Quantitative Magnetic Resonance Imaging for Segmentation and White Matter Extraction of the Hypothalamus

Melanie Spindler^{1*}, Christiane Thiel^{1,2,3}

¹ Biological Psychology, Department of Psychology, School of Medicine and Health Sciences, Carl von Ossietzky Universität Oldenburg, Oldenburg, Germany

² Cluster of Excellence “Hearing4all”, Carl von Ossietzky Universität Oldenburg, Oldenburg, Germany

³ Research Centre Neurosensory Science, Carl von Ossietzky Universität Oldenburg, Oldenburg, Germany

* Correspondence:

Melanie Spindler

University of Oldenburg

melanie.spindler@uol.de

Abstract

Since the hypothalamus is involved in many neuroendocrine, metabolic, and affective disorders, detailed hypothalamic imaging has become of major interest to better characterize disease-induced tissue damages and abnormalities. Still, image contrast of conventional anatomical magnetic resonance imaging lacks morphological detail, thus complicating complete and precise segmentation of the hypothalamus. The hypothalamus' position lateral to the third ventricle and close proximity to white matter tracts including the optic tract, fornix, and mammillothalamic tract display one of the remaining shortcomings of hypothalamic segmentation, as reliable exclusion of white matter is not yet possible. Recent studies found that quantitative magnetic resonance imaging (qMRI), a method to create maps of different standardized tissue contents, improved segmentation of cortical and subcortical brain regions. So far, this has not been tested for the hypothalamus. Therefore, in this study, we investigated the usability of qMRI and diffusion MRI for the purpose of detailed manual segmentation and data-driven parcellation of the hypothalamus and compared our results to recent state-of-the-art segmentations. Our results show that qMRI and diffusion parameters indeed differ between hypothalamic subunits, and that qMRI is helpful for hypothalamic segmentation. In addition, we provide a data-driven clustering algorithm to reliably exclude white matter from hypothalamic tissue. We propose that qMRI poses a useful addition to detailed hypothalamic segmentation and volumetry.

Introduction

The human hypothalamus is a small diencephalic structure located to the left and right of the third ventricle, above the pituitary gland and below the thalamus. It consists of approximately 15 distinct nuclei (Baroncini et al., 2012), which are, to a large extent, functionally separable and involved in a variety of metabolic, endocrine, and psychiatric diseases (e.g., mood disorders, obesity) (Bao & Swaab, 2019; Saper & Lowell, 2014; Seong et al., 2019). Hence, it has become increasingly important to distinguish different hypothalamic regions in neuroimaging research. In the last years, segmentation of the hypothalamus was mostly based on T_1 - and T_2^* -weighted magnetic resonance images (MRI), using manual or semi-automated techniques and anatomical landmarks identified in histological examinations. However, differences in MRI acquisition strategies can result in variability across studies, and image contrast in the hypothalamic region lacks morphological detail, resulting in inter-rater bias and inaccurate segmentations. Especially lateral boundaries of the hypothalamus are not easily identifiable and often approximated using surrounding anatomical landmarks, such as the optic tract (Goldstein et al., 2007; Lemaire et al., 2011; Wolff et al., 2018). Therefore, and because manual segmentation is very time-consuming, the first automated tool for both hypothalamus segmentation and parcellation was recently developed (Billot et al., 2020). The tool uses a convolutional neural network to form five subunits trained on manually parcellated T_1w images following a previously published protocol for manual segmentation (Bocchetta et al., 2015; Makris et al., 2013). While this enables fast segmentation of large datasets, the subunits generated are largely influenced by the training dataset and spatially constrained to follow the orientations and ratios learned on the training data. As a result, individual size or shape differences, which may be relevant especially in patient samples, could be overlooked. The same applies to a recently published atlas-based segmentation. Neudorfer et al. (2020) developed a high-resolution atlas for hypothalamic nuclei and surrounding structures based on 900 subjects from the Human Connectome Project, thereby enabling detailed hypothalamus parcellation on nucleus level. However, due to the small structure of the hypothalamus, pinpoint registration is another critical aspect that can lead to misrepresentation of the underlying tissue. To better integrate underlying tissue information into segmentation, data-driven clustering methods have been developed, which were based on diffusion-weighted (Schönknecht et al., 2013), or functional imaging data (Osada et al., 2017). Nonetheless, one remaining drawback of available procedures for hypothalamic segmentation and volumetry is the reliable exclusion of cerebrospinal fluid, the fornix, and other white matter. The fornix is a large white matter bundle passing from the hippocampus through the hypothalamus and to the mammillary bodies (Christiansen et al., 2016). In more posterior parts of the hypothalamus, the fornix is fully surrounded by hypothalamic tissue, appearing more diffuse and weaker in contrast on T_1w images, further complicating its reliable and complete extraction. The optic and mammillothalamic tracts are additional white matter bundles in

close proximity to the hypothalamus that need to be taken into consideration, the former bordering the inferior to lateral regions of the hypothalamus, and the latter arising from the superior edge of the mammillary bodies. Especially diffusion parameters such as fractional anisotropy or mean diffusivity, which are frequently used to investigate structure-function relationships in the hypothalamus, could be largely influenced by partial volume contaminations from white matter and CSF. In recent years, MRI protocols sensitive to microstructural properties have been developed, including quantitative MRI (qMRI). Based on multiparameter maps including magnetization transfer saturation (MT sat.), proton density (PD), longitudinal relaxation rate ($R_1: 1/T_1$), and effective transverse relaxation rate ($R_2^*=1/T_1^*$), different tissue contents are computed (Callaghan et al., 2014). MT sat. is sensitive to the amount of myelin and macromolecular content, thereby serving as a marker for fiber integrity and myelination (Hagiwara et al., 2018; Lema et al., 2016). Consequently, MT sat. is higher in white compared to grey matter and could be a useful tool to distinguish hypothalamic microstructure from surrounding white matter. In contrast, PD is sensitive to free water content, and can be used to characterize tissue damage (e.g., edema) and CSF. Finally, R_1 is sensitive to myelin and water, and together with R_2^* reflects iron, which is often used as a marker for ageing and neurodegenerative diseases (Ulla et al., 2013; Ward et al., 2014). Hence, contrarily to conventional T_1w and T_2^*w images, grey values in qMRI provide meaningful information about the underlying tissue microstructure, while achieving comparable spatial resolution. Therefore, qMRI has previously been used for easier identification or segmentation of both cortical and subcortical areas including somatomotor cortex (Carey et al., 2017), brainstem (Lambert et al., 2013), pallidum, putamen, and substantia nigra (Helms et al., 2009). To our knowledge, however, no study has investigated hypothalamic and surrounding tissue microstructure using qMRI.

