Title: Tracing diurnal differences in brain anatomy with voxel-based morphometry - associations with sleep characteristics and data quality

Authors and affiliations

Michal Rafal Zareba [1,2,*], Magdalena Fafrowicz [3], Tadeusz Marek [3], Halszka Oginska [3], Ewa Beldzik [3], Aleksandra Domagalik [2,*]

[1] Institute of Zoology and Biomedical Research, Faculty of Biology, Jagiellonian University, 30-387 Kraków, Poland

[2] Centre for Brain Research, Jagiellonian University, 31-501 Kraków, Poland

[3] Department of Cognitive Neuroscience and Neuroergonomics, Institute of Applied Psychology, Jagiellonian University, 30-348 Kraków, Poland

* corresponding authors: michal.rafal.zareba@gmail.com (M.R.Z.), aleksandra.domagalik@uj.edu.pl (A.D.)

Abstract

In recent decades magnetic resonance imaging (MRI) has established itself as the golden standard for studying human brain structure. As a result, it has become increasingly more important to identify technical and physiological factors that may influence study outcomes and contribute to misleading conclusions. This work set out to investigate and compare the variability in voxel-based morphometry (VBM) associated with data quality and time-of-day (TOD) when images are collected using a large longitudinal dataset (N = 72). To test the replicability of the results connected to each factor, the analyses were conducted in two comparable cohorts (each N = 36). The replicated TOD differences were further tested for their correlation with indices of sleep. We report that TOD is a significant and partially replicable predictor of the variability in VBM with the magnitude of the observed differences regionally connected to sleep length. Although data quality also has a prevalent impact on VBM, the affected areas were inconsistent across the two datasets. Interestingly, certain overlapping regions were identified in both analyses, showing that longitudinal results could be misattributed to image quality unless this factor is accounted for. All in all, both TOD and data quality are crucial variables to be considered and controlled for at the stage of experimental planning, as well as data acquisition and analysis. Inclusion of these parameters in the study designs can greatly improve the reliability and replicability of the reported MRI results.

Keywords

voxel-based morphometry; time-of-day; sleep length; data quality; grey matter; white matter

1. Introduction

Recent decades have seen a dramatic rise in the use of magnetic resonance imaging (MRI) for studying the human brain structure and function in health and disease. Good contrast between the grey matter (GM), white matter (WM) and cerebrospinal fluid (CSF) along with the method's non-invasiveness and ability to measure brain activity have deemed it a golden standard for research purposes. In particular, one of the most widely used anatomical techniques

is T1-weighted voxel-based morphometry (VBM), which measures the relative volume of GM and WM in each voxel in the brain. Its usefulness stems from a combination of short data acquisition times, fully automated image processing pipelines and studies relating the computed parameters to histological data (Keifer et al., 2015; Asan et al., 2021).

As the involvement of MRI in brain research has become ubiquitous, more and more attention has been paid to variables that may influence study outcomes. It has been shown that VBM data can be affected by technical factors such as pulse sequence used for data acquisition, magnetic field strength, signal-to-noise ratio (SNR) or head motion (Goto et al., 2013; Reuter et al., 2015; Seiger et al., 2015; Pardoe et al., 2016; Pardoe and Martin, 2022). On the other hand, the development of chronobiology and the observation of diurnal variations at all levels of the organisation of biological systems has made us aware of an important point: we are in a continuous process, we are a change, and our anatomy is just a 'snapshot' of physiology. Indeed, previous MRI studies have found widespread time-of-day (TOD)-dependent differences in task-related brain activity, resting-state functional connectivity, as well as WM microstructure (Marek et al., 2010; Thomas et al., 2018; Fafrowicz et al., 2019; Orban et al., 2020; Voldsbekk et al., 2020; Farahani et al., 2021).

To the best of our knowledge, only one work has investigated the TOD variability in VBM measures (Trefler et al., 2016). The work reported that 4 hours of controlled rest were associated with GM volume decreases in several areas within the frontal, temporal and parietal lobes, as well as in the cerebellum and caudate nucleus. A follow up study by the same research group has suggested TOD-related increase in mean diffusivity of CSF-like water pool along the brain surface, major blood vessels and perivascular spaces as a plausible mechanism for the observed effect (Thomas et al., 2018). Nevertheless, in the course of the day, we also perform numerous cognitive tasks and these processes have been associated with swelling of neurons and glial cells (Xie et al., 2013), new dendritic spines formation (Hübener and Bonhoeffer, 2010), as well as increases of myelin sheath around electrically active axons as a result of oligodendrocyte proliferation and differentiation (Gibson et al., 2014; Xiao et al., 2016), all happening in the time frame of hours after the activity. The described examples of experience-

dependent plasticity should be associated with increased GM and WM volumes as measured with VBM (Keifer et al., 2015; Asan et al., 2021).

Interestingly, both experience-dependent plasticity and CSF-like water diffusivity are related to indices of sleep. Rapid eye movement (REM) sleep has been linked to oligodendrocyte proliferation and removal of excessive synaptic connections to create the space for future memory storage (Bellesi et al., 2013; Li et al., 2017). In turn, poor sleep quality has been linked to enlarged perivascular spaces and increased CSF tracer enrichment (Berezuk et al., 2015; Eide et al., 2022). On the other hand, a different study reported the existence of region-specific sleep-related alterations in the apparent diffusion coefficient across the brain, with the magnitudes of the regional differences in the metric both positively and negatively related to sleep fragmentation (Demiral et al., 2019). Despite the growing interest in the subject, the relationship between sleep and TOD variability in brain structure still remains unexplored.

The first goal of the current study was to build on the aforementioned work by Trefler and colleagues (Trefler et al., 2016), who investigated VBM GM TOD effect in a tightly-controlled laboratory setting. By doubling the between-session difference in circadian time to 8 hours and allowing the participants to engage in their daily non-strenuous activities, we wanted to create a scenario which to a larger extent resembles the typical studies where TOD effect is not controlled at the stage of experiment planning. We additionally wanted to extend the state of the literature by incorporating WM volume into the models for the first time. We hypothesised that both TOD increases and decreases in GM and WM volume could be observed as a result of, respectively, experience-dependent plasticity and permeation of CSF-like water into the brain parenchyma. Secondly, we aimed to investigate the variability in GM and WM volumes associated with a number of image quality metrics to see whether the observed TOD effects could be misattributed to differences in data quality. To further ensure the validity of our results, the original dataset (N = 72) was split into two equal halves, named discovery and replication cohorts, and we attempted to replicate the findings observed in one half of the dataset in the remaining participants of the study. In the case of replicated TOD effects that were unrelated to differences in data quality, we correlated them with two indices of sleep health, i.e. sleep length

and sleep fragmentation. We expected the volumetric TOD increases, predominantly linked to the synaptic plasticity, to be positively associated with the sleep metrics. As for the volumetric TOD decreases, mainly related to the diffusivity of CSF-like water, both positive and negative correlations were expected, similarly to the previously reported results (Demiral et al., 2019). The current work was performed in an extended cohort (N = 72) from earlier works showing widespread TOD variation in resting-state functional connectivity (Fafrowicz et al., 2019; Farahani et al., 2021). To the best of our knowledge, it is the largest dataset used to investigate the interindividual variability in the TOD effect for the brain anatomy in a voxel- or vertex-wise manner.

