UNIVERSITY OF AMSTERDAM

BACHELOR THESIS IN BRAIN & COGNITION

Parkinson's disease and the Medial Geniculate Nucleus: A 7-Tesla MRI comparison of structural volume between patients and healthy controls

Author:
O. PAPAKYRIAKOU[1]
[1]*Integrative Model-based Cognitive Neuroscience research unit, University of Amsterdam*

Supervisor:
Dr. A. J. M. ALKEMADE

Abstract

Parkinson's is a degenerative neurological disorder, affecting a considerable amount of the elderly and involving the impairment of several brain structures. However, the subcortical structures involved have not been thoroughly investigated, even though they play a major role in normal functioning, like hearing. This study examined whether Parkinson's disease is related to a decrease of structural volume in the medial geniculate nucleus (MGN), the thalamic relay of audition. Participants included 11 patients with Parkinson's and 11 healthy controls. Their brain scans had already been collected at the Spinoza center for Neuroimaging in Amsterdam, using a 7-Tesla MRI scanner. Structural volume, measured in cubic millimeters, was calculated for every subject by manually delineating the MGN in each hemisphere of the scan. For each group, the data were collapsed by taking the mean of the two hemispheres, and they were analyzed with a Bayesian Mann-Whitney U test. Results showed that the data are equally likely under both competing hypotheses, that of a decreased MGN volume in Parkinson's patients compared to healthy controls, and that of no difference between the two groups, $BF10 = 1.22$. This suggests that it is not yet clear whether Parkinson's is related to a decrease of structural volume in the MGN, highlighting the need for further research on the topic. Elucidating the involvement of the MGN in Parkinson's is important for the prognosis of the disease since it could help doctors alleviate symptoms related to audition.

12 ECTS
Submitted January 30, 2022

odysseas.papakyriakou@student.uva.nl

Realized between October 2021 and January 2022

Parkinson's disease and the Medial Geniculate Nucleus: A 7 Tesla MRI comparison of structural volume between patients and healthy controls

Parkinson's disease (PD) is a degenerative disorder of the central nervous system, affecting the elderly with a prevalence of 1-3%, and the general population with 0.3% (Raza et al., 2019). The disorder mainly results in dysfunctions of the somatomotor system including bradykinesia, tremors, and instability, but may also be accompanied by dementia and cognitive impairment, especially in advanced stages (Raza et al., 2019). Besides the well-established link between somatomotor impairments and dopamine deficits in the basal ganglia, grey matter volume seems to be reduced in brain areas that are linked to cognitive impairments in PD (Lee et al., 2013). This is highly relevant for the prognosis of the disease, since it could guide doctors in their prescriptions (Gerrits et al., 2014). Therefore, investigating which brain areas show a volume decrease in relation to PD is of great importance.

When examining the cognitive impairments associated with PD, previous research has focused on problems related to memory and verbal fluency (Gerrits et al., 2016), vision (Pezzoli et al., 2021) or executive functions (Elgh et al., 2009). However, PD cognitive symptoms often include a broad spectrum of auditory impairments, which seem to have evaded scientific scrutiny (Jafari et al., 2020). These impairments may affect the whole auditory cycle, from the cochlear level to cortical auditory areas (Jafari et al., 2020). The cochlear and brainstem involvement in PD has been investigated using pure tone audiometry and brainstem evoked response audiometry respectively (Shetty et al., 2019), while the cortical involvement has been studied with EEG (de Keyser et al., 2021). Interestingly, the involvement of the medial geniculate nucleus (MGN), the thalamic relay between the brainstem and the cortical auditory areas, has not been investigated in relation to PD.

Moreover, the human subcortex, including the MGN, is largely underrepresented in brain atlases and neuroimaging studies (Bazin et al., 2020). Particularly, of the approximately 455 individual structures within the human subcortex, only 7% have been mapped with MRI (Alkemade et al., 2021). This becomes an important issue when considering that subcortical structures are crucial for normal behavior and they are also potential targets of deep brain stimulation for PD and other disorders (Miletić et al., 2022). Importantly, there have been efforts to automate the delineation of subcortical structures using several algorithms, including MASSP (Multi-contrast Anatomical Subcortical Structure Parcellation), which has now been successfully

trained on 17 structures (Bazin et al., 2020). The fact that MASSP has not been trained on the MGN, coupled with the lack of publications on PD and the MGN, motivates the investigation of how the MGN differs between people with PD and healthy controls using 7 Tesla MRI.

Magnetic Resonance Imaging (MRI) is a non-invasive technology that produces three dimensional detailed anatomical images. The technique uses computer-generated radio waves to reconstruct images of tissues and organs by generating a strong magnetic field (Kraff et al., 2015). Because the quality of the produced image depends on the strength of the magnetic field, 7T MRI scanners typically outperform optimized 3T MRI scanners (Isaacs et al., 2020), allowing for enhanced details in imaging of tiny subcortical structures, like the MGN.