In the present study we first aimed to evaluate whether qMRI sequences can aid to better delineate the hypothalamus. Second, the sensitivity of qMRI and diffusion MRI to identify possible microstructural differences in hypothalamic subunits is explored. Third, we determined whether quantitative and diffusion MRI measures could be used to automatically extract the fornix and other white matter tracts from the hypothalamus using a data-driven clustering approach to improve available segmentation procedures.

Methods

Participants

12 healthy participants aged between 21 and 35 years without history of neurological or psychiatric disorders were included in this study (7 females, 5 males). One male and female were left-handed, as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971). The remaining subjects were right-handed. The study was approved by the ethics committee of the University of Oldenburg (Drs. 50/2017), and all participants provided written informed consent prior to participation.

MRI Acquisition

Data were collected on a 3T whole-body MRI scanner (Magnetom Prisma, Siemens Medical Systems, Germany) with a 64-channel head coil for signal reception and a body coil for transmission. During scanning, the participants were asked to lay still with their eyes closed and their head was padded to minimize head motion.

T₁-weighted images were acquired using a magnetization-prepared rapid gradient-echo sequence (MP-RAGE) with isotropic voxel size of 0.75 mm³. T₂^{*}-weighted images were obtained using a turbo spin echo sequence with voxel size of 0.6x0.6x2.0 mm and averaged across two acquisitions to increase the signal-to-noise ratio. The field of view (FOV) was placed such that it covered the entire hypothalamic region. Following the protocol of Callaghan et al. (2019) and Tabelow et al. (2019), we used multi-echo 3D FLASH (fast low-angle shot) sequences at 0.8 mm³ resolution with six different echo times for each acquisition for qMRI (T1w, PDw, MT sat.). A SIEMENS product sequence based on spin-echo/stimulated echo (SE/STE) imaging was acquired at the beginning of each session with 4.0x4.0x5.0 mm resolution. Prior to each T1w, PDw, and MTw measurements, two low-resolution sensitivity maps from the head and body coils were acquired (voxel size: 4.0x4.0x5.0 mm).

Multi-shell, multi-slice (iPAT factor 3) diffusion-weighted imaging was acquired in anterior-to-posterior phase encoding direction, with 132 diffusion directions (forty directions each for b₁₀₀₀, b₂₀₀₀, and b₃₀₀₀ with 12 interspersed b₀ images) and 2.0 mm³ voxel size. An additional single b₀ image was obtained in posterior-to-anterior phase encoding direction with otherwise identical parameters. The total scanning time was ~45 minutes, and a detailed description of the imaging parameters is provided in Table 1.

Table 1: MRI measurements in the order in which they were acquired. FOV: Field of view, TR: Repetition time

	Flip angle	FOV (mm ²)	Echo times (ms)	TR (ms)	No. slices
SE/STE	[90, 120, 60, 135, 45]	256x256	[14, 14]	2000	18
T1w	21	256x256	[2.57, 5.82, 9.86, 13.9, 17.94, 21.98]	26	279
PDw	4	256x256	[2.57, 5.82, 9.86, 13.9, 17.94, 21.98]	26	279
MT sat.	6	256x256	[2.57, 5.82, 9.86, 13.9, 17.94, 21.98]	44	279

Sensitivity Head	6	256x256	1.99	4.1	44
Sensitivity Body	6	256x256	1.99	4.1	44
MP-RAGE	9	240x240	2.07	2000	224
T2*	150	224x320	79	5000	40
b₀	-	232x232	75	3900	81
Diffusion	-	232x232	75	3900	81

MRI Processing

First, images were automatically reoriented to AC-PC orientation in mid-sagittal plane (*Auto-Reorient* module), and multiparameter maps were generated (*create hMRI* module) using the hMRI toolbox (Tabelow et al., 2019) implemented in the Statistical Parametric Mapping toolbox (SPM 12) (*Statistical Parametric Mapping*, 2007). B1 transmit field correction was performed based on the SE/STE sequence. Data were corrected for RF sensitivity bias using the head and body sensitivity maps.

The T_{1w} image was segmented into GM, WM, and CSF, and bias corrected (light regularization: 0.001) using *Segment* in SPM. Multiparameter maps were then registered and interpolated to anatomical (T₁) space using the FMRIB Software Library's (FSL, v. 5.0.9) *flirt* with rigid body (6 df) registration (Greve & Fischl, 2009; Smith et al., 2004). Trilinear interpolation was used for reslicing from 0.8 mm³ isotropic to 0.75 mm³ isotropic resolution. For registration, the T_{1w} image was skullstripped in SPM (*Image Calculator*) using the tissue probability maps for GM, WM, and CSF generated in the segmentation step. Diffusion-weighted data were eddy and movement corrected using FSL's *topup* and *eddy* (Andersson & Sotiropoulos, 2016). *Epi_reg* was used to register the diffusion data to T1-space, and diffusion gradients were rotated according to the transformation matrix. Using *qboot*, multishell constant solid angle orientation distribution functions (ODF) were reconstructed with a spherical harmonics basis of 6. Based on mean ODF in each voxel, generalized fractional anisotropy (GFA) was computed as a quantitative descriptor of the diffusion process in each voxel (Tuch, 2004). GFA was chosen instead of tensor FA as a more sensitive measure for crossing fibers and partial volume contaminations.