2. Materials and methods

2.1. Participants

77 volunteers were recruited for the study. All participants were right-handed, had normal or corrected to normal vision, no history of neurological and psychiatric disorders, and were drug-free. The following additional inclusion criteria were applied: age between 19 and 35 years, no shift work, not having been on a flight crossing more than two time zones within the past two months, declared regular daily schedule with no sleep debt (from 6.5 to 9 hours of sleep each night). The level of daytime sleepiness was controlled with Epworth Sleepiness Scale (ESS; Johns, 1991) and sleep quality with Pittsburgh Sleep Quality Index (PSQI; Buysse et al., 1989). The regularity of participants' sleep-wake schedule was checked for one week prior to MRI scanning using MotionWatch8 actigraphs (CamNtech Ltd.). After inspection of the MRI data, five participants were excluded from the study due to unsatisfactory removal of meninges during the brain segmentation process. The final sample consisted of 72 participants (45 females) and was divided into the discovery (N = 36) and replication (N = 36) datasets to assess the replicability of the findings. A summary of the demographics for the two cohorts and the full sample is provided in the Supplementary Material Table 1.

Each participant underwent two scanning sessions: one in the morning, and the other one in the evening, about 1 and 10 hours after awakening, respectively. The beginning of

experimental sessions (morning and evening) was adjusted to the declared and objective (derived from actigraphic data) sleep-wake patterns of participants on a regular working day. The morning scans were acquired between 8:00 AM and 10:20 AM, whereas the evening scans were collected between 5:00 PM and 7:20 PM. The session order was counterbalanced across participants in the full sample, as well as across the discovery and replication cohorts. Participants completed the study on one day (when the morning session was first) or on two consecutive days (when the evening session was first). On the night before the morning session participants slept in the rooms located in the same building as the MRI laboratory. They were only allowed to engage in non-strenuous activities during the experimental days. Participants abstained from alcohol (48 h) and caffeine (24 h) before the first session. Caffeine and alcohol were also banned during the study days. The study was conducted under the Declaration of Helsinki and approved by the Ethics Committee for Biomedical Research at the Military Institute of Aviation Medicine (Warsaw, Poland; 26 February 2013) and the Research Ethics Committee of the Institute of Applied Psychology at Jagiellonian University (Kraków, Poland; 21 February 2017).

As the current dataset was a part of a larger investigation into the effects of chronotype on brain morphometric measures (Zareba et al., 2021), we additionally tested in the full sample whether there was any interaction between the circadian preference (Oginska, 2011; Oginska et al., 2017) and TOD volumetric brain differences. No such effect was found ($p > 0.05$) and thus chronotype was not included in the models.

2.2. Actigraphy-derived sleep measurements

The sleep-wake patterns of the participants were monitored with MotionWatch8 actigraphs (CamNtech Ltd.). The devices were given to the subjects one week prior to MRI scanning; participants wore them non-stop on the wrist of the non-dominant hand. By pressing the 'event' button', they marked the moments when the lights went out and when they woke up. The following parameters were extracted from full week registrations with use of the MotionWare Software: (1) average fall asleep time; (2) average wake-up time; (3) time in bed;

(4) ‘actual’ or ‘total’ sleep time; (5) sleep efficiency (actual sleep time expressed as a percentage of time in bed); (6) sleep latency (time between switching off the lights and falling asleep); (7) sleep fragmentation index based on night-time movement data (higher values indicate greater mobility during the night - it can be used as an indication of sleep quality). Sleep length (actual sleep time) and sleep fragmentation index were chosen as the variables of interest for correlation with the replicated TOD findings. The complete actigraphic data from 3 participants was not available due to battery or connection failure, meaning the sleep parameters were available for 69 subjects. None of the sleep measures differed between the discovery and replication cohorts ($p > 0.05$). Their summary for both groups and the full sample is provided in the Supplementary Material Table 1.

2.3. MRI data acquisition

The high resolution, structural images used in the described investigation were acquired using a 3T scanner (Magnetom Skyra, Siemens) equipped with a 64-channel head/neck coil with a T1 MPRAGE sequence (field of view = 256 x 248 mm; 176 sagittal slices; 1x1x1.1 mm3 voxel size; repetition time = 2300 ms, echo time = 2.98 ms, inversion time = 900 ms, flip angle = 9°, GRAPPA acceleration factor 2). The anatomical MRI data was collected as a part of a longer protocol that additionally included other MRI modalities.

2.4. MRI data quality assessment

All the images that entered the analyses were found to have satisfactory (n = 19) or good data quality (n = 125) according to the quality check reports generated by the Computational Anatomy Toolbox (CAT12 version 12.8; run under MATLAB 2019b). Nevertheless, to exclude the possibility that the differences in the individual image quality affected the results, we applied the Preprocessed Connectomes Project Quality Assessment Protocol (PCP QAP; Shehzad et al., 2015) to further investigate the matter. For each T1-weighted scan, six quality metrics were calculated: signal-to-noise ratio (SNR; mean GM intensity divided by the standard deviation (SD) of background voxels), contrast-to-noise ratio

(CNR; the difference between the mean WM and GM intensities divided by the SD of background voxels), foreground-to-background energy ratio (FBER; variance of voxels inside of head divided by variance of voxels outside of it), voxel smoothness as full width at half maximum (FWHM), entropy focus criterion (EFC; Shannon entropy of voxel signal in proportion to the highest possible entropy for a similarly sized image; it indicates motion and ghosting artefacts) and fraction of artefact background voxels (QI1; the proportion of background voxels with higher signal intensities than the most frequently occuring value). Their distributions across the dataset are shown in the Supplementary Material Figures 1-6.