The study by Gerrits et al. (2014) explored the relationship between grey matter volume of 93 PD patients and cognitive performance. Grey matter volume was measured with voxel-based morphometry (VBM) on 3T MRI scans of PD patients and healthy controls, matched for age and sex. The cognitive functions involved visuospatial memory, learning, and executive functions, measured with several diagnostic tests. Results showed that there were positive correlations between volume in cortical and subcortical structures and performance in the tests. The study also showed that grey matter volume in several brain regions is reduced in PD patients compared to healthy controls. The results suggest that the lower the performance on a cognitive function is, the lower the grey matter volume of the corresponding area is in PD patients, and that PD patients show a reduction in grey matter volume compared to healthy controls.

Golden et al. (2015) investigated the cortical neuroanatomical correlates of auditory spatial processing in 20 patients with Alzheimer's disease. Even though the clinical population in this study did not suffer from Parkinson's, the results are still informative about how grey matter volume is related to auditory impairments. Grey matter volume was measured using VBM on 3T MRI scans and auditory spatial processing was assessed with several. There were three groups in the study: patients with Alzheimer's, patients with posterior cortical atrophy, and healthy controls. Results showed that grey matter volume in the right inferior parietal lobe was positively associated with performance on these tests. This means that the worse patients perform in the auditory spatial discrimination task, the less cortical grey matter they have in this area, which is informative about how grey matter reduction due to PD may relate to auditory impairments.

A review by Ibarretxe-Bilbao et al. (2011) examined the brain structural correlates of cognitive impairments in PD patients. The studies reviewed measured grey matter volume using

either VBM, or manual parcellations from MRI scans. The results show a positive correlation between cortical grey matter volume and the corresponding cognitive function, including memory, vision, and verbal fluency. This means that the worse patients perform in tasks measuring these functions, the less grey matter volume they have in a cortical area.

Although these studies point to a reduction of grey matter volume when there is cognitive impairment, they do not investigate subcortical structures, like the MGN. A study by Mahoney et al. (2011) investigated changes in grey matter volume of the MGN in people with tinnitus, an auditory impairment involving the false perception of sound (Uluyol et al., 2016). Importantly, tinnitus has a higher prevalence and is experienced more severely among people with PD compared to the general population (Uluyol et al., 2016). The study by Mahoney et al. (2011) included 43 patients with dementia, of which seven had tinnitus, as assessed with the Tinnitus Handicap Inventory. Grey matter volume was measured using VBM on 1.5T MRI scans, and results showed that patients with tinnitus had significantly reduced grey matter volume in the MGN compared to patients without tinnitus. This suggests that tinnitus is related to a reduction of structural volume in the MGN.

The results of the abovementioned studies point to a positive relationship between grey matter volume and cognitive functions. However, the studies examining PD patients provide evidence for this association only for cortical areas (Gerrits et al., 2014; Golden et al., 2015; Ibarretxe-Bilbao et al., 2011). The study by Mahoney et al. (2011) provides evidence for a reduction of structural volume in the MGN in people with tinnitus, but it does not examine PD patients. Moreover, the brain scans used to examine the relationship between structural volume in the MGN and tinnitus were obtained with a 1.5T MRI, thus resulting in reduced spatial resolution and contrast, as compared to 7T MRI (Isaacs et al., 2020). This is critically important when it comes to tiny subcortical structures like the MGN, thus creating the need for further investigation using more sensitive instruments.

It becomes apparent that the MGN has not been investigated in relation to Parkinson's and that there is need to investigate the MGN with highly sensitive instruments. Therefore, the research question of this study is whether PD patients show a reduction in structural volume of the MGN as captured with 7T MRI. Based on the current literature, it is hypothesized that there will be a decrease of structural volume in the MGN in people with PD as compared to healthy controls. An additional goal of this study is the production of MGN delineations that can be used as training

data for the MASSP algorithm, expanding its capabilities to include the MGN and allowing for its extension on PD patients.

Method

Participants

Each group, PD patients and healthy controls, consisted of 11 subjects. Their age and sex distributions are shown in tables 1a and 1b.

Table 1a

Age and Sex Distribution of PD Patients

Age (years)	Female	Male	Total
49-60	6	0	6
61-73	0	5	5
Total	6	5	11

Table 1b

Age and Sex Distribution of AHEAD Participants

Age (years)	Female	Male	Total
18-30	1	1	2
31-40	1	1	2
41-50	1	1	2
51-60	1	0	1
61-70	2	1	3
71-80	0	1	1
Total	6	5	11

For the healthy control group, 11 participants were collected from the Amsterdam ultra-high field adult lifespan database (AHEAD), made up of 105 subjects covering the entire adult lifespan (Alkemade et al., 2020). The selection of the 11 participants was based on the training data that MASSP needs, that is why their age distribution does not match that of the PD patients.

Materials

The scans that were reused in this study were originally acquired using a 7 Tesla MRI scanner, tailored to producing images of the human subcortex (Alkemade et al., 2020). It was a Philips Achieva scanner with a 32-channel head array coil, located at the Spinoza centre for Neuroimaging in Amsterdam. A multi-echo magnetization-prepared rapid gradient echo (MP2RAGEME) sequence was used to obtain T1-weighted and T2* contrasts (Caan et al., 2019), allowing the computation of contrasts from a single scan with an isotropic resolution of 0.7 mm and a reconstructed voxel size of $0.64 \times 0.64 \times 0.7 = 0.287$ mm^3. MP2RAGEME consists of two rapid gradient echo images, acquired in sagittal plane after a 180° inversion pulse and excitation pulses with inversion times of 670 ms and 3675.4 ms (Alkemade et al., 2020).