Hypothalamus Segmentation

Hypothalamus masks were manually defined for the left and right hemisphere using established anatomical landmarks (Goldstein et al., 2007; Schindler et al., 2013) based on the T_{1w} and T₂*w image in 3D Slicer v. 10.0.2 (<https://www.slicer.org/>, (Fedorov et al., 2012)). The fornix was included when surrounded by hypothalamic tissue. Using the tissue probability map for CSF, the masks were refined such that voxels with a CSF probability >10 % were excluded. To implement a standardized procedure employing qMRI, a segmentation protocol for the hypothalamus and white matter was developed by re-evaluation of the hypothalamus masks based on qMRI, and manual segmentation of the fornix

passing through the hypothalamus mask was performed. Following this procedure, an additional second rater performed the hypothalamus and fornix segmentations to determine inter-rater reliability based on Dice coefficients. To that aim, the voxel overlap between the CSF-corrected hypothalamus and fornix masks of both raters was calculated.

Statistical Analysis

Hypothalamic subunits were generated following the CNN architecture proposed by Billot et al. (2020) using the bias-corrected T_1w image (anterior-inferior (a-iHyp), anterior-superior (a-sHyp), inferior tuberal (iTub), superior tuberal (sTub), and posterior (Post) Hypothalamus). Those subunits were then used to independently evaluate tissue microstructure of the hypothalamus based on quantitative parameters. Mean values of MT sat., PD, R_1 , R_2^* , and GFA were compared between subunits using one-way ANOVAs and Tukey's HSD post hoc tests where appropriate. To account for multiple comparisons, $p < .01$ (Bonferroni adjustment) was considered significant for each ANOVA.

For automated white matter extraction in the hypothalamus, voxel values for each qMRI parameter within the hypothalamus masks were first normalized to a mean of zero and standard deviation of one. Then, spectral clustering was computed separately for each hemisphere based on MT sat., PD, R_1 , R_2^* , and GFA. Spectral clustering is a graph-based algorithm that partitions the data into k clusters using the eigenvectors of the normalized Laplacian matrix (Shi & Malik, 2000; von Luxburg, 2007). It was implemented with k -means using *spectralcluster* in MATLAB 2018a (Mathworks Inc.). An advantage of spectral clustering compared to other methods is that underlying clusters are not restricted to convex regions because the algorithm does not make assumptions on the cluster shape. As distance metric, the Euclidean distance was chosen. First, optimal k ($k = \{2, 3, 4, 5, 6\}$) was determined minimizing the Davies Bouldin index (DBI) (Davies & Bouldin, 1979). The DBI describes a ratio between within- and between-cluster distances and is a common method to evaluate clustering performance. Resulting clusters were compared to the manual segmentation of the fornix to integrate a priori knowledge about underlying anatomical structure. Then, the number of clusters leading to the best representation of the fornix structure was chosen. Afterwards, clusters were labeled according to their relative position. Using Dice coefficients, the overlap between the manually segmented fornix and the clustering outcome was calculated.

Finally, white matter extraction in all three automated procedures (Billot et al. (2020), Neudorfer et al. (2020), spectral clustering) was evaluated. To that end, the T_1w image was affine and nonlinearly registered to the Montreal Neurological Institute (MNI) template in 0.5mm^3 resolution provided by Neudorfer et al. (2020). Registration was visually checked on easily identifiable landmarks including the mammillary bodies and anterior commissure. The same transforms were then applied to the MT

sat. image and fornix clusters from this study. Overlap was determined using the Dice coefficient, and 95% Hausdorff distance, which describes the 95th percentile of the surface distance. Compared to the Dice coefficient, the Hausdorff distance is less influenced by registration mismatches by taking into consideration the relative position of the individual shapes. Because white matter including the fornix was discarded following Billot et al. (2020), it was not possible to determine the overlap with our results, or the atlas-based segmentation by Neudorfer et al. (2020). Therefore, results obtained with Billot et al. (2020) were only visually evaluated. If not stated otherwise, statistical analysis was performed with R v. 3.6.3 (<https://www.R-project.org/>).

Results

Hypothalamus segmentation and microstructure

Manually segmented hypothalamus masks were overlaid onto the MT saturation, PD, R_1 , and R_2^* maps to explore the possibility of using these images as aid in segmentation. In all participants, MT saturation and R_1 images provided a better visual contrast to laterally surrounding structures as well as white matter structures such as the optic tract, mammillothalamic tract, and fornix, compared to conventional T_1 w and T_2^* w images (Figure 1). Based on these findings, a segmentation protocol was developed employing MT sat. and R_1 images (Table 2). Hypothalamus and fornix masks created following this protocol were compared between two raters using Dice coefficients as a measure of

inter-rater reliability. Mean hypothalamus volume was 789 ± 107 mm³ (rater 1) and 782 ± 82 mm³ (rater 2). Mean Dice coefficients suggest high correspondence between the hypothalamus (0.90 ± 0.01) and the fornix (0.76 ± 0.06) of both raters. For further analyses, segmentation results obtained by the first rater were used.

Figure 1: The left (blue) and right (red) masks of the hypothalamus of one participant overlaid onto the T_1 - and T_2^* -weighted images as well as onto the MT sat., PD, R_1 , and R_2^* maps. Image intensities for the T_1 - and T_2^* -weighted images are displayed in arbitrary units (a.u.), MT sat. and PD in percent units (p.u.), and R_1 and R_2^* in 1/second (s⁻¹).

Table 2: Segmentation protocol for the hypothalamus and surrounding white matter based on qMRI. For segmentation, all images are shifted to AC-PC orientation in the midsagittal plane. Segmentations were performed in 3D Slicer v. 10.0.2, with the R1 image loaded into the foreground and MT sat. image into the background. Visibility of fore- and background was toggled as specified in “Modality”.