The within-subject comparisons revealed significant between-session differences in the image quality for all the metrics but FWHM (see Supplementary Material Table 2 for more details). As moderate and strong correlations were found between the quality indices (see Supplementary Material Table 3), principal component analysis (PCA) was performed in R (princomp function, version 4.0.3) to reduce the data dimensionality and identify the factors most contributing to scan quality variability (see Supplementary Material Tables 4-5 for results). In line with the Kaiser criterion (i.e. component eigenvalues higher than 1), the first three components, accounting together for 90.7% of the variance, were chosen for further analyses. The first component (PC1) was most strongly associated with positive loadings from SNR, CNR and FBER, the second component (PC2) was most strongly related to negative loadings from EFC and QI1, whereas the third component (PC3) was mostly associated with a positive loading from FWHM. PC1 was found to have lower values for the evening session scans (paired t-test, T-stat = -6.06, $p < 0.001$), the opposite was observed for PC2 (paired t-test, T-stat = 5.08, $p < 0.001$) and no between-session differences were found for PC3 (paired t-test, T-stat = 0.89, $p = 0.376$). Thus, PC1 and PC2 were included in the VBM models as variables of interest. The two cohorts did not differ in terms of the PC1 values in both sessions, nor in the PC2 values for the morning session ($p > 0.05$). A significant difference was observed for the PC2 in the evening session (two sample t-test, T-stat = 2.82, family-wise error rate (FWE)-corrected $p = 0.037$) with the replication cohort displaying higher values (lower data quality) compared to the discovery cohort. Nevertheless, the PC2 in the discovery cohort spanned across

a broader range of values (from -3.02 to 1.92) than in the replication cohort (from -1.80 to 2.42). Additionally, the two groups had similar between-session differences in both PC1 and PC2 values ($p > 0.05$). As such, we believe that this factor has a limited impact on the results of the subsequently described analyses.

2.5. Voxel-based morphometry

The brain images underwent processing with the standard VBM pipeline (Ashburner and Friston, 2000) in CAT12. Firstly, each volume was segmented into GM, WM and CSF, which was followed by normalisation into the MNI152 space using diffeomorphic anatomical registration through exponentiated lie algebra (DARTEL). Such normalised GM and WM segments were subsequently modulated with the Jacobian determinant to account for the resulting signal intensity changes. The preprocessed GM and WM volumes were inspected visually to ensure satisfactory quality of the segmentation. Lastly, each of the segmented, normalised, and modulated images underwent smoothing with a 4-mm Gaussian filter.

The statistical analysis was performed separately for each tissue class using the 3dLME program (Chen et al., 2013) available in AFNI (Cox, 1996). Voxels with estimated mean study-wise volumes lower than 0.1 were excluded from the analyses to prevent the inclusion of non-brain data. In the discovery dataset, two models were made for each tissue class. The first model incorporated PC1 and PC2 as the variables of interest. The second model aimed to investigate the TOD effect using a paired t-test. In both cases sex, age and total intracranial volume (ICV) were controlled as covariates. The multiple comparison correction was achieved through cluster-level FWE < 0.05 following the initial voxel-level thresholding ($p < 0.001$) to allow for the results comparison at the identical statistical cut-off.

In the case of the data quality-related findings, we prioritised assessing the replicability of the results over estimating their unbiased effect sizes, and thus the model from the discovery dataset was repeated in the replication cohort with two adjustments. Firstly, the significant clusters from the discovery models were saved as binary masks containing all the suprathreshold voxels for the given tissue type and were subsequently used as small-volume

correction (SVC) masks in order to limit the multiple comparison problem. Secondly, to further increase the statistical power, the multiple comparison correction was done with voxel-level false discovery rate (FDR) < 0.05 and minimal cluster extent (k) of 10 voxels.

In the case of the TOD findings, we aimed to simultaneously test the replicability of the results and estimate their effect sizes in an unbiased manner using Cohen's d. The significant clusters from the GM and WM TOD models were saved separately as binary masks and were subsequently used to extract the cluster-averaged brain tissue volumes in the replication dataset. The statistical significance was assessed with paired t-tests.

3. Results

3.1. Impact of data quality on voxel-based morphometry

The variability in VBM associated in the discovery dataset with image quality is presented in Table 1 and Figure 1. PC1 was related to differences in GM volume across frontal, parietal, temporal and occipital lobes, as well as insula, thalamus and cerebellum. PC1 was also connected to differences in WM volume in the orbitofrontal cortex, internal capsule and fornix. PC2 was in turn related to variability of GM and WM volume in the occipital lobe. The SVC analysis in the replication dataset yielded no significant findings (voxel-level FDR = 0.05, k = 10). For comparison, the results of the whole-brain analysis in the second cohort are presented in the Supplementary Material Table 6 and Supplementary Material Figure 7.

Table 1. Data quality-related differences in VBM in the discovery dataset (cluster-level FWE < 0.05).

Tissue class	MNI	Voxels	Location	Component	F-stat
GM	42, -2, 7	274	R insula	PC1	50.14
GM	2, 23, 31	224	B anterior cingulate cortex	PC1	27.94

GM	6, -14, -1	157	R thalamus	PC1	28.47
GM	64, 8, 28	146	R precentral gyrus	PC1	21.86
GM	47, -5, -15	133	R superior temporal gyrus	PC1	26.20
GM	-43, -73, -28	123	L cerebellar lobule VIIa crus I L inferior occipital gyrus	PC1	26.99
GM	47, -73, 36	109	R angular gyrus	PC1	20.02
GM	-58, -43, 21	105	L superior temporal gyrus	PC1	26.60
GM	-4, -68, 68	105	L precuneus	PC1	56.55
GM	-19, -30, 77	105	L postcentral gyrus	PC1	34.48
GM	-18, -82, -21	114	L lingual gyrus	PC2	36.26
GM	21, -100, 15	105	R superior occipital gyrus	PC2	36.60
WM	12, -6, -3	291	R internal capsule	PC1	37.00
WM	15, 51, -14	178	R orbitofrontal cortex	PC1	32.06
WM	-1, -2, -2	103	L internal capsule L fornix	PC1	36.25
WM	-13, -82, -15	176	L occipital lobe, subgyral	PC2	30.28

Abbreviations: GM, grey matter; WM, white matter; B, bilateral; R, right; L, left; PC1, principal component 1; PC2, principal component 2.