To ensure data privacy, numbers were assigned to participants and all skull information was removed. This resulted in a separate image using a brain extraction tool (BET), creating a

binary mask to remove skull information. Furthermore, skull stripping resulted in the removal of identifying facial features, like the ears and nose (Alkemade et al., 2020).

Moreover, for the estimation of structural volume of the MGN, manual parcellations, also called delineations, were performed using the FSLeyes software. The contrast used was the R1, the inverse of the T1-weighted contrast, as this is recommended for parcellating several subcortical structures (Alkemade et al., 2021). To minimize bias, the parcellations were performed under standardized conditions, while randomizing the order of the participants' scans and the starting hemisphere. Particularly, the parcellations were done during daylight, in the same individual space and with a display brightness of 90%. These parcellations were used to estimate the structural volume of the MGN using the "nighres" library (Huntenburg et al., 2018) with a Python script made by senior staff members of the University of Amsterdam.

Procedure

Before the commencement of the study, the research received approval from the ethical review board of the University of Amsterdam. This allowed the collection of brain scans for the AHEAD participants (Alkemade et al., 2020). For the recruitment of PD patients, the research obtained the approval from the medical ethical review board of the Amsterdam Medical Centre due to sensitivity with patients' medical records.

The healthy control participants were recruited through the University of Amsterdam, newsletters, and the social media of the Dutch Parkinson Foundation. Participants were between 18 and 80 years and did not have any medical issues, as assessed through a self-reported health statement before the study. The PD patients were recruited through a newsletter, the Dutch Parkinson Foundation, and through a presentation at a yearly meeting, discussing the purpose of the study. Participants were excluded if they had not signed the informed consent before the beginning of the study or if there were possible interfering factors, like pacemakers, pregnancy, or claustrophobia. Participant eligibility was assessed with a questionnaire, and for the healthy control participants it was aimed to have at least 6 male and 6 female participants per age-group. Participants were fully informed about the purpose of the study in the informed consent form, so there was no need for debriefing. After the scanning, participants received 20 euros each for their participation and a compensation for travel expenses if there were any.

Analysis

As mentioned previously, volumetric data about the MGN of 11 PD patients were compared to 11 healthy controls. Initially, the intention was to match the participants of the two groups by age and sex, resulting in a matched-pairs design. This proved inefficient because the age distribution of the two groups was dissimilar, resulting in at least five pairs of an age difference greater than 10 years. This exceeds the recommended age limit when matching for age in case-control studies (Mansournia et al., 2018). Therefore, the data were not matched for the analysis, resulting in a between-subjects design.

Moreover, even though aging has been shown to be associated with decreased brain volume (Scahill et al., 2003), age was not included as a covariate in this analysis. This is because the assumption that the groups of the independent variable should not differ on the covariate (Field, 2018) is violated. More specifically, healthy controls have a mean age of 47.9 with a minimum of 21, while PD patients have a mean age of 59.9, with a minimum of 49.

Therefore, the analysis that was run was a Bayesian Mann-Whitney U test, the non-parametric version of the independent samples t-test (Field, 2018). The non-parametric test was preferred over its parametric counterpart because distributional assumptions cannot be accurately tested with only 11 observations per group (Field, 2018). These assumptions need to be tested and met within the Bayesian framework as well (van Doorn et al., 2020), so a Bayesian independent samples t-test would not suffice. The Bayesian framework was preferred over the frequentist one, because it easily allows for a continuous updating of support for one hypothesis over the other (Etz et al., 2018). This is of critical importance when the sample size is so small and there might be additional data collection in the future.

For each group, a Wilcoxon signed-rank test, the non-parametric equivalent of the paired-samples t-test, was conducted on the volume of the two hemispheres to establish whether there is a lateralization effect. If that was the case, the Bayesian Mann-Whitney U test would be conducted twice, comparing the volume of each hemisphere between the two groups separately. This was not the case, so the Bayesian Mann-Whitney U test was conducted once, on the mean of the two hemispheres. Since the alternative hypothesis is one-sided, the prior distribution is truncated at 0 to only allow for positive effect size values. It is defined with a Cauchy distribution, centred at 0, with a scale of $r = 1/\sqrt{2}$, according to the guidelines provided by van Doorn et al. (2021).

All data processing and the analysis was done in Python3 (van Rossum & Drake, 2009). The code for the Bayesian Mann-Whitney U test was based on the algorithm described by van Doorn et al. (2020) on Bayesian rank-based hypothesis testing, as well as on the source code of JASP, available on its GitHub page (JASP Team, 2022). In the spirit of open science, all Python scripts and datasets used are available on my GitHub page[1] for reproducibility purposes.

Results

Firstly, the results of the Wilcoxon signed-rank tests conducted on the MGN volume of the two hemispheres are presented in table 2. For healthy controls, the difference between the left (M = 122.81, SD = 37.43) and right (M = 120.76, SD = 37.83) hemispheres was not statistically significant, W = 30, p = 0.83. Similarly, for PD patients, the difference between the left (M = 99.99, SD = 36.16) and the right hemispheres (M = 105.94, SD = 32.54) was not statistically significant, W = 27, p = 0.64. This allowed to conduct one Bayesian Mann-Whitney U test on the mean of the two hemispheres.