Modality	Segmentation
(1) MT sat. coronal (0-5 p.u.)	Begin on the first coronal slice where the anterior commissure (AC) appears continuous. Draw superior edge of the hypothalamus below AC, and inferior edge above optic chiasm (OC). When the optic tracts bifurcate, the hypothalamus is segmented above the infundibular stalk and medial to the optic tracts. Continue until interventricular foramen and fornix.
(2) R ₁ coronal (0.2-1.6 s ⁻¹)	Start again on the first coronal slice where the AC appears continuous and draw the lateral extent of the hypothalamus on both sides of the third ventricle. Medially, the left and right masks cover the third ventricle up to the midline. Check exclusion of olfactory tubercule. Continue until interventricular foramen and fornix.
(3) MT sat. coronal (0-5 p.u.)	Check for exclusion of optic or AC fibers. On the first slice with the fornix visible, draw until below the inferior tip of the fornix. When the fornix is laterally surrounded by hypothalamic tissue, trace around it (exclude voxels of approx. > 2.5 p.u.). Continue until mammillary bodies.
(4) R ₁ coronal (0.2-1.6 s ⁻¹)	Draw lateral edges of the hypothalamus until the appearance of the mammillary bodies (MBs). Check for boundaries: hypothalamic sulcus and subthalamic nucleus (superior), substantia innominata (lateral). Inferiorly, draw until infundibular stalk. Medially, the masks touch at the midline.
(5) MT sat. coronal (0-5 p.u.)	At the first slice with MBs visible, the superior, inferior, and lateral boundary of the mask is defined as the extent of the MBs. Exclude the mammillothalamic tract (visible on the superior edge of the MBs. Check for and exclude substantia nigra (lateral). MBs touch at the midline.
Review in triplanar view	
Coronal	R ₁ : Lateral extent of the hypothalamus MT sat.: Optic tract excluded?
Sagittal	R ₁ : Correct delineation of the hypothalamic sulcus? Check posterior extent of MBs. Excluded infundibular stalk? MT sat.: mammillothalamic tract and fornix excluded?
Axial	R ₁ : Check lateral extent of the inferior hypothalamus and the extent of the MBs
CSF Correction	
Tissue Probability Map CSF	Exclude voxels > 0.10 to correct for fluid-filled spaces, e.g. the third ventricle

In the second step, we determined whether tissue microstructure differs between subregions of the hypothalamus. We analysed mean MT sat., PD, R₁, R₂^{*}, and GFA were in the subunits identified with the algorithm of Billot et al. (2020) using five one-way ANOVAs (Bonferroni-corrected $p < .01$ is considered significant). Results suggest significant subunit differences in MT sat. ($F(4,55)=33, p<.001$), R₁ ($F(4,55)=14.46, p<.001$), R₂^{*} ($F(4,55)=7.154, <.001$), and GFA ($F(4,55)=10.3, p<.001$), and no

differences in PD ($F(4,55)=2.611$, $p=.0452$). Tukey's HSD post hoc test results are depicted in Figure 2 and suggest that most subunits differ in mean MT sat., followed by R_1 .

Figure 2: Boxplot showing microstructural properties of five hypothalamic subunits generated following Billot et al. (2020). Tukey's HSD post hoc tests results: *** $p<.001$, ** $p<.01$, * $p<.05$. A-iHyp: anterior-inferior hypothalamus, a-sHyp: anterior-superior hypothalamus, iTub: inferior tuberal hypothalamus, sTub: superior tuberal hypothalamus, Post: posterior hypothalamus.

White matter extraction and comparison to existing approaches

Our final analysis aimed to automatically detect white matter within the hypothalamus masks and to draw comparisons with regard to existing automated segmentation protocols. To that end, spectral clustering based on qMRI and GFA was performed. First, optimal k was determined using the DBI for each participant and hemisphere. DBI ranged between 2 and 6 (2: 58 %, 3: 25 %, 4: 8.33 %, 5: 4.17 %, 6: 4.17 % of cases), whereby $k=4$ yielded the highest agreement with manual fornix segmentations, as determined by the mean Dice coefficients: For $k=2$, the mean Dice coefficient was 0.35 and 0.42 for the left and right fornix, respectively, For $k=3$, the mean Dice coefficient was 0.60 (left) and 0.73 (right), and 0.81 (left) and 0.82 (right) for $k=4$. Hence, clustering was implemented with four clusters. For all $n=12$ participants, four unique subunits were identified, which were labelled as superior-lateral, inferior-medial, fornix, and posterior/mammillary based on their position (Figure 3). The cluster labelled as fornix also included white matter of the mammillothalamic tract located at the superior

base of the mammillary bodies (Figure 3 bottom row, Figure 4 right), and for one participant white matter belonging to the anterior commissure that was not completely excluded (not shown).

Figure 3: Spectral clustering result in coronal slices exemplary for one participant from anterior to posterior beginning at the top left from left to right and top to bottom.

Spectral clustering outcome on white matter extraction was then compared to those of established procedures (Billot et al., 2020; Neudorfer et al., 2020) and to our manual segmentation. Visual inspection of results obtained from Billot et al. determined that “holes” in the hypothalamic subunits are mostly consistent with the position of the fornix as determined with our automated approach, but that there is residual white matter especially in the superior-tuberal (fornix, Figure 4), and posterior subunits (ends of the mammillothalamic tract).

Next, we compared our manual and clustering results to the corresponding atlas-based masks of the fornix and mammillothalamic tract in Neudorfer et al. (2020) according to Dice coefficients and 95% Hausdorff distances using two one-way ANOVAs and Tukey’s HSD post-tests. Results suggest that Dice coefficients differed ($F(2,33)=115.1$, $p<.001$), with manual and spectral clustering sharing the largest overlap compared to manual and atlas ($p<.001$) and spectral clustering and atlas ($p<.001$) overlap. There was no difference between manual-atlas and spectral clustering-atlas Dice coefficients ($p=.960$). Hausdorff distances were significantly different between approaches ($F(2,33)=4.056$, $p=.027$). Here, however, only manual-spectral clustering Hausdorff distance was smaller compared to manual-atlas Hausdorff distance ($p=.025$). Hausdorff distances between manual-clustering and clustering-atlas were similar ($p=.724$) as well as between manual-atlas and clustering-atlas ($p=.131$) (Table 3). This suggests that while Dice coefficients were considerably lower for the atlas-based segmentation, the Hausdorff distances were comparable to our clustering results.