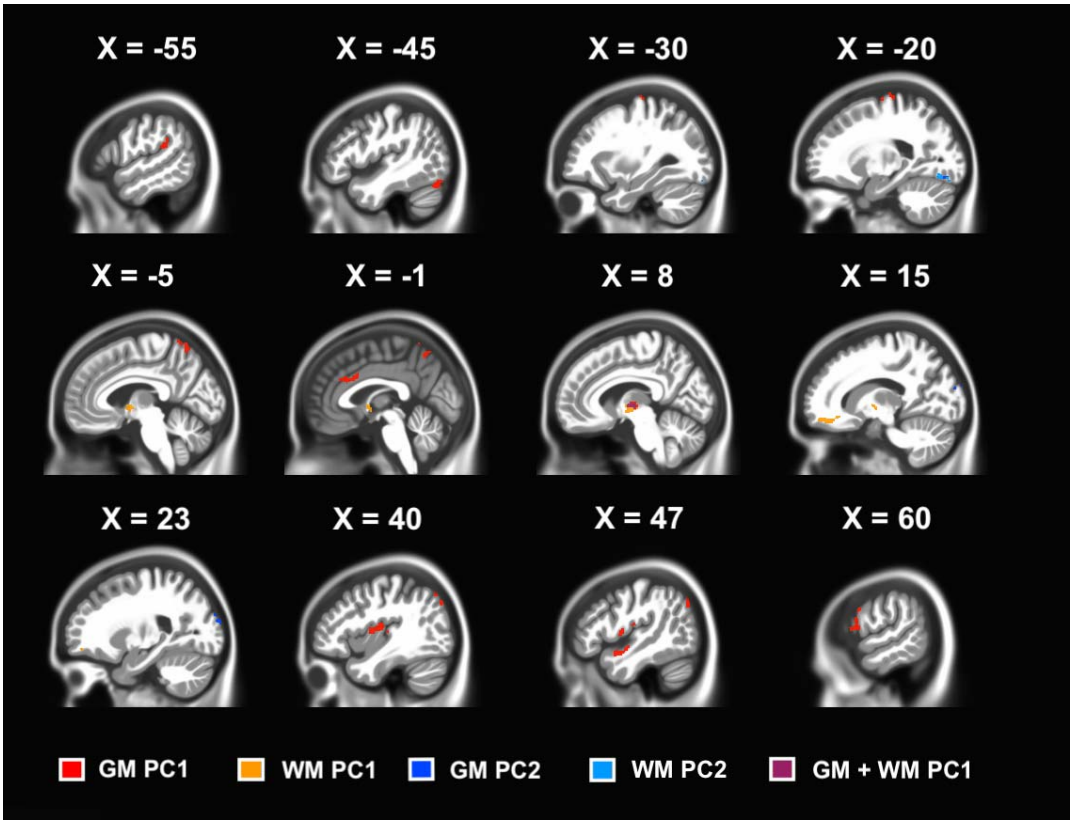

Figure 1. The variability in VBM associated with data quality in the discovery dataset (cluster-level FWE < 0.05). Principal component 1 (PC1) was related to signal-to-noise ratio, contrast-to-noise ratio and foreground-to-background energy ratio. Principal component 2 (PC2) was associated with entropy focus criterion, a measure related to motion and ghosting artefacts, and fraction of artefact background voxels. Abbreviations: GM, grey matter; WM, white matter.

3.2. Time-of-day differences in voxel-based morphometry

The TOD differences in VBM in the discovery dataset are depicted in Table 2 and Figure 2. The evening scans were associated with lower GM volume in the bilateral retrosplenial cortex, together with the right occipital pole, inferior temporal gyrus, thalamus and cerebellar lobule VIIa. Furthermore, the evening session was characterised by decreased WM volume in bilateral parietal lobes and increased WM volume in the right internal capsule. The clusters of decreased GM volume in the right thalamus and increased WM volume in the right internal capsule largely overlapped with each other (see Figure 2). Interestingly, these two areas

were also affected by the PC1 metric of data quality in the previous analysis. The comparisons in the second dataset replicated 2 out of 7 clusters from the primary investigation, i.e. morning-to-evening decreases in GM volume in the bilateral retrosplenial cortex (MNI: -6, -47, 4; Cohen's d = -0.633; FWE = 0.004) and WM volume in the right parietal lobe (MNI: 38, -30, 41; Cohen's d = -0.568; FWE = 0.011). All the unreplicated findings were far from reaching the significance (uncorrected p values ranging from 0.158 to 0.584).

Table 2. Time-of-day (TOD) differences in the voxel-based morphometry (VBM) in the discovery dataset (cluster-level FWE < 0.05).

Tissue class	**MNI**	**Voxels**	**Location**	**Directionality**	**t-stat**
GM	-6, -47, 4	213	B retrosplenial cortex	M > E	4.93
GM	2, -85, -34	151	R cerebellar lobule VIIa crus II R occipital pole	M > E	4.65
GM	8, -11, -1	123	R thalamus	M > E	4.94
GM	44, -42, -38	89	R cerebellar lobule VIIa crus I R inferior temporal gyrus	M > E	4.30
WM	8, 3, -2	365	R internal capsule	E > M	5.24
WM	38, -30, 41	101	R parietal lobe, sublobar	M > E	4.56
WM	-28, -72, 22	95	L parietal lobe, sublobar	M > E	4.64

Abbreviations: GM, grey matter; WM, white matter; B, bilateral; R, right; L, left; M, morning; E, evening.

