jun; 56(3):907–922. https://linkinghub.elsevier.com/retrieve/pii/S1053811911002023, doi: 10.1016/j.neuroimage.2011.02.046.

Pauli WM, Nili AN, Tyszka JM. A high-resolution probabilistic in vivo atlas of human subcortical brain nuclei. Scientific Data. 2018 Dec; 5(1). http://www.nature.com/articles/sdata201863, doi: 10.1038/sdata.2018.63.

Qiu A, Brown T, Fischl B, Ma J, Miller MI. Atlas Generation for Subcortical and Ventricular Structures With Its Applications in Shape Analysis. IEEE Transactions on Image Processing. 2010 Jun; 19(6):1539–1547. http://ieeexplore.ieee.org/document/5404310/, doi: 10.1109/TIP.2010.2042099.

Sitek KR, Gulban OF, Calabrese E, Johnson GA, Lage-Castellanos A, Moerel M, Ghosh SS, De Martino F. Mapping the human subcortical auditory system using histology, post mortem MRI and in vivo MRI at 7T. bioRxiv. 2019 Jun; http://biorxiv.org/lookup/doi/10.1101/568139, doi: 10.1101/568139.

Smith SM. Fast robust automated brain extraction. Human Brain Mapping. 2002 Nov; 17(3):143–155. http://doi.wiley.com/10.1002/hbm.10062, doi: 10.1002/hbm.10062.

Steiner H, Tseng KY, editors. Handbook of Basal Ganglia Structure and Function, Second Edition, vol. 24 of Handbook of Behavioral Neuroscience. Elsevier; 2017. https://linkinghub.elsevier.com/retrieve/pii/C20140023593, doi: 10.1016/C2014-0-02359-3.

Tardif CL, Gauthier CJ, Steele CJ, Bazin PL, Schäfer A, Schaefer A, Turner R, Villringer A. Advanced MRI techniques to improve our understanding of experience-induced neuroplasticity. NeuroImage. 2016 May; 131:55–72. https://linkinghub.elsevier.com/retrieve/pii/S1053811915007661, doi: 10.1016/j.neuroimage.2015.08.047.

Turner R. Myelin and Modeling: Bootstrapping Cortical Microcircuits. Frontiers in Neural Circuits. 2019 May; 13. https://www.frontiersin.org/article/10.3389/fncir.2019.00034/full, doi: 10.3389/fncir.2019.00034.

Van Essen DC, Smith SM, Barch DM, Behrens TEJ, Yacoub E, Ugurbil K. The WU-Minn Human Connectome Project: An overview. NeuroImage. 2013 Oct; 80:62–79. https://linkinghub.elsevier.com/retrieve/pii/S1053811913005351, doi: 10.1016/j.neuroimage.2013.05.041.

Visser E, Keuken MC, Douaud G, Gaura V, Bachoud-Levi AC, Remy P, Forstmann BU, Jenkinson M. Automatic segmentation of the striatum and globus pallidus using MIST: Multimodal Image Segmentation Tool. NeuroImage. 2016 Jan; 125:479–497. https://linkinghub.elsevier.com/retrieve/pii/S1053811915009143, doi: 10.1016/j.neuroimage.2015.10.013.

Visser E, Keuken MC, Forstmann BU, Jenkinson M. Automated segmentation of the substantia nigra, subthalamic nucleus and red nucleus in 7 T data at young and old age. NeuroImage. 2016 Oct; 139:324–336. https://linkinghub.elsevier.com/retrieve/pii/S1053811916302889, doi: 10.1016/j.neuroimage.2016.06.039.

Wang BT, Poirier S, Guo T, Parrent AG, Peters TM, Khan AR. Generation and evaluation of an ultra-high-field atlas with applications in DBS planning. In: Styner MA, Angelini ED, editors. *Proceedings of SPIE Medical Imaging* San Diego, California, United States; 2016. p. 97840H. http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.2217126, doi: 10.1117/12.2217126.

Weiskopf N, Suckling J, Williams G, Correia MM, Inkster B, Tait R, Ooi C, Bullmore ET, Lutti A. Quantitative multi-parameter mapping of R1, PD*, MT, and R2* at 3T: a multi-center validation. Frontiers in Neuroscience. 2013; 7. http://journal.frontiersin.org/article/10.3389/fnins.2013.00095/abstract, doi: 10.3389/fnins.2013.00095.

Xiao Y, Fonov V, Bériault S, Subaie FA, Chakravarty MM, Sadikot AF, Pike GB, Collins DL. Multi-contrast unbiased MRI atlas of a Parkinson's disease population. International Journal of Computer Assisted Radiology and Surgery. 2015 Mar; 10(3):329–341. http://link.springer.com/10.1007/s11548-014-1068-y, doi: 10.1007/s11548-014-1068-y.

Yeatman JD, Wandell BA, Mezer AA. Lifespan maturation and degeneration of human brain white matter. Nature Communications. 2014 Dec; 5(1). http://www.nature.com/articles/ncomms5932, doi: 10.1038/ncomms5932.

Yelnik J, Bardinet E, Dormont D, Malandain G, Ourselin S, Tandé D, Karachi C, Ayache N, Cornu P, Agid Y. A three-dimensional, histological and deformable atlas of the human basal ganglia. I. Atlas construction based on immunohistochemical and MRI data. NeuroImage. 2007 Jan; 34(2):618–638. https://linkinghub.elsevier.com/retrieve/pii/S1053811906009700, doi: 10.1016/j.neuroimage.2006.09.026.

Zhang Y, Wei H, Cronin MJ, He N, Yan F, Liu C. Longitudinal atlas for normative human brain development and aging over the lifespan using quantitative susceptibility mapping. NeuroImage. 2018 May; 171:176–189. https://linkinghub.elsevier.com/retrieve/pii/S1053811918300089, doi: 10.1016/j.neuroimage.2018.01.008.