Table 2

Results of the Wilcoxon Signed-rank Test, Comparing the Left and Right Hemispheric MGN Volume, Measured in Cubic Millimeters (mm^3) for Each Group

Group	Left mean	Left sd	Right mean	Right sd	W	p-value
HC	122.81	37.43	120.76	37.83	30	0.83
PD	99.99	36.16	105.94	32.54	27	0.64

Note. HC refers to healthy controls, PD to Parkinson's disease patients, and sd to the standard deviation.

The results of the Bayesian Mann-Whitney U test are presented in table 3. On average, the MGN volume of the healthy controls (M = 121.78, SD = 34.70) was higher compared to that of the PD patients (M = 102.97, SD = 29.00). The Bayesian Mann-Whitney U test, examining this difference between the two groups, suggests that the data are 1.22 times more likely under the hypothesis that PD patients have a decreased MGN volume in relation to healthy controls, as compared to the hypothesis that there is no difference, BF_{10} = 1.22. This evidence is hardly supportive of any hypothesis over the other, so the data are equally likely under either hypothesis. Moreover, the posterior distribution of δ, the standardized effect size for the difference between the sum ranks of the groups, has a median of 0.46 and the 95% credible interval ranges from 0.03

[1] https://github.com/OdysseasPapakyriakou/BScThesisAnalysis

to 1.16. Therefore, assuming that the effect exists, there is 95% certainty that the true value of δ lies between 0.03 and 1.16.

Table 3

Results of the Bayesian Mann-Whitney U Test, Comparing the Mean Hemispheric MGN Volume of the two Groups, Measured in Cubic Millimeters (mm^3)

BF_{10}	Posterior median	95% CRI	mean HC	sd HC	mean PD	sd PD
1.22	0.46	[0.03, 1.16]	121.78	34.70	102.97	29.00

Note. BF_{10} refers to the Bayes Factor in support of the alternative hypothesis, CRI to the credible interval, HC to the healthy controls, PD to Parkinson's disease patients, and SD to the standard deviation.

Figure 1a shows a boxplot of the hemispheric means of the MGN per group, along with the observations. To explore the relationship between age and MGN volume per group, scatter plots and correlation coefficients were produced. The results, presented in figure 1b, suggest a moderate positive relationship between age and MGN volume for the healthy controls, r = 0.59, while it points towards a very weak negative relationship between age and MGN volume for the PD patients, r = -0.15.

Figure 1a

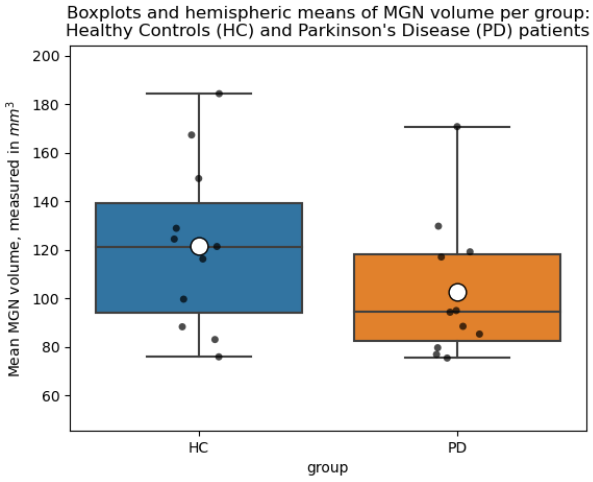
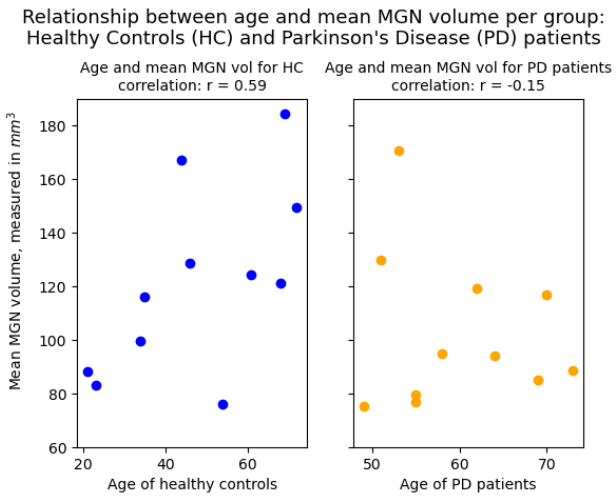


Figure 1b



For confirmation, the Bayesian Mann-Whitney U test was also performed in JASP, with BF_{10} = 1.14. The median of the posterior distribution of δ is 0.44 and the 95% credible interval ranges from 0.03 to 1.19. This small difference between the results is because the algorithm used for the analysis, in JASP and in Python, performs many simulations, so it does not give the same

outcome every time. However, since it uses 5 chains of 1000 iterations each, the results should converge around the same number across different implementations, which is the case here.