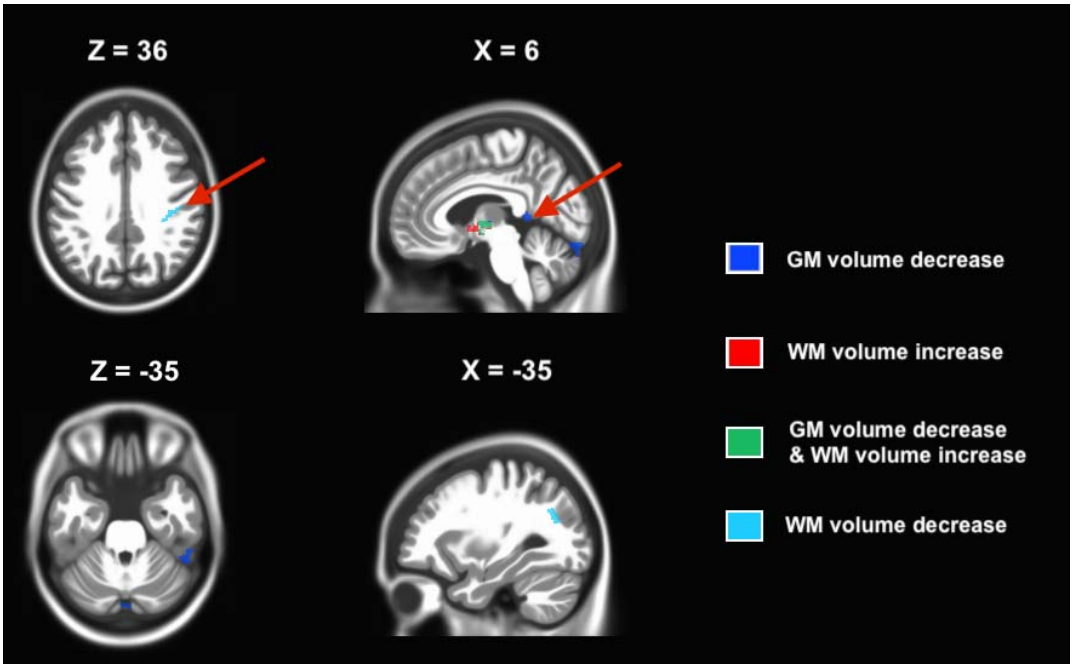

Figure 2. The time-of-day (TOD) effect on the voxel-based morphometry (VBM) in the discovery dataset (cluster-level FWE < 0.05). Red arrows indicate clusters that were also significant in the replication cohort. Abbreviations: GM, grey matter; WM, white matter.

3.3. Association between time-of-day results and sleep indices

Sleep time, sleep fragmentation and the replicated TOD volumetric differences were all normally distributed according to the Kolmogorov–Smirnov test ($p > 0.05$), and thus Pearson's correlations were calculated to evaluate the associations between sleep indices and structural brain differences. The TOD GM effect in the bilateral retrosplenial cortex was significantly correlated with sleep time ($r = 0.34$, FWE = 0.016; see Figure 3) but not sleep fragmentation index ($r = -0.18$, FWE = 0.52). Neither sleep time ($r = -0.01$, FWE = 1), nor sleep fragmentation index ($r = -0.11$, FWE = 1) was related to the TOD difference in WM volume in the right parietal lobe.

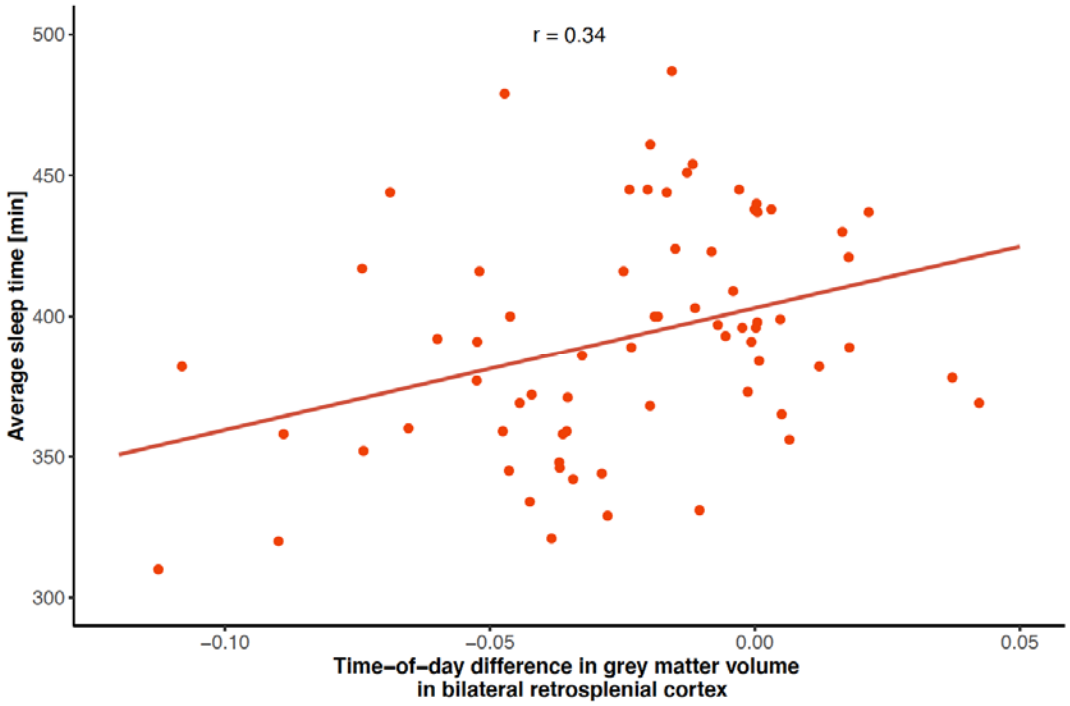

Figure 3. The TOD change in the GM volume in the bilateral retrosplenial cortex was positively correlated with the average sleep time over the course of the previous week (r = 0.34).

4. Discussion

4.1. Pronounced and dataset-specific effects of image quality on voxel-based morphometry

Our investigation revealed that in the discovery dataset PC1, a measure related to SNR, CNR and FBER, was associated with GM volume variation spread across the four brain lobes, subcortical areas and cerebellum, as well as differences in frontal and subcortical WM volume. In turn, PC2, related to the in-scanner motion and characteristics of background noise, was connected to the estimated GM and WM volumes only in the occipital lobe. Thus, our work corroborates the previous studies showing that VBM can be affected by these factors (Goto et al., 2013; Reuter et al., 2015; Pardoe et al., 2016; Pardoe and Martin, 2022). However, in contrast with these reports, our study also made an important attempt to establish whether the data quality would similarly impact the estimated volumetric metrics across separate datasets. Despite comparable levels of quality metrics in the two cohorts, none of the results found in the discovery group was replicated. While in the discovery dataset a vast majority of the data

quality-related results was linked to PC1, the whole-brain analysis in the replication cohort yielded significant associations for GM and WM volumes only with PC2. This indicates that the impact of image quality on the brain morphometric measures is likely dataset-specific and difficult to directly compare between the studies. As in the case of the discovery cohort two of the clusters from the TOD analysis overlapped with the areas affected by PC1, our work underlines the need for these hardware-related contributions to be strictly controlled at the stage of experiment planning and data acquisition. The TOD differences in the data quality metrics have been reported before (Vogelbacher et al., 2018) and have been suggested to result from factors such as room temperature and heating up of the scanner after numerous imaging sessions (Clare et al., 2006; Shuter et al., 2008).

4.2. Influence of time-of-day on voxel-based morphometry

To the best of our knowledge, our study is the first one to simultaneously consider the TOD effect on GM and WM VBM. After excluding the results overlapping with the data quality-related findings, GM volume reduction was found in the bilateral retrosplenial cortex, as well as the right cerebellar hemisphere, occipital pole and inferior temporal gyrus, while WM volume decrease was seen bilaterally in the parietal white matter. Our work corroborates the earlier reports showing TOD effects on GM structure as assessed with VBM and surface-based morphometry (Trefler et al., 2016; Elvsåshagen et al., 2017). In the case of WM, we extend the current state of the field by showing that WM volume is also affected by TOD, similarly to the diffusion of water molecules within this tissue (Voldsbekk et al., 2020).

