Optimizing thalamic segmentation and relating derived volumes to executive functioning in Parkinson´s disease

Díaz Castiella, María

Optimizing thalamic segmentation and relating derived volumes to executive functioning in Parkinson´s disease

María Díaz Castiella

Master Thesis Clinical Neuropsychology

Faculty of Behavioural and Social Sciences – Leiden University

July, 2021

Student number: s2931877

External Supervisor: Dr. Chris Vriend, Department of Anatomy and Neurosciences; Amsterdam UMC

First Examiner: Dr. Marit Ruitenberg, Health, Medical and Neuropsychology Unit; Leiden University

Acknowledgments

I wish to express my gratitude to the department of Anatomy and Neurosciences, section Neuropsychiatry of Amsterdam UMC, for providing me the opportunity of doing this thesis project. I would also like to express my most sincere appreciation to Dr. Chris Vriend and Dr. Marit Ruitenberg for their outstanding professionalism and their supervision and guidance throughout all the stages of this thesis project.

Abstract

Apart from its motor symptoms, Parkinson´s disease (PD) is characterized by cognitive deficits; among these, executive functioning (EF) deficits are common. Stroke studies have shown a role of the whole thalamus, and specifically the medial nuclear group (Mng), in EF. Although in PD studies have shown volume loss in the thalamus and Mng, little is known about the relationship between this phenomenon and EF deficits in PD. We aimed to address this gap and examine the relationship between thalamic and Mng volumes and EF in PD. As the thalamus is difficult to segment due to poor tissue contrast, we also aimed to explore whether using Phase-Sensitive Inversion Recovery MRI scans in addition to T1 MRI scans (T1+PSIR MRI) would optimize automatic thalamic segmentation over using T1 MRI scans alone. A total of 76 PD participants underwent T1 and PSIR MRI scans and completed tests measuring verbal fluency, resistance to cognitive interference and planning abilities, which were used to calculate an EF composite score. We found that using T1+PSIR MRI scans resulted in a more accurate thalamic segmentation as shown by less thalamic overestimation especially in the lateral and anteroventral nuclear groups, and smaller thalamic and nuclear groups volumes. We used the derived volumes to relate to global EF performance. Results showed that total thalamic and total Mng volumes were not significantly related to global EF performance (β = .08, t = 1.24, p = .22; β = .16, t = 1.71, p = .09, respectively). Results of exploratory analyses showed a positive significant relationship between left (β = .16, t = 2.26, p = .02) and right (β = .18, t = 2.67, p = .01) thalamic volumes and left (β = .25, t = 2.84, p = .01) Mng volume and verbal fluency performance. In this thesis project we demonstrated that automatic thalamic segmentation can be optimized by using PSIR MRI scans. Although the association between thalamic and Mng volumes and global EF in PD is not supported in this thesis project, volumes seem to be specifically associated with verbal fluency. We propose that this could be due to the role of the Mng in cognitive flexibility. Supported by future longitudinal studies, our findings could have implications for early diagnosis and effective management of verbal fluency deficits in PD.

Keywords: Parkinson´s disease, thalamus, medial, executive functioning, verbal fluency.

Layman abstract

Background: Apart from the well-known motor symptoms, many people with Parkinson´s disease (PD) experience problems with complex mental skills like self-control or planning. In PD, the volumes of some brain structures are smaller than in the general population. The thalamus is one of these brain structures. A smaller thalamic volume, specifically one part of the thalamus called the medial nuclear group (Mng), has been related to problems in complex mental skills in other disorders with motor symptoms. However, we do not know much about this relationship in PD. In addition, getting a good estimation of thalamic volumes is difficult because this structure is not very visible in a brain scan.

Objectives: One objective of this thesis project was to examine if thalamic and Mng volumes are related to global complex mental skills in PD. Another objective was to examine if we could get a better estimation of thalamic volumes by combining information from two different types of brain scans together instead of only one.