Table 3: Mean and standard deviation of Dice coefficients and Hausdorff distances for each comparison between white matter extraction procedures. Manual segmentation was compared to spectral clustering (Manual-Clustering), and to atlas-based extraction following Neudorfer et al. (Manual-Atlas). Additionally, atlas-based extraction was compared to spectral clustering results (Clustering-Atlas).

	Manual-Clustering	Manual-Atlas	Clustering-Atlas
Dice Coefficient	0.82/0.08	0.45/0.06	0.44/0.06
95% Hausdorff Distance	1.65/0.54	2.26/0.49	1.82/0.59

Figure 4: Evaluation of automated white matter extraction with spectral clustering. The Magnetization Transfer Saturation image of three consecutive sagittal slices in one subject overlayed with outcomes of three automated procedures. Top: the cluster of the fornix obtained by our spectral clustering approach. Middle: Fornix obtained from the high-resolution atlas by Neudorfer et al. (2020). Bottom: Subunits generated with the automated clustering approach of Billot et al. (2020). Note that large parts of the superior tuberal hypothalamus consist of residual white matter. A-iHyp: anterior-inferior hypothalamus, a-sHyp: anterior-superior hypothalamus, iTub: inferior tuberal hypothalamus, sTub: superior tuberal hypothalamus, Post: posterior hypothalamus.

Discussion

To our knowledge, this is the first study investigating quantitative MRI in the hypothalamus. We aimed to investigate whether knowledge from qMRI can aid in hypothalamic segmentation and improve segmentation accuracy especially with regards to surrounding white matter tracts (e.g., fornix). In addition, we aimed to develop an automated procedure for exclusion of white matter and compared our results to state-of-the-art automated approaches. We proposed that hypothalamic segmentation including MT saturation and R_1 maps can yield more accurate and reproducible results especially in lateral regions of the hypothalamus and surrounding white matter structures than T_1w and T_2^*w images. Contrast in conventional structural imaging is generally low in the hypothalamic region and generated as a mixture of different parameters, thereby representing arbitrary gray values. By using standardized maps, image contrast can be better controlled and more easily reproduced across different sites (Gracien et al., 2020; Weiskopf et al., 2013). Intensity standardization methods for T_1 -weighted images are available, however, comparability between different sequences and standardization approaches is still limited. Because of the weak visibility of lateral hypothalamic boundaries, the size of the hypothalamus in neuroimaging studies is often underestimated due to conservative anatomical landmarks. This way, marginal areas of the hypothalamus are excluded, which is especially critical when examining the lateral hypothalamic area. This leads to a large range in reported hypothalamic volume (approximately 0.6-3.6 cm³), which further reduces comparability across studies (Schindler et al., 2017). As T_1w images do not provide sufficient contrast to completely delineate the fornix, we suggest the usage of MT sat. and R_1 imaging for hypothalamus segmentation where possible. To that aim, we implemented a standardized segmentation procedure that takes advantage of the higher contrast obtained. We showed that underlying microstructure differs between hypothalamic subunits following the automated clustering approach by Billot et al. (2020), suggesting that qMRI could indeed be a useful tool to characterize hypothalamic subregions and tissue abnormalities. Still, the subunits analyzed are likely to be influenced by residual white matter from the fornix and mammillothalamic tract, which could explain the higher MT sat. and GFA in the superior tuberal and posterior subunits compared to the other subunits. We therefore propose implementation of manual or automated post-processing of obtained subunits via qMRI and CSF probability maps for detailed hypothalamic analyses. As the subunits generated are also largely influenced by the training dataset and spatially constrained to follow the orientations and ratios learned on the training data, individual size or shape differences could be overlooked, and fornical fibers as well as fluid-filled spaces are not reliably excluded in the images used. Additionally, it is unclear which voxels are most influential for the formation of the subunits/labeling of each voxel, and further research should be done to determine the underlying problem solving strategies of the network (Lapuschkin et al., 2019). Nevertheless, Billot et al. (2020) present a useful, computationally inexpensive, and time efficient

opportunity for hypothalamic parcellation especially for large datasets where manual techniques would be unfeasible. To integrate knowledge from qMRI into the CNN architecture, the training dataset could possibly be extended to also cover MT sat. and R_1 information to make this approach more sensitive to different tissue types.

A spectral clustering algorithm was used to allow for automated extraction of white matter from hypothalamic tissue based on underlying tissue properties. Using this approach, we were able to delineate the mamillothalamic tract and fornix in all participants. When compared to manual segmentation and a hypothalamic atlas (Neudorfer et al., 2020), our results showed a high level of agreement with manual segmentations. Additionally, Dice coefficients between our clustering approach and the atlas were rather low, however, Hausdorff distances were similar – a mismatch that could be explained by possible registration-related inaccuracies and the individual shapes of the hypothalamic region and fornix that are not grasped by the atlas. These results suggest that while there is only moderate spatial overlap, the shape of the fornix is accurately traced.

Due to the proximity of the hypothalamus to the third ventricle, exclusion of voxels with a high probability of CSF is of particular relevance to control partial volume effects. In existing manual and automated procedures, this is often overlooked. Here, we excluded voxels with a CSF probability $>10\%$, but this also implies that the hypothalamus gets comparatively small and that nuclei close to the third ventricle, such as the periventricular nucleus, will only be partially included. The proximity to the third ventricle could also be problematic in the analysis of diffusion parameters, which are usually based on a tensor function that can be largely influenced by fluid. We therefore decided to use GFA as a possibly more sensitive measure. Nevertheless, clustering did not always yield fully spatially separable clusters, which might be due to the chosen algorithm, as spectral clustering does not take the spatial position of the voxels into account. Another reason could be related to the number of clusters used in this study. For most of the cases, the Davies-Bouldin index suggested an optimal number of two clusters, but results did not achieve anatomically reasonable clusters, and we argue that underlying differences in hypothalamic tissue might be subtle and prone to noise. Hence, further data-driven automated procedures for hypothalamic parcellation should be explored to achieve the most sensitive outcome.