Despite doubling the difference in circadian time between the imaging sessions to 8 hours, we observed no regional increases in either GM or WM volume, alike the earlier study by Trefler and colleagues (Trefler et al., 2016). This indicates that in the TOD context typical for MRI studies the effects associated with increased CSF-like water diffusivity might have greater impact on the signal characteristics than the experience-dependent plasticity, especially that CSF permeates the superficial cortical layers even during wakefulness and the amount of fluid detectable in perivascular spaces of WM increases with TOD (Xie et al., 2013; Barisano et al.,

2021). While in the case of our study participants were engaged in a variety of daily cognitive processes between the imaging sessions, it would be interesting to investigate how different the results would be if the subjects instead engaged in a cognitive training that affects only specific neural networks, for example by building on some earlier diffusion studies in humans and animals (Hofstetter et al., 2013). We believe that this lack of strict control of between-session activity of the participants, aimed at mimicking the designs of a vast majority of MRI studies, may partially contribute to the fact that only 2 out of 7 VBM TOD clusters were replicated in the second dataset (with medium effect sizes). However, such controlled conditions are often difficult to arrange in the research settings. Thus, we suggest that limiting the impact of TOD on VBM should be achieved through at least scheduling participants' scans to a narrow window of time since their (habitual) wake up times.

4.3. Association between sleep and daily differences in voxel-based morphometry

Last but not least, our analysis revealed that the daily change in the GM volume in the bilateral retrosplenial cortex had a moderate correlation with the average sleep time of the participants during the previous week ($r = 0.34$). This finding serves as a further premise for the biological basis of the observed TOD results and makes a significant contribution to the field by identifying a factor that can modulate the magnitude of the VBM TOD effect.

Contrary to our hypothesis, the aforementioned relationship was the only significant association between the daily differences in tissue volume and measures of sleep health. It cannot be excluded that variation in sleep length and sleep fragmentation exhibits a stronger effect on the diffusivity of CSF-like water pool in general rather than the TOD fluctuation in this metric. Poor sleep quality has been associated with enlarged perivascular spaces and increased CSF tracer enrichment (Berezuk et al., 2015; Eide et al., 2022). While the amount of fluid detectable within perivascular spaces increases with TOD (Barisano et al., 2021) and has been shown to be under circadian control (Hablitz et al., 2020), no studies have examined the impact of sleep health on the TOD differences in the discussed measures. Future studies are thus warranted to fill this crucial gap in our understanding of the described phenomena.

4.4. Limitations

As discussed before, in our study participants were allowed to perform their typical non-strenuous activities during the experimental days. Thus, one limitation of our work lies in the fact that its design does not enable a clear differentiation between circadian- and experience-dependent effects. Nevertheless, it can also be considered one of its major strengths as neuroimaging studies often do not account for the time of data acquisition and participants' activities performed on the day of the scan prior to the visit in the laboratory. Thus, our findings constitute a representative of the effects that can be otherwise misattributed to experimental manipulation.

As our datasets were collected only in young individuals and ageing has been shown to impact the glymphatic functions and synaptic plasticity (Bloss et al., 2011; Salman et al., 2022), we believe that the results presented in this manuscript may not be fully replicable in the older groups. Furthermore, despite our interest in investigating the relationship between the sleep data and TOD variability in GM and WM volume, we restricted the currently described analyses only to the individuals without any sleep-related disorders. Future studies are therefore warranted to enrich our knowledge by responding to these limitations.

5. Conclusions

Our work shows that TOD is a significant and partially replicable predictor of the variability in VBM, and the magnitude of the daily differences in tissue volume can be regionally linked to sleep length. In turn, data quality also has a prevalent effect on VBM, however, the impact of this factor is likely dataset-specific and difficult to compare between the studies. All in all, both TOD and data quality are crucial factors to be considered and controlled for at the stage of experimental planning, as well as data acquisition and analysis.

Funding

This work was supported by the National Science Centre, Poland (NCN) [grant number 2013/08/M/HS6/00042].

Data availability

The dataset used in this study will be made available upon manuscript acceptance.

CRediT author statement

MRZ: Conceptualization, Methodology, Investigation, Formal Analysis, Validation, Data Curation, Writing - Original Draft, Visualization. **MF:** Conceptualisation, Methodology, Investigation, Data Curation, Writing - Review & Editing, Funding Acquisition. **TM:** Conceptualisation, Methodology, Writing - Review & Editing, Funding Acquisition. **HO:** Investigation, Formal Analysis, Writing - Review & Editing. **EB:** Investigation, Formal Analysis, Validation, Writing - Review & Editing. **AD:** Conceptualisation, Methodology, Investigation, Formal Analysis, Validation, Data Curation, Writing - Review & Editing, Supervision. All authors contributed to the article and approved the submitted version.

Acknowledgements

This research was supported in part by PL-Grid Infrastructure. We thank Prof. Patricia Reuter-Lorenz for her constructive suggestions during the planning and development of the Harmonia project and her valuable support. We also thank Anna Beres, Piotr Faba, Koryna Lewandowska, Monika Ostrogorska, Barbara Sikora-Wachowicz for their assistance with the fMRI data collection, Aleksandra Zyrkowska for help with the process of participant selection, and Magdalena Debowska for help in collecting actigraphy data.

Declarations of interest

Authors report no conflict of interests.

Glossary

MRI, magnetic resonance imaging; GM, grey matter; WM, white matter; CSF, cerebrospinal fluid; VBM, voxel-based morphometry; SNR, signal-to-noise ratio; TOD, time-of-day; REM, rapid eye movement; ESS, Epworth Sleepiness Scale; PSQI, Pittsburgh Sleep Quality Index; CAT12, Computational Anatomy Toolbox; PCP QAP, Preprocessed Connectomes Project Quality Assessment Protocol; SD, standard deviation; CNR, contrast-to-noise ratio; foreground-to-background energy ratio, FBER; full width at half maximum, FWHM; entropy focus criterion, EFC; fraction of artefact background voxels, QI1; principal component analysis, PCA; FWE, family-wise error rate; DARTEL, diffeomorphic anatomical registration through exponentiated lie algebra; ICV, intracranial volume; SVC, small-volume correction; FDR, false-discovery rate.

References

Asan, L., Falfán-Melgoza, C., Beretta, C.A., Sack, M., Zheng, L., Weber-Fahr, W., Kuner, T., Knabbe, J., 2021. Cellular correlates of gray matter volume changes in magnetic resonance morphometry identified by two-photon microscopy. Sci Rep. 11(1), 4234. doi: 10.1038/s41598-021-83491-8.