Results: The estimation of thalamic volumes was better combining information from two different types of brain scans together instead of only one. There was no significant relationship between thalamic and Mng volumes and global complex mental skills. When we explored this in more detail, we found that smaller thalamic volume, and smaller Mng volume only in the left side of the brain, were significantly related to worse verbal fluency.

Conclusion: We demonstrated that combining information from two different types of brain scans can help to get a better estimation of thalamic volumes. Although the association between thalamic and Mng volumes and global complex mental skills in PD is not supported in this thesis project, volumes seem to be specifically associated with verbal fluency. We propose that this is because the Mng has a role in mental flexibility, which is key for verbal fluency. Supported by future studies that examine the relationship between thalamic and Mng volumes and verbal fluency in PD over time, our findings could have implications for early diagnosis and effective management of problems in verbal fluency in PD.

Introduction

Parkinson´s disease (PD) is the second most prevalent neurodegenerative disease worldwide and it is characterized by motor symptoms including resting tremor, bradykinesia and postural instability (Jankovic, 2008). Non-motor symptoms including sensory, automatic, psychiatric, and cognitive deficits are also common and affect functioning and quality of life (Schapira et al., 2017). Cognitive deficits are present in 15-43% newly diagnosed patients with more than 40% developing mild cognitive impairment within five years after diagnosis and up to 80% of these developing dementia (Aarsland et al., 2003; Pedersen et al., 2017). Cognitive deficits in PD involve problems in the domains of executive functioning (EF), memory, visuospatial functioning, and processing speed (Ding et al., 2015). These deficits are heterogenous and can be classified in two distinct cognitive profiles: one characterized by primary memory, visuospatial and semantic deficits, and another by EF deficits, which are most reported (Sollinger et al., 2010; van Balkom et al., 2016). EF deficits in PD have extensively been recognized as frontal-type as they resemble the ones seen in patients with frontal lobe damage and are consistent with the disruption of the cortico-striatal-thalamo-cortical circuits in PD (Owen, 2004). EF deficits involve problems in cognitive flexibility, verbal fluency, inhibition, working memory, and planning abilities, and are related to reduced instrumental activities, social participation and quality of life in PD (Foster et al., 2011; Smith et al., 2020). Indeed, deficits in verbal fluency and resistance to cognitive interference have been identified as risk factors for developing dementia in PD (Lee et al., 2014).

Studies show that EF deficits in PD are related to the disruption of the cortico-striatal-thalamo-cortical circuits due to dopamine denervation in the caudate nucleus (Chung et al., 2018). Other mechanisms contribute to these deficits and disruption of these circuits, such as decreased functional connectivity, neuropathology, white matter abnormalities, and cortical and subcortical atrophy (Fang et al., 2020; Foo et al., 2017; Song et al., 2011). Specifically, caudate atrophy has been related to the disruption of the cortico-striatal-thalamo-cortical circuits due to its connections with the prefrontal cortex (Grahn et al., 2008). In addition, thalamic atrophy could contribute to the disruption of these circuits through its input to and from the striatum and projections to the prefrontal cortex (Owens-Walton et al., 2021). As thalamic atrophy in PD has been shown to predict cognitive deficits even before these are evident in neuropsychological test performances, neuroimaging could be promising as an assisting tool for identifying at risk patients, and consequently, facilitate early treatment implementation (Foo et al., 2017; Lanskey et al., 2018; Vasconcellos et al., 2018).

This thesis project focuses on the association between the thalamus and EF in PD. Although the thalamus was initially recognised as a relay station, it is now known for its modulatory influence and involvement in higher-order cognitive processes (Varela, 2014). Studies in other neurodegenerative movement disorders such as in Huntington´s disease have shown that atrophy in the *whole* thalamus bilaterally is related to deficits in verbal fluency, inhibition, cognitive flexibility, working memory, and planning abilities (Furlong et al., 2020; Kassubek et al., 2005). In particular,