Discussion

Interpretation of the results

The aim of this research was to examine whether people with Parkinson's show a reduction in the structural volume of the MGN or not. The results suggest that it is not clear whether PD patients show a decrease of structural volume in the MGN. This is because the Bayes Factor is very close to one, so it does not provide supportive or contradictive evidence for either hypothesis over the other (Wagenmakers et al., 2018). Within the Bayesian framework, it is common to use the findings of previous research to make a more informed decision about the prior in future research. Since the Bayes Factor in this study is very close to 1, future researchers should not use the posterior median for the effect size to inform their prior (Wagenmakers et al., 2018), especially since the wide range of the 95% credible interval reflects high uncertainty. If the Bayes Factor was smaller than 1/3 or larger than 3, then there would be supportive evidence for the found effect size and for supporting one hypothesis over the other.

Moreover, the exploratory analysis on the relationship between age and MGN volume suggests that volume is positively related to age across the adult lifespan for healthy controls, while it is weakly negatively related to age for PD patients. Of course, one should consider that the age of the healthy controls starts from 21 years of age, while that of the PD patients starts from 49 years of age, so for a clearer conclusion one would estimate the correlation between age and MGN volume for healthy controls with an age similar to that of the PD patients. However, this change of direction in the relationship is an important observation for future research because it suggests that age may not be included as a covariate, since the assumption of homogeneity of regression slopes (Field, 2018) is violated. Instead, the optimal way to address the influence of age would be to design a study where participants can be matched for age.

The fact that participants were not matched for age in this study has important implications for the interpretation of the result. Since age seems to be negatively correlated with MGN volume in PD patients, in case of a higher BF_{10}, one could not conclude whether the effect was because of PD or because of age. That is why future research in this area should strive to match participants by age. In this study this was not done because the healthy control participants were selected based on the needs of the MASSP algorithm (Bazin et al., 2020).

Observing the boxplot and scatterplot of the PD patients in figures 1a and 1b one might wonder about the presence of an outlier, namely, the observation that seems to have an MGN volume of approximately 170. The z-score of this observation is 2.45, which is not extremely large (Field, 2018), so it might reflect a real variability in the population. Additionally, there are very few observations to judge whether this value deviates from the rest, and the non-parametric tests are very robust to extreme values (Field, 2018). For these reasons the observation was not perceived as an outlier, and it was retained in the analysis.

Connection to previous research

Although the previous research presented above has provided indirect support for a reduction of MGN volume in PD patients, the results of this study neither support nor contradict previous research. This claim is only possible within the Bayesian framework, since within the frequentist framework it is not possible to know whether a statistically not significant p-value means support for the null hypothesis or low statistical power (Wagenmakers et al., 2018). On the contrary, a Bayes Factor of around 1 suggests insensitive data (Wagenmakers et al., 2018), which could be due to the very small sample size, the high variability in the data, as shown via the standard deviations of each group, or because participants were not matched for age and sex, and of course because of any combination of these reasons.

The hypothesis that PD patients may show a decrease in MGN volume was based on literature suggesting that PD patients show a decrease in cortical volume (Gerrits et al., 2014; Golden et al., 2015; Ibarretxe-Bilbao et al., 2011), that tinnitus patients show a decrease in MGN volume (Mahoney et al., 2011) and that PD patients often suffer from tinnitus (Uluyol et al., 2016). However, it is not known whether the PD patients in this study also experienced tinnitus, while their cortical volume was not measured at all. Therefore, the fact that this study yielded an inconclusive result does not cast doubt on the findings of previous research, but it accentuates the need for further research on the topic.

A simulation to explore the role of sample size

Notably, all previous studies examined, included at least 20 participants in each group, while this study included 11, possibly not allowing to detect the difference in MGN volume, if it exists. To further explore the significance of sample size, MGN volume observations were simulated based on the sample mean and standard deviation of each group assuming a normal distribution for 11, 15, 20, 30, 50, and 75 participants per group. Even though in most cases the

sample size was enough to run parametric statistics, the Bayesian non-parametric Mann-Whitney U test was conducted again on this simulated data for a more direct comparison, with the same hypotheses. The results are presented in table 4.

Interestingly, with 11 simulated observations per group it is possible to obtain a result in the other direction, slightly supporting the hypothesis that PD is not related to decreased MGN volume. Therefore, one cannot really know whether the data in this study reflect a real effect in the population, suggesting that the distributions of the two populations may be similar. However, assuming that the data obtained are representative of the population, one would need at least 30 participants per group to support the hypothesis that PD patients show a decrease in MGN volume with a Bayes Factor larger than 3, which is the minimum criterion to support a hypothesis (Wagenmakers et al., 2018).

Table 4

Results of the Bayesian Mann-Whitney U Test on Simulated Data, Comparing the Mean Hemispheric MGN Volume of the two Groups, Measured in Cubic Millimeters (mm^3)

n	BF_{10}	posterior median	95% CRI	mean HC	sd HC	mean PD	sd PD
11	0.73	0.35	[0.02, 1.01]	116.60	35.68	102.99	28.32
15	2.41	0.56	[0.05, 1.25]	126.69	35.28	101.87	24.40
20	1.60	0.46	[0.04, 1.02]	120.72	31.50	105.19	29.12
30	16.10	0.67	[0.19, 1.19]	127.07	25.70	105.66	28.26
50	21.92	0.58	[0.17, 0.95]	119.10	32.69	98.92	27.40
75	76.05	0.58	[0.26, 0.91]	120.28	29.81	99.47	24.09

Note. n refers to the sample size per group, BF_{10} to the Bayes Factor in support of the alternative hypothesis, CRI to the credible interval, HC to healthy controls, PD to Parkinson's disease patients, and sd to the standard deviation.