As qMRI reflects different characteristics of tissue microarchitecture, another application of MT sat., PD, R_1 , and R_2^* in the hypothalamus could be related to physiological or psychological outcomes, (e.g., stress, psychiatric disorders, obesity) to deepen the understanding of hypothalamic plasticity in health and disease. Diffusion parameters are commonly used to investigate structure-function relationships (Chen et al., 2015; Clark et al., 2011; Poletti et al., 2020; Thomas et al., 2019), but are sensitive to changes of the diffusion process due to CSF and white matter. Still, other quantitative MRI parameters

have only been integrated to a very limited extent. For example, recently Kullmann and colleagues investigated the association of water content in the brain and obesity and found associations of higher body-mass index with increased water content in the hypothalamus, possibly related to inflammatory processes (Kullmann et al., 2020). In our study of healthy participants, microstructural properties in hypothalamic subregions differed. It is therefore conceivable that in disease, only parts of the hypothalamic area are affected, and its microstructure changed, or that specific areas show an increase or decrease in volume when affected. Therefore, qMRI could pose a useful tool in investigating structure-function relationships of the hypothalamus.

In summary, we propose that qMRI can improve hypothalamic segmentation and parcellation and possibly enhance inter-study comparability and interpretability with particular focus on the fornix and other surrounding white matter that, to date, display one of the biggest weaknesses in detailed hypothalamic analyses.

References

- Andersson, J. L. R., & Sotiropoulos, S. N. (2016). An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *NeuroImage*, 125, 1063–1078. <https://doi.org/10.1016/j.neuroimage.2015.10.019>
- Bao, A.-M., & Swaab, D. F. (2019). The human hypothalamus in mood disorders: The HPA axis in the center. *IBRO Reports*, 6, 45–53. <https://doi.org/10.1016/j.ibror.2018.11.008>
- Baroncini, M., Jissendi, P., Balland, E., Besson, P., Pruvo, J.-P., Francke, J.-P., Dewailly, D., Blond, S., & Prevot, V. (2012). MRI atlas of the human hypothalamus. *NeuroImage*, 59(1), 168–180. <https://doi.org/10.1016/j.neuroimage.2011.07.013>
- Billot, B., Bocchetta, M., Todd, E., Dalca, A. V., Rohrer, J. D., & Iglesias, J. E. (2020). Automated segmentation of the Hypothalamus and associated subunits in brain MRI. *NeuroImage*, 117287. <https://doi.org/10.1016/j.neuroimage.2020.117287>
- Bocchetta, M., Gordon, E., Manning, E., Barnes, J., Cash, D. M., Espak, M., Thomas, D. L., Modat, M., Rossor, M. N., Warren, J. D., Ourselin, S., Frisoni, G. B., & Rohrer, J. D. (2015). Detailed volumetric analysis of the hypothalamus in behavioral variant frontotemporal dementia. *Journal of Neurology*, 262(12), 2635–2642. <https://doi.org/10.1007/s00415-015-7885-2>
- Callaghan, M. F., Freund, P., Draganski, B., Anderson, E., Cappelletti, M., Chowdhury, R., Diedrichsen, J., FitzGerald, T. H. B., Smittenaar, P., Helms, G., Lutti, A., & Weiskopf, N. (2014). Widespread age-related differences in the human brain microstructure revealed by quantitative magnetic resonance imaging. *Neurobiology of Aging*, 35(8), 1862–1872. <https://doi.org/10.1016/j.neurobiolaging.2014.02.008>
- Callaghan, M. F., Lutti, A., Ashburner, J., Balteau, E., Corbin, N., Draganski, B., Helms, G., Kherif, F., Leutritz, T., Mohammadi, S., Phillips, C., Reimer, E., Ruthotto, L., Seif, M., Tabelow, K., Ziegler, G., & Weiskopf, N. (2019). Example dataset for the hMRI toolbox. *Data in Brief*, 25, 104132. <https://doi.org/10.1016/j.dib.2019.104132>

376 Carey, D., Krishnan, S., Callaghan, M. F., Sereno, M. I., & Dick, F. (2017). Functional and Quantitative
377 MRI Mapping of Somatomotor Representations of Human Supralaryngeal Vocal Tract.
378 *Cerebral Cortex (New York, NY)*, 27(1), 265–278. <https://doi.org/10.1093/cercor/bhw393>

379 Chen, D. Q., Strauss, I., Hayes, D. J., Davis, K. D., & Hodaie, M. (2015). Age-Related Changes in
380 Diffusion Tensor Imaging Metrics of Fornix Subregions in Healthy Humans. *Stereotactic and*
381 *Functional Neurosurgery*, 93(3), 151–159. <https://doi.org/10.1159/000368442>

382 Christiansen, K., Metzler-Baddeley, C., Parker, G. D., Muhlert, N., Jones, D. K., Aggleton, J. P., & Vann,
383 S. D. (2016). Topographic separation of fornical fibers associated with the anterior and
384 posterior hippocampus in the human brain: An MRI-diffusion study. *Brain and Behavior*, 7(1).
385 <https://doi.org/10.1002/brb3.604>

386 Clark, K. A., Nuechterlein, K. H., Asarnow, R. F., Hamilton, L. S., Phillips, O. R., Hageman, N. S., Woods,
387 R. P., Alger, J. R., Toga, A. W., & Narr, K. L. (2011). Mean diffusivity and fractional anisotropy
388 as indicators of disease and genetic liability to schizophrenia. *Journal of psychiatric research*,
389 45(7), 980–988. <https://doi.org/10.1016/j.jpsychires.2011.01.006>

390 Davies, D. L., & Bouldin, D. W. (1979). A Cluster Separation Measure. *IEEE Transactions on Pattern*
391 *Analysis and Machine Intelligence*, PAMI-1(2), 224–227.
392 <https://doi.org/10.1109/TPAMI.1979.4766909>