Ashburner, J., Friston, K.J., 2000. Voxel-based morphometry--the methods. Neuroimage 11, 805-821. doi: 10.1006/nimg.2000.0582

Barisano, G., Sheikh-Bahaei, N., Law, M., Toga, A.W., Sepehrband, F., 2021. Body mass index, time of day and genetics affect perivascular spaces in the white matter. J Cereb Blood Flow Metab. 41(7), 1563-1578. doi: 10.1177/0271678X20972856.

Bellesi, M., Pfister-Genskow, M., Maret, S., Keles, S., Tononi, G., Cirelli, C., 2013. Effects of sleep and wake on oligodendrocytes and their precursors. J Neurosci. 33(36), 14288-14300. doi: 10.1523/JNEUROSCI.5102-12.2013.

Berezuk, C., Ramirez, J., Gao, F., Scott, C.J., Huroy, M., Swartz, R.H., Murray, B.J., Black, S.E., Boulos, M.I., 2015. Virchow-Robin Spaces: Correlations with Polysomnography-Derived Sleep Parameters. Sleep 38(6), 853-858. doi: 10.5665/sleep.4726.

Bloss, E.B., Janssen, W.G., Ohm, D.T., Yuk, F.J., Wadsworth, S., Saardi, K.M., McEwen, B.S., Morrison, J.H., 2011. Evidence for reduced experience-dependent dendritic spine plasticity in the aging prefrontal cortex. J Neurosci. 31(21), 7831-7839. doi: 10.1523/JNEUROSCI.0839-11.2011.

Buysse, D.J., Reynolds, C.F. 3rd, Monk, T.H., Berman, S.R., Kupfer, D.J., 1989. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res. 28(2), 193-213. doi: 10.1016/0165-1781(89)90047-4.

Chen, G., Adleman, N.E., Saad Z.S., Leibenluft, E., Cox, R.W., 2014. Applications of multivariate modeling to neuroimaging group analysis: a comprehensive alternative to univariate general linear model. Neuroimage. 99, 571-88. doi: 10.1016/j.neuroimage.2014.06.027.

Clare, S., Evans, J., Jezzard, P., 2006. Requirements for room temperature shimming of the human brain. Magn Reson Med. 55(1), 210-214. doi: 10.1002/mrm.20735. Erratum in: Magn Reson Med. 2006, 55(6), 1466.

Cox, R.W., 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput Biomed Res. 29(3), 162-73. doi: 10.1006/cbmr.1996.0014.

Demiral, S.B., Tomasi, D., Sarlls, J., Lee, H., Wiers, C.E., Zehra, A., Srivastava, T., Ke, K., Shokri-Kojori, E., Freeman, C.R., Lindgren, E., Ramirez, V., Miller, G., Bandettini, P., Horovitz, S., Wang, G.J., Benveniste, H., Volkow, N.D., 2019. Apparent diffusion coefficient changes in human brain during sleep - Does it inform on the existence of a glymphatic system? Neuroimage 185, 263-273. doi: 10.1016/j.neuroimage.2018.10.043.

Eide, P.K., Pripp, A.H., Berge, B., Hrubos-Strøm, H., Ringstad, G., Valnes, L.M., 2022 Altered glymphatic enhancement of cerebrospinal fluid tracer in individuals with chronic poor sleep quality. J Cereb Blood Flow Metab. 42(9), 1676-1692. doi: 10.1177/0271678X221090747.

Elvsåshagen, T., Zak, N., Norbom, L.B., Pedersen, P.Ø., Quraishi, S.H., Bjørnerud, A., Alnæs, D., Doan, N.T., Malt, U.F., Groote, I.R., Westlye, L.T.., 2017. Evidence for cortical structural plasticity in humans after a day of waking and sleep deprivation. Neuroimage. 156, 214-223. doi: 10.1016/j.neuroimage.2017.05.027.

Fafrowicz, M., Bohaterewicz, B., Ceglarek, A., Cichocka, M., Lewandowska, K., Sikora-Wachowicz, B., Oginska, H., Beres, A., Olszewska, J., Marek, T., 2019. Beyond the Low Frequency Fluctuations: Morning and Evening Differences in Human Brain. Front Hum Neurosci. 13, 288. doi: 10.3389/fnhum.2019.00288.

Farahani, F.V., Fafrowicz, M., Karwowski, W., Bohaterewicz, B., Sobczak, A.M., Ceglarek, A., Zyrkowska, A., Ostrogorska, M., Sikora-Wachowicz, B., Lewandowska, K., Oginska, H., Beres, A., Hubalewska-Mazgaj, M., Marek, T., 2021. Identifying Diurnal Variability of Brain Connectivity Patterns Using Graph Theory. Brain Sci. 11(1), 111. doi: 10.3390/brainsci11010111.

Gibson, E.M., Purger, D., Mount, C.W., Goldstein, A.K., Lin, G.L., Wood, L.S., Inema, I., Miller, S.E., Bieri, G., Zuchero, J.B., Barres, B.A., Woo, P.J., Vogel, H., Monje, M., 2014. Neuronal activity promotes oligodendrogenesis and adaptive myelination in the mammalian brain. Science. 344(6183), 1252304. doi: 10.1126/science.1252304.

Goto, M., Abe, O., Aoki, S., Miyati, T., Takao, H., Hayashi, N., Mori, H., Kunimatsu, A., Ino, K., Yano, K., & Ohtomo, K., 2013. Bilateral pre- and postcentral gyrus volume positively correlates with T2-SNR of putamen in healthy adults. Neuroradiology, 55(2), 245–250. doi: 10.1007/s00234-012-1126-5.

Hablitz, L.M., Plá, V., Giannetto, M., Vinitsky, H.S., Stæger, F.F., Metcalfe, T., Nguyen, R., Benrais, A., Nedergaard, M., 2020. Circadian control of brain glymphatic and lymphatic fluid flow. Nat Commun. 11(1), 4411. doi: 10.1038/s41467-020-18115-2.

Hofstetter, S., Tavor, I., Tzur Moryosef, S., Assaf, Y., 2013. Short-term learning induces white matter plasticity in the fornix. J Neurosci. 33(31), 12844-50. doi: 10.1523/JNEUROSCI.4520-12.2013.

Hübener, M., Bonhoeffer, T., 2010. Searching for engrams. Neuron. 67(3), 363-71. doi: 10.1016/j.neuron.2010.06.033.