Limitations and suggestions for future research

The study has several limitations. Firstly, as mentioned previously, the research design did not allow to consider the influence of age and sex on MGN volume by matching participants by these factors. The most important implication of this is that if the study found enough evidence to support the reduction of MGN volume in PD patients, it would still be unclear whether this evidence resulted from age or from Parkinson's, or from sex. Therefore, future research should focus on matching participants by age and sex. For this study, the scans from both groups had already been obtained, so it would have been possible to delineate brain images of participants that

could have been matched by age and sex. However, this was intentionally not done because priority was given to the training and validation needs of the MASSP algorithm.

As becomes even more apparent from the simulation, another important limitation of this study is its small sample size. The main reason for this is that there were only 11 brain scans available from people with Parkinson's. But even if there were more, it would be unfeasible to manually delineate so many brain scans within the three months in which this thesis had to be completed. According to the simulation, future research should focus on collecting at least 30 participants per group, so that a more conclusive result can be obtained.

Despite the convenience of having already obtained brain scans, this posed a few restrictions on the study. Specifically, it was impossible to know whether the PD patients suffered from tinnitus since no such data had been collected. This would have been critical information, since one of the reasons to hypothesize a reduction in MGN volume was the high prevalence of tinnitus in PD patients (Uluyol et al., 2016) and the link between tinnitus and decreased MGN volume (Mahoney et al., 2011). Moreover, there were several movement artifacts in the brain scans, for which there was no correction. Optimally, such artifacts need to be corrected for, so that the scans provide a more accurate depiction of the relevant structure. Importantly, however, there seemed to be no significant artifacts around the MGN area. Another point is that the brain scans for the healthy controls and the PD patients were collected at different times, which may have resulted in slight differences in the acquisition protocols (Alkemade et al., 2017, 2020). Additionally, PD patients may also take several other drugs or suffer from other disorders, which could potentially affect MGN volume. Although these factors may have affected the quality of the brain scans, and should be considered in future research, there is no reason to suspect that they may have biased the results of this study in a systematic way.

Another issue that may have affected the quality of the data is the accuracy of the delineations. This can be estimated by comparing the delineations used in this study with others' delineations on the same subjects and structure. This inter-rater accuracy can be conveyed with the Dice similarity coefficient, indicating the amount of overlap between the delineations in a range between zero and one (Alkemade et al., 2021). However, the Dice coefficient is largely affected by the size of the structure, so "the dilated Dice score is a potentially more representative measure for the agreement between raters regardless of size" (Alkemade et al., 2021, p. 38).

The dilated Dice score, between the delineations in this analysis and those of a more experienced rater, is available for eight of the eleven PD patients, but for none of the eleven healthy controls. Therefore, considering the delineations for each hemisphere separately, there is a total of 22*2 = 44 delineations, and there are dilated Dice scores for 8*2 = 16 of these. Scores higher than 0.75 are considered as an acceptable inter-rater agreement score (Alkemade et al., 2021), and there are two scores out of 16 below 0.75. Assuming that these 16 scores are representative of the total 44 scores, this suggests that 2/16 = 0.125, or 12.5%, of the delineations show a considerable difference between the delineations of a more experienced rater. This might indicate a data quality issue that may have affected the results, and it will be addressed in the future by having other raters delineate the same brain scans, as was done in the paper by Alkemade et al. (2021).

Strengths

Despite the limitations of this study, there are some strengths to be considered. Most importantly, it is the first study that examines how Parkinson's may relate to the volume of the MGN, thus providing future researchers with important information. For example, the median of the posterior distribution shows that if there is a reduction in MGN volume this is rather small, 0.46, with high uncertainty about it 95% CRI = [0.03, 1.16], and that a study of similar design would need at least 30 subjects per group to provide strong evidence for its existence.

Additionally, the delineations for this study will be funnelled into the MASSP algorithm, thus reliably extending its use to the MGN and to PD patients. This means that MASSP could reliably delineate other brain scans by convenience, thus generating delineations for healthy controls that can be matched to the PD patients by age and sex. This is a highly relevant implication of this study since it allows for an improved research design with more reliable results. Lastly, MASSP's extension to a subcortical structure, like the MGN, could potentially enhance the design of human brain atlases by including more subcortical structures.

Conclusion

Conclusively, the result of this study does not allow to answer whether PD patients show a decrease in MGN volume, thus accentuating the need for further research on the topic. It also provides evidence for different kinds of relationships between age and MGN volume in the two groups, thus suggesting that age might not meet the assumptions of a covariate. Moreover, the extension of the MASSP algorithm to the MGN will allow future researchers to delineate this brain area reliably and automatically, which can have important academic and societal implications.