the *medial nuclear group (Mng),* which is formed by the mediodorsal, intralaminar and midline thalamic nuclei, has been implicated in these EF subdomains due to its strong connections with the prefrontal cortex and the striatum (van der Werf et al., 2002; van der Werf, Scheltens et al., 2003). Animal lesion studies have shown a primarily role of the Mng in cognitive flexibility, although in stroke studies the Mng has also been involved in verbal fluency, inhibition and planning abilities (Edelstyn et al., 2014; Saalmann, 2014; van der Werf, Scheltens et al., 2003). Apart from cognitive flexibility, within the Mng, the mediodorsal nucleus has been involved in working memory, and the intralaminar and midline nuclei in processes of arousal and awareness including attention (Parnaudeau et al., 2018; van der Werf et al., 2002). This suggests that the integrity of the Mng might be key for EF (van der Werf, Scheltens et al., 2003). Post-mortem and animal models of PD have shown a significant atrophy in the intralaminar and midline nuclei, and some in-vivo studies have shown atrophy in the mediodorsal nucleus (Chen et al., 2020; Halliday, 2009; Villalba et al., 2014). In PD, whole thalamic atrophy has been related to verbal fluency deficits, but although some researchers have suggested that specifically the atrophy in the Mng might be related to EF deficits in PD, this relationship remains unexplored (Galvan & Smith, 2011; Gerrits et al., 2016; Ibarretxe-Bilbao et al., 2010).

The primary objective of this thesis project was to examine whether thalamic and Mng volumes are associated with global EF in PD. We hypothesized that a smaller total thalamic volume was related to worse global EF performance in PD. Secondly, we hypothesized that a smaller total Mng volume was related to worse global EF performance in PD. Global EF performance was operationalized by an EF composite score including EF measures of verbal fluency, resistance to cognitive interference and planning abilities. We additionally performed exploratory analyses to examine the specific relationship between left or right thalamic and left or right Mng volumes and performance in each of the individual EF measures included in the EF composite score. Findings could have implications for people with PD and professionals as it could provide with a better understanding of the neural underpinnings of EF deficits in PD. Followed by future studies it could potentially facilitate early diagnosis, monitoring progression and early treatment implementation. As the thalamus is difficult to segment due to poor tissue contrast of its white and grey matter content and the surrounding white matter, a secondary objective of this thesis project was to optimize thalamic segmentation. Phase-Sensitive Inversion Recovery (PSIR) MRI scans have a T1-weighted inversion recovery sequence with phase-sensitive reconstruction by maintaining tissue magnetization polarity. By maintaining the polarity of the signal intensity, PSIR MRI scans can provide a better signal to noise ratio and superior grey-white matter contrast (Hou et al., 2005). In this thesis project we explored whether automatic thalamic segmentation could be optimized by using PSIR MRI scans in addition to MP-RAGE T1-weighted (T1+PSIR) MRI scans compared to using MP-RAGE T1-weighted (T1) MRI scans alone.

Methods

Design

This thesis project is part of a study called COGTIPS (van Balkom et al., 2019), a double-blind randomized active controlled trial that examines the effects of an online cognitive training on EF and other cognitive deficits in PD. This thesis project used a cross-sectional design: the pre-training assessment of the COGTIPS study. The independent variables were total thalamic or total Mng volumes and the dependent variable was an EF composite score.

Participants

In the COGTIPS study participants were enrolled based on their interest and they were selected based on the presence of mild subjective cognitive complaints measured by the Parkinson´s Disease Cognitive Functional Rating Scale (Kulisevsky et al., 2013) and mild to moderate disease stage measured by the modified Hoehn & Yahr stage (Hoehn & Yahr, 1998). An overview of inclusion and exclusion criteria is shown in Appendix A. Those who completed neuroimaging assessment pre-training in the COGTIPS study were included in this thesis project. A total of 85 participants, 33 males and 52 females, ranging in age from 44 to 80 years, were initially included in this thesis project. Data exclusions and demographics and clinical characteristics of participants are presented in the results section.