393 Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.-C., Pujol, S., Bauer, C., Jennings,
394 D., Fennessy, F., Sonka, M., Buatti, J., Aylward, S., Miller, J. V., Pieper, S., & Kikinis, R. (2012).
395 3D Slicer as an Image Computing Platform for the Quantitative Imaging Network. *Magnetic*
396 *resonance imaging*, 30(9), 1323–1341. <https://doi.org/10.1016/j.mri.2012.05.001>

397 Goldstein, J. M., Seidman, L. J., Makris, N., Ahern, T., O'Brien, L. M., Caviness, V. S., Kennedy, D. N.,
398 Faraone, S. V., & Tsuang, M. T. (2007). Hypothalamic Abnormalities in Schizophrenia: Sex
399 Effects and Genetic Vulnerability. *Biological Psychiatry*, 61(8), 935–945.
400 <https://doi.org/10.1016/j.biopsych.2006.06.027>

401 Gracien, R.-M., Maiworm, M., Brüche, N., Shrestha, M., Nöth, U., Hattingen, E., Wagner, M., &
 402 Deichmann, R. (2020). How stable is quantitative MRI? – Assessment of intra- and inter-
 403 scanner-model reproducibility using identical acquisition sequences and data analysis
 404 programs. *NeuroImage*, 207, 116364. <https://doi.org/10.1016/j.neuroimage.2019.116364>
 405 Greve, D. N., & Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based
 406 registration. *NeuroImage*, 48(1), 63–72. <https://doi.org/10.1016/j.neuroimage.2009.06.060>
 407 Hagiwara, A., Hori, M., Kamagata, K., Warntjes, M., Matsuyoshi, D., Nakazawa, M., Ueda, R., Andica,
 408 C., Koshino, S., Maekawa, T., Irie, R., Takamura, T., Kumamaru, K. K., Abe, O., & Aoki, S.
 409 (2018). Myelin Measurement: Comparison Between Simultaneous Tissue Relaxometry,
 410 Magnetization Transfer Saturation Index, and T 1 w/T 2 w Ratio Methods. *Scientific Reports*,
 411 8(1), 10554. <https://doi.org/10.1038/s41598-018-28852-6>
 412 Helms, G., Draganski, B., Frackowiak, R., Ashburner, J., & Weiskopf, N. (2009). Improved
 413 segmentation of deep brain grey matter structures using magnetization transfer (MT)
 414 parameter maps. *Neuroimage*, 47(1), 194–198.
 415 <https://doi.org/10.1016/j.neuroimage.2009.03.053>
 416 Kullmann, S., Abbas, Z., Machann, J., Shah, N. J., Scheffler, K., Birkenfeld, A. L., Häring, H.-U., Fritsche,
 417 A., Heni, M., & Preissl, H. (2020). Investigating obesity-associated brain inflammation using
 418 quantitative water content mapping. *Journal of Neuroendocrinology*, n/a(n/a), e12907.
 419 <https://doi.org/10.1111/jne.12907>
 420 Lambert, C., Lutti, A., Helms, G., Frackowiak, R., & Ashburner, J. (2013). Multiparametric brainstem
 421 segmentation using a modified multivariate mixture of Gaussians. *NeuroImage: Clinical*, 2,
 422 684–694. <https://doi.org/10.1016/j.nicl.2013.04.017>
 423 Lapuschkin, S., Wäldchen, S., Binder, A., Montavon, G., Samek, W., & Müller, K.-R. (2019). Unmasking
 424 Clever Hans predictors and assessing what machines really learn. *Nature Communications*,
 425 10(1), 1096. <https://doi.org/10.1038/s41467-019-08987-4>

426 Lema, A., Bishop, C., Malik, O., Mattoscio, M., Ali, R., Nicholas, R., Muraro, P. A., Matthews, P. M.,
 427 Waldman, A. D., & Newbould, R. D. (2016). A compararison of magnetization transfer
 428 methods to assess brain and cervical cord microstructure in multiple sclerosis. 226.
 429 <https://doi.org/10.1111/jon.12377>

430 Lemaire, J.-J., Frew, A. J., McArthur, D., Gorgulho, A. A., Alger, J. R., Salomon, N., Chen, C., Behnke, E.
 431 J., & De Salles, A. A. F. (2011). White matter connectivity of human hypothalamus. *Brain*
 432 *Research*, 1371, 43–64. <https://doi.org/10.1016/j.brainres.2010.11.072>

433 Makris, N., Swaab, D. F., van der Kouwe, A., Abbs, B., Boriel, D., Handa, R. J., Tobet, S., & Goldstein, J.
 434 M. (2013). Volumetric parcellation methodology of the human hypothalamus in
 435 neuroimaging: Normative data and sex differences. *NeuroImage*, 69, 1–10.
 436 <https://doi.org/10.1016/j.neuroimage.2012.12.008>

437 Neudorfer, C., Germann, J., Elias, G. J. B., Gramer, R., Boutet, A., & Lozano, A. M. (2020). A high-
 438 resolution in vivo magnetic resonance imaging atlas of the human hypothalamic region.
 439 *Scientific Data*, 7(1), 305. <https://doi.org/10.1038/s41597-020-00644-6>

440 Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory.
 441 *Neuropsychologia*, 9(1), 97–113. [https://doi.org/10.1016/0028-3932\(71\)90067-4](https://doi.org/10.1016/0028-3932(71)90067-4)

442 Osada, T., Suzuki, R., Ogawa, A., Tanaka, M., Hori, M., Aoki, S., Tamura, Y., Watada, H., Kawamori, R.,
 443 & Konishi, S. (2017). Functional subdivisions of the hypothalamus using areal parcellation and
 444 their signal changes related to glucose metabolism. *NeuroImage*, 162, 1–12.
 445 <https://doi.org/10.1016/j.neuroimage.2017.08.056>

446 Poletti, S., Melloni, E., Mazza, E., Vai, B., & Benedetti, F. (2020). Gender-specific differences in white
 447 matter microstructure in healthy adults exposed to mild stress. *Stress*, 23(1), 116–124.
 448 <https://doi.org/10.1080/10253890.2019.1657823>