Firstly, it would be easier to conduct research on the relationship between PD and MGN volume, thus elucidating this unexplored area of study, while the use of MASSP on the MGN could result in its inclusion in human brain atlases. This is very important, since 93% of the human subcortex remains unrepresented in brain atlases, even though it plays a role in normal functioning and in the treatment of several disorders, for example with deep brain stimulation (Alkemade et al., 2021). From a societal standpoint, facilitating research about the relation between PD and MGN volume is important for the prognosis of the disease. For example, if the MGN volume is indeed reduced, knowing this could help doctors and patients in treating symptoms related to audition. Therefore, despite its inconclusive result, this research provides significant insights in the academic community with the potential of having critical societal relevance.

References

Alkemade, A., de Hollander, G., Keuken, M. C., Schäfer, A., Ott, D. V. M., Schwarz, J., Weise, D., Kotz, S. A., & Forstmann, B. U. (2017). Comparison of T2∗-weighted and QSM contrasts in Parkinson's disease to visualize the STN with MRI. *PLoS ONE*, *12*(4). https://doi.org/10.1371/journal.pone.0176130

Alkemade, A., Mulder, M. J., Groot, J. M., Isaacs, B. R., van Berendonk, N., Lute, N., Isherwood, S. J., Bazin, P. L., & Forstmann, B. U. (2020). The Amsterdam Ultra-high field adult lifespan database (AHEAD): A freely available multimodal 7 Tesla submillimeter magnetic resonance imaging database. *NeuroImage*, *221*. https://doi.org/10.1016/j.neuroimage.2020.117200

Alkemade, A., Mulder, M. J., Trutti, A. C., & Forstmann, B. U. (2021). Manual delineation approaches for direct imaging of the subcortex. *Brain Structure and Function*. https://doi.org/10.1007/s00429-021-02400-x

Bazin, P. L., Alkemade, A., Mulder, M. J., Henry, A. G., & Forstmann, B. U. (2020). Multi-contrast anatomical subcortical structures parcellation. *ELife*, *9*, 1–23. https://doi.org/10.7554/ELIFE.59430

de Keyser, K., de Letter, M., Santens, P., Talsma, D., Botteldooren, D., & Bockstael, A. (2021). Neurophysiological investigation of auditory intensity dependence in patients with Parkinson's disease. *Journal of Neural Transmission*, *128*(3), 345–356. https://doi.org/10.1007/s00702-021-02305-3

Elgh, E., Domellöf, M., Linder, J., Edström, M., Stenlund, H., & Forsgren, L. (2009). Cognitive function in early Parkinson's disease: A population-based study. *European Journal of Neurology*, *16*(12), 1278–1284. https://doi.org/10.1111/j.1468-1331.2009.02707.x

Gerrits, N. J. H. M., van der Werf, Y. D., Hofman, M., Foncke, E. M. J., Klein, M., Berendse, H. W., & van den Heuvel, O. A. (2014). Gray matter differences contribute to variation in cognitive performance in Parkinson's disease. *European Journal of Neurology*, *21*(2), 245–252. https://doi.org/10.1111/ene.12269

Gerrits, N. J. H. M., van Loenhoud, A. C., van den Berg, S. F., Berendse, H. W., Foncke, E. M. J., Klein, M., Stoffers, D., van der Werf, Y. D., & van den Heuvel, O. A. (2016). Cortical thickness, surface area and subcortical volume differentially contribute to cognitive heterogeneity in Parkinson's disease. *PLoS ONE*, *11*(2). https://doi.org/10.1371/journal.pone.0148852

Golden, H. L., Nicholas, J. M., Yong, K. X. X., Downey, L. E., Schott, J. M., Mummery, C. J., Crutch, S. J., & Warren, J. D. (2015). Auditory spatial processing in Alzheimer's disease. *Brain*, *138*(1), 189–202. https://doi.org/10.1093/brain/awu337

JASP Team. (2022). JASP (Version 16.0)[Computer software]. https://github.com/jasp-stats/jaspTTests/blob/master/R/ttestbayesianindependentsamples.R

Huntenburg, J. M., Steele, C. J., & Bazin, P. L. (2018). Nighres: processing tools for high-resolution neuroimaging. *GigaScience*, *7*(7). https://doi.org/10.1093/gigascience/giy082

Ibarretxe-Bilbao, N., Junque, C., Marti, M. J., & Tolosa, E. (2011). Brain structural MRI correlates of cognitive dysfunctions in Parkinson's disease. *Journal of the Neurological Sciences*, *310*(1–2), 70–74. https://doi.org/10.1016/j.jns.2011.07.054

Isaacs, B. R., Mulder, M. J., Groot, J. M., van Berendonk, N., Lute, N., Bazin, P. L., Forstmann, B. U., & Alkemade, A. (2020). 3 versus 7 Tesla magnetic resonance imaging for parcellations of subcortical brain structures in clinical settings. *PLoS ONE*, *15*(11 November). https://doi.org/10.1371/journal.pone.0236208

Jafari, Z., Kolb, B. E., & Mohajerani, M. H. (2020). Auditory Dysfunction in Parkinson's Disease. In *Movement Disorders* (Vol. 35, Issue 4, pp. 537–550). John Wiley and Sons Inc. https://doi.org/10.1002/mds.28000