Measures

Age, sex, disease duration, and medication use were obtained through interviews and questionnaires, and education level was measured by the Dutch Verhage Classification (Verhage, 1964). Global cognitive deficits were screened by the Montreal Cognitive Assessment (Nasreddine et al., 2005) and participants were classified into cognitive intact, level II single or multiple domain cognitive impairment, and probable dementia using a neuropsychological battery (see Appendix B) and considering recent diagnostic criteria (Emre et al., 2007; Litvan et al., 2012). Motor severity was obtained by the Unified Parkinson´s Disease Rating Scale-Part III (Fahn & Elton, 1987) and disease stage by the modified Hoehn & Yahr stage. Psychiatric comorbidities including depression, anxiety, impulsivity, and apathy were assessed by the Beck Depression Inventory (Beck et al., 1961), the Parkinson´s Anxiety Scale (Leentjens et al., 2014), the Questionnaire for Impulsive-Compulsive Disorders in Parkinson´s Disease (Weintraub et al., 2012), and the Apathy Scale (Starkstein et al., 1992), respectively. All these measures have been validated in PD and have moderate to high validity and reliability (Gill et al., 2008; Visser et al., 2006).

Structural imaging

Image acquisition. The MRI scans were acquired on a Discovery MR750 3.0T MRI scanner (General Electric, Milwaukee) with a 32-channel head coil at the Amsterdam University Medical Center (Amsterdam UMC), location Vrije Universiteit (VU), Amsterdam, The Netherlands. The scans were 3D sagittal MP-RAGE T1-weighted sequence according to ADNI-3 protocol with the following parameters: TI = 900 ms, TE = min full echo, flip angle = 8º, 168 slices (1 x 1 x 1 mm, matrix size

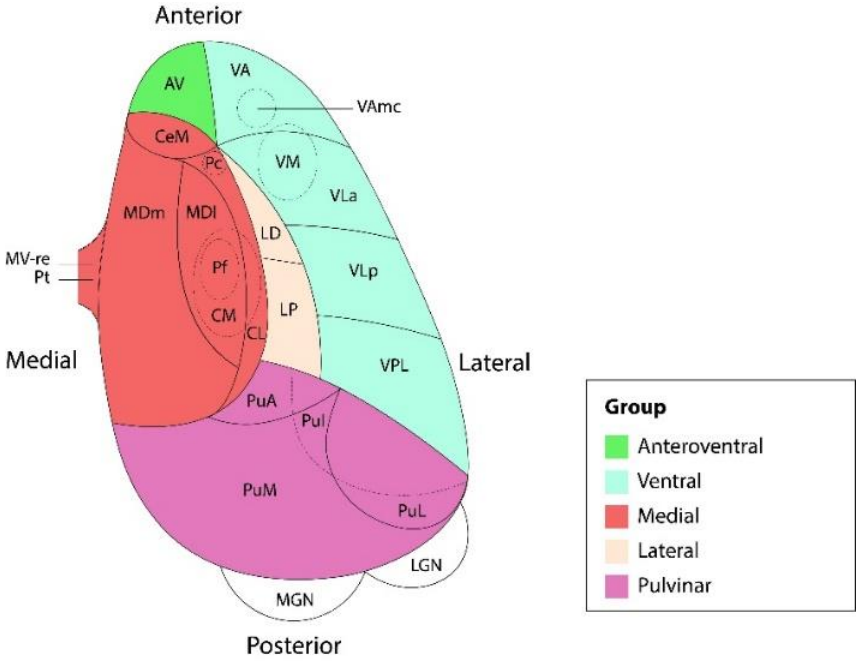
256 x 256) and 3D Cube sagittal PSIR with the following parameters: TI = 650 ms, TR = 3000 ms, TE = minimum, 168 slices (1 x 1 x 1 mm, matrix size 256 x 256).