449 Saper, C. B., & Lowell, B. B. (2014). The hypothalamus. *Current Biology*, 24(23), R1111–R1116.
 450 <https://doi.org/10.1016/j.cub.2014.10.023>

451 Schindler, S., Schönknecht, P., Schmidt, L., Anwander, A., Strauß, M., Trampel, R., Bazin, P.-L., Möller,
 452 H. E., Hegerl, U., Turner, R., & Geyer, S. (2013). Development and Evaluation of an Algorithm
 453 for the Computer-Assisted Segmentation of the Human Hypothalamus on 7-Tesla Magnetic
 454 Resonance Images. *PLoS ONE*, 8(7), e66394. <https://doi.org/10.1371/journal.pone.0066394>
 455 Schindler, S., Schreiber, J., Bazin, P.-L., Trampel, R., Anwander, A., Geyer, S., & Schönknecht, P.
 456 (2017). Intensity standardisation of 7T MR images for intensity-based segmentation of the
 457 human hypothalamus. *PLOS ONE*, 12(3), e0173344.
 458 <https://doi.org/10.1371/journal.pone.0173344>
 459 Schönknecht, P., Anwander, A., Petzold, F., Schindler, S., Knösche, T. R., Möller, H. E., Hegerl, U.,
 460 Turner, R., & Geyer, S. (2013). Diffusion imaging-based subdivision of the human
 461 hypothalamus: A magnetic resonance study with clinical implications. *European Archives of*
 462 *Psychiatry and Clinical Neuroscience*, 263(6), 497–508. [https://doi.org/10.1007/s00406-012-](https://doi.org/10.1007/s00406-012-0389-5)
 463 0389-5
 464 Seong, J., Kang, J. Y., Sun, J. S., & Kim, K. W. (2019). Hypothalamic inflammation and obesity: A
 465 mechanistic review. *Archives of Pharmacal Research*, 42(5), 383–392.
 466 <https://doi.org/10.1007/s12272-019-01138-9>
 467 Shi, J., & Malik, J. (2000). Normalized cuts and image segmentation. *IEEE Transactions on Pattern*
 468 *Analysis and Machine Intelligence*, 22(8), 888–905. <https://doi.org/10.1109/34.868688>
 469 Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H.,
 470 Bannister, P. R., De Luca, M., Drobnjak, I., Flitney, D. E., Niazy, R. K., Saunders, J., Vickers, J.,
 471 Zhang, Y., De Stefano, N., Brady, J. M., & Matthews, P. M. (2004). Advances in functional and
 472 structural MR image analysis and implementation as FSL. *NeuroImage*, 23 Suppl 1, S208-219.
 473 <https://doi.org/10.1016/j.neuroimage.2004.07.051>
 474 *Statistical Parametric Mapping*. (2007). Elsevier. [https://doi.org/10.1016/B978-0-12-372560-](https://doi.org/10.1016/B978-0-12-372560-8.X5000-1)
 475 8.X5000-1

Tabelow, K., Balteau, E., Ashburner, J., Callaghan, M. F., Draganski, B., Helms, G., Kherif, F., Leutritz, T., Lutti, A., Phillips, C., Reimer, E., Ruthotto, L., Seif, M., Weiskopf, N., Ziegler, G., & Mohammadi, S. (2019). HMRI – A toolbox for quantitative MRI in neuroscience and clinical research. *NeuroImage*, 194, 191–210. <https://doi.org/10.1016/j.neuroimage.2019.01.029>

Thomas, K., Beyer, F., Lewe, G., Zhang, R., Schindler, S., Schönknecht, P., Stumvoll, M., Villringer, A., & Witte, A. V. (2019). *Altered hypothalamic microstructure in human obesity* [Preprint]. Neuroscience. <https://doi.org/10.1101/593004>

Tuch, D. S. (2004). Q-ball imaging. *Magnetic Resonance in Medicine*, 52(6), 1358–1372. <https://doi.org/10.1002/mrm.20279>

Ulla, M., Bonny, J. M., Ouchchane, L., Rieu, I., Claise, B., & Durif, F. (2013). Is R2* a New MRI Biomarker for the Progression of Parkinson’s Disease? A Longitudinal Follow-Up. *PLoS ONE*, 8(3). <https://doi.org/10.1371/journal.pone.0057904>

von Luxburg, U. (2007). A tutorial on spectral clustering. *Statistics and Computing*, 17(4), 395–416. <https://doi.org/10.1007/s11222-007-9033-z>

Ward, R. J., Zucca, F. A., Duyn, J. H., Crichton, R. R., & Zecca, L. (2014). The role of iron in brain ageing and neurodegenerative disorders. *The Lancet. Neurology*, 13(10), 1045–1060. [https://doi.org/10.1016/S1474-4422\(14\)70117-6](https://doi.org/10.1016/S1474-4422(14)70117-6)

Weiskopf, N., Suckling, J., Williams, G., Correia, M. M., Inkster, B., Tait, R., Ooi, C., Bullmore, E. T., & Lutti, A. (2013). Quantitative multi-parameter mapping of R1, PD*, MT, and R2* at 3T: A multi-center validation. *Frontiers in Neuroscience*, 7. <https://doi.org/10.3389/fnins.2013.00095>

Wolff, J., Schindler, S., Lucas, C., Binninger, A.-S., Weinrich, L., Schreiber, J., Hegerl, U., Möller, H. E., Leitzke, M., Geyer, S., & Schönknecht, P. (2018). A semi-automated algorithm for hypothalamus volumetry in 3 Tesla magnetic resonance images. *Psychiatry Research: Neuroimaging*, 277, 45–51. <https://doi.org/10.1016/j.psychresns.2018.04.007>

502 **Acknowledgments**

503 This work was supported by the Neuroimaging Unit of the Carl von Ossietzky Universität Oldenburg
504 funded by grants from the German Research Foundation (3T MRI INST 184/152-1 FUGG). The authors
505 wish to thank Tina Schmitt, Gülsen Yanc, and Katharina Grote for helping with MRI data acquisition.