Kraff, O., Fischer, A., Nagel, A. M., Mönninghoff, C., & Ladd, M. E. (2015). MRI at 7 Tesla and above: Demonstrated and potential capabilities. In *Journal of Magnetic Resonance Imaging* (Vol. 41, Issue 1, pp. 13–33). John Wiley and Sons Inc. https://doi.org/10.1002/jmri.24573

Lee, E. Y., Sen, S., Eslinger, P. J., Wagner, D., Shaffer, M. L., Kong, L., Lewis, M. M., Du, G., & Huang, X. (2013). Early cortical gray matter loss and cognitive correlates in non-demented Parkinson's patients. *Parkinsonism and Related Disorders*, *19*(12), 1088–1093. https://doi.org/10.1016/j.parkreldis.2013.07.018

Mahoney, C. J., Rohrer, J. D., Goll, J. C., Fox, N. C., Rossor, M. N., & Warren, J. D. (2011). Structural neuroanatomy of tinnitus and hyperacusis in semantic dementia. *Journal of Neurology, Neurosurgery and Psychiatry*, *82*(11), 1274–1278. https://doi.org/10.1136/jnnp.2010.235473

Mansournia, M. A., Jewell, N. P., & Greenland, S. (2018). Case–control matching: effects, misconceptions, and recommendations. *European Journal of Epidemiology*, *33*(1), 5–14. https://doi.org/10.1007/s10654-017-0325-0

Miletić, S., Bazin, P.-L., Isherwood, S. J. S., Keuken, M. C., Alkemade, A., & Forstmann, B. U. (2022). Charting human subcortical maturation across the adult lifespan with in vivo 7 T MRI. *NeuroImage*, *249*, 118872. https://doi.org/10.1016/j.neuroimage.2022.118872

Pezzoli, S., Sánchez-Valle, R., Solanes, A., Kempton, M. J., Bandmann, O., Shin, J. il, Cagnin, A., Goldman, J. G., Merkitch, D., Firbank, M. J., Taylor, J. P., Pagonabarraga, J., Kulisevsky, J., Blanc, F., Verdolini, N., Venneri, A., & Radua, J. (2021). Neuroanatomical and cognitive correlates of visual hallucinations in Parkinson's disease and dementia with Lewy bodies: Voxel-based morphometry and neuropsychological meta-analysis. In *Neuroscience and Biobehavioral Reviews* (Vol. 128, pp. 367–382). Elsevier Ltd. https://doi.org/10.1016/j.neubiorev.2021.06.030

Raza, C., Anjum, R., & Shakeel, N. ul A. (2019). Parkinson's disease: Mechanisms, translational models and management strategies. In *Life Sciences* (Vol. 226, pp. 77–90). Elsevier Inc. https://doi.org/10.1016/j.lfs.2019.03.057

Scahill, R. I., Frost, C., Jenkins, R., Whitwell, J. L., Rossor, M. N., & Fox, N. C. (2003). A Longitudinal Study of Brain Volume Changes in Normal Aging Using Serial Registered Magnetic Resonance Imaging. In *Arch Neurol* (Vol. 60). https://jamanetwork.com/

Shetty, K., Krishnan, S., Thulaseedharan, J. V., Mohan, M., & Kishore, A. (2019). *Asymptomatic Hearing Impairment Frequently Occurs in Early-Onset Parkinson's Disease*. https://doi.org/10.14802/jmd.18048/J

Uluyol, S., Kilicaslan, S., Baydar, C., & Sener, U. (2016). Hearing impairment and tinnitus severity in Parkinson's disease (2016). *Turkish Journal of Geriatrics*, *19*(4), 228–331.

van Doorn, J., Ly, A., Marsman, M., & Wagenmakers, E. J. (2020). Bayesian rank-based hypothesis testing for the rank sum test, the signed rank test, and Spearman's ρ. *Journal of Applied Statistics*, *47*(16), 2984–3006. https://doi.org/10.1080/02664763.2019.1709053

van Doorn, J., van den Bergh, D., Böhm, U., Dablander, F., Derks, K., Draws, T., Etz, A., Evans, N. J., Gronau, Q. F., Haaf, J. M., Hinne, M., Kucharský, Š., Ly, A., Marsman, M., Matzke, D., Gupta, A. R. K. N., Sarafoglou, A., Stefan, A., Voelkel, J. G., & Wagenmakers, E. J. (2021). The JASP guidelines for conducting and reporting a Bayesian analysis. In *Psychonomic Bulletin and Review* (Vol. 28, Issue 3, pp. 813–826). Springer. https://doi.org/10.3758/s13423-020-01798-5

Van Rossum, G., & Drake, F. L. (2009). *Python 3 Reference Manual*. CreateSpace.

Wagenmakers, E. J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Selker, R., Gronau, Q. F., Dropmann, D., Boutin, B., Meerhoff, F., Knight, P., Raj, A., van Kesteren, E. J., van Doorn, J., Šmíra, M., Epskamp, S., Etz, A., Matzke, D., … Morey, R. D. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin and Review*, *25*(1), 58–76. https://doi.org/10.3758/s13423-017-1323-7