Johns, M.W., 1991. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep. 14(6), 540-5. doi: 10.1093/sleep/14.6.540.

Keifer, O.P. Jr, Hurt, R.C., Gutman, D.A., Keilholz, S.D., Gourley, S.L., Ressler, K.J., 2015. Voxel-based morphometry predicts shifts in dendritic spine density and morphology with auditory fear conditioning. Nat Commun. 6, 7582. doi: 10.1038/ncomms8582.

Li, W., Ma, L., Yang, G., Gan, W.B., 2017. REM sleep selectively prunes and maintains new synapses in development and learning. Nat Neurosci. 20(3), 427–437. doi: 10.1038/nn.4479.

Marek, T., Fafrowicz, M., Golonka, K., Mojsa-Kaja, J., Oginska, H., Tucholska, K., Urbanik, A., Beldzik, E., Domagalik, A., 2010. Diurnal patterns of activity of the orienting and executive attention neuronal networks in subjects performing a Stroop-like task: a functional magnetic resonance imaging study. Chronobiol Int. 27(5), 945-58. doi: 10.3109/07420528.2010.489400.

Oginska, H., 2011. Can you feel the rhythm? A short questionnaire to describe two dimensions of chronotype. Pers. Individ. Differ. 50(7), 1039–1043. https://doi.org/10.1016/j.paid.2011.01.020

Oginska, H., Mojsa-Kaja, J., Mairesse, O., 2017. Chronotype description: In search of a solid subjective amplitude scale. Chronobiol Int. 34(10), 1388-1400. doi: 10.1080/07420528.2017.1372469.

Orban, C., Kong, R., Li, J., Chee, M.W.L., Yeo, B.T.T., 2020. Time of day is associated with paradoxical reductions in global signal fluctuation and functional connectivity. PLoS Biol. 18(2), e3000602. doi: 10.1371/journal.pbio.3000602.

Pardoe, H.R., Kucharsky Hiess, R., Kuzniecky, R., 2016. Motion and morphometry in clinical and nonclinical populations. NeuroImage, 135, 177–185. doi: 10.1016/j.neuroimage.2016.05.005.

Pardoe, H.R., Martin, S.P., 2022. In-scanner head motion and structural covariance networks. Human brain mapping, 43(14), 4335–4346. doi: 10.1002/hbm.25957.

Reuter, M., Tisdall, M. D., Qureshi, A., Buckner, R. L., van der Kouwe, A. J. W., Fischl, B., 2015. Head motion during MRI acquisition reduces gray matter volume and thickness estimates. NeuroImage, 107, 107–115. doi: 10.1016/j.neuroimage.2014.12.006.

Salman, M.M., Kitchen, P., Halsey, A., Wang, M.X., Törnroth-Horsefield, S., Conner, A.C., Badaut, J., Iliff, J.J., Bill, R.M., 2022. Emerging roles for dynamic aquaporin-4 subcellular relocalization in CNS water homeostasis. Brain. 145(1), 64-75. doi: 10.1093/brain/awab311.

Seiger, R., Hahn, A., Hummer, A., Kranz, G. S., Ganger, S., Küblböck, M., Kraus, C., Sladky, R., Kasper, S., Windischberger, C., & Lanzenberger, R., 2015. Voxel-based morphometry at ultra-high fields. a comparison of 7T and 3T MRI data. NeuroImage. 113, 207–216. Doi: 10.1016/j.neuroimage.2015.03.019.

Shehzad, Z., Giavasis, S., Li, Q., Benhajali, Y., Yan, C., Yang, Z., Milham, M., Bellec, P., Craddock, C., 2015. The Preprocessed Connectomes Project Quality Assessment Protocol - a resource for measuring the quality of MRI data. Front Neurosci. doi: 10.3389/conf.fnins.2015.91.00047.

Shuter, B., Yeh, I.B., Graham, S., Au, C., Wang, S.C., 2008. Reproducibility of brain tissue volumes in longitudinal studies: effects of changes in signal-to-noise ratio and scanner software. Neuroimage. 41(2), 371-379. doi: 10.1016/j.neuroimage.2008.02.003.

Thomas, C., Sadeghi, N., Nayak, A., Trefler, A., Sarlls, J., Baker, C.I., Pierpaoli, C., 2018. Impact of time-of-day on diffusivity measures of brain tissue derived from diffusion tensor imaging. Neuroimage. 173, 25-34. doi: 10.1016/j.neuroimage.2018.02.026.

Trefler, A., Sadeghi, N., Thomas, A.G., Pierpaoli, C., Baker, C.I., Thomas, C., 2016. Impact of time-of-day on brain morphometric measures derived from T1-weighted magnetic resonance imaging. Neuroimage. 133, 41-52. doi: 10.1016/j.neuroimage.2016.02.034.

Vogelbacher, C., Möbius, T.W.D., Sommer, J., Schuster, V., Dannlowski, U., Kircher, T., Dempfle, A., Jansen, A., Bopp, M.H.A, 2018. The Marburg-Münster Affective Disorders Cohort Study (MACS): A quality assurance protocol for MR neuroimaging data. Neuroimage. 172, 450-460. doi: 10.1016/j.neuroimage.2018.01.079.

Voldsbekk, I., Maximov, I.I., Zak, N., Roelfs, D., Geier, O., Due-Tønnessen, P., Elvsåshagen, T., Strømstad, M., Bjørnerud, A., Groote, I., 2020. Evidence for wakefulness-related changes to extracellular space in human brain white matter from diffusion-weighted MRI. Neuroimage. 212, 116682. doi: 10.1016/j.neuroimage.2020.116682.

Xiao, L., Ohayon, D., McKenzie, I.A., Sinclair-Wilson, A., Wright, J.L., Fudge, A.D., Emery, B., Li, H., Richardson, W.D., 2016. Rapid production of new oligodendrocytes is required in the earliest stages of motor-skill learning. Nat Neurosci. 19(9), 1210-1217. doi: 10.1038/nn.4351.

Xie, L., Kang, H., Xu, Q., Chen, M.J., Liao, Y., Thiyagarajan, M., O'Donnell, J., Christensen, D.J., Nicholson, C., Iliff, J.J., Takano, T., Deane, R., Nedergaard, M., 2013. Sleep drives metabolite clearance from the adult brain. Science. 342(6156), 373-7. doi: 10.1126/science.1241224.

Zareba, M.R., Fafrowicz, M., Marek, T., Beldzik, E., Oginska, H., Domagalik, A., 2021. Late chronotype is linked to greater cortical thickness in the left fusiform and entorhinal gyri. Biol. Rhythm Res. doi: 10.1080/09291016.2021.1990501.