Image pre-processing. We used FreeSurfer 7.1.1 software package (Fischl, 2012) for image pre-processing. We examined the T1 and PSIR MRI scans for errors in scanning and motion artifacts. Cortical reconstruction processes of T1 MRI scans were done by the standard "recon-all" script which includes removal of non-brain tissue, normalization, Talairach transformation, registration to a Gaussian classifier array, and white matter segmentation. PSIR MRI scans were skull and neck stripped and registered to the T1 MRI scans using an in-house pipeline. We used an algorithm based on a probabilistic atlas by Iglesias et al. (2018) (https://FreeSurfer.net/fswiki/ThalamicNuclei) for automatic thalamic segmentation. We performed thalamic segmentation using two different approaches: one using T1 MRI scans alone and the other using T1+PSIR MRI scans. As stated in Iglesias et al. (2018) "first the atlas is spatially warped following a deformation model. Second, a segmentation is drawn for each voxel independently, following the categorical distribution specified by the (deformed) atlas at each location. And third, image intensities are drawn independently at each voxel, as independent samples of Gaussian mixture models conditioned on the underlying segmentation" (p. 320). This algorithm results in left and right whole thalamic volumes and a parcellation of the thalamus into 25 different subnuclei volumes (mm^3). We computed the volumes of five left and right nuclear groups by summing subnuclei volumes (see Figure 1): The ventral nuclear group volume was a sum of the ventral anterior, ventral anterior magnocellular, ventral lateral anterior and posterior, ventral posterolateral, and ventromedial subnuclei volumes; The medial nuclear group volume was a sum of the central medial, central lateral, paracentral, centromedian, parafascicular, paratenial, medial ventral and mediodorsal medial magnocellular, and mediodorsal lateral parvocellular subnuclei volumes; The lateral nuclear group volume was a sum of the laterodorsal and lateral posterior subnuclei volumes. Lastly, the pulvinar nuclear group volume was a sum of the lateral and medial geniculate, limitans and anterior, medial, lateral, and inferior pulvinar subnuclei volumes. The anteroventral nuclear group volume was formed by the anteroventral subnucleus volume alone. We summed left and right thalamic and nuclear groups volumes to obtain the total volume of the whole thalamus and each nuclear group, respectively. We obtained cerebrospinal fluid volume using the FAST tool in FSL (Jenkinson et al., 2012; Zhang et al., 2011). To compare the quality of thalamic segmentation using the two different approaches, we multiplied cerebrospinal fluid and white matter volumes (mm^3) by the total thalamic volume obtained from the two different approaches to obtain the amount of cerebrospinal fluid and white matter overlap of the resulting total thalamic segmentations. We also obtained the number of outliers of left and right volumes of each nuclear group and summed them to obtain the total number of outliers of the resulting thalamic segmentation of each approach. For examining the relationship between thalamic and Mng volumes and EF performance, volumes were operationalized by thalamic segmentation using T1+PSIR MRI

scans, with lower values as measured in mm^3 representing lower volumes. Intracranial volume (ICV) in mm^3 was also obtained to control for inter-individual differences in brain size.

Figure 1

Thalamic nuclear groups.



Note. Figure by Weeland et al. (2021).

Cognitive measures

Global EF performance was operationalized by an EF composite score. This was computed by an average of *z*-scores, obtained by comparison to a norm group, of different EF test performances. These included the total number of correct responses of the Controlled Oral Word Association test (Schmand et al., 2008), the interference score of the Stroop Colour-Word test (Hammes et al., 1971), and the total accuracy score of the Tower of London test (Trujillo et al., 2015). The interpretation of raw scores of these measures is presented in Appendix B. Lower *z*-scores in the EF composite score represent worse global EF performance.

Controlled Oral Word Association test (COWAT; Schmand et al., 2008). This test consists of three different letters: D, A and T, and for each letter participants should name as many words as possible in one minute, except for names of people or places. The total number of correct responses was used as an indicator of verbal fluency. Lower *z*-scores represent worse performance. The internal consistency of this test is .82 and the convergent validity for EF subdomains of shifting and updating factors, working memory, and fluid reasoning is .23, .18 and .25, respectively (Aita et al., 2019; Schmand et al., 2008).

Stroop Colour-Word test (Stroop; Hammes et al., 1971). This test consists of three cards with 100 items in rows from left to right: "card I" which has words printed in black with names of the colours red, green, yellow, and blue; "card II" with different patches with colours red, green, yellow, and blue; "card III" in which the words with names of the colours are printed in an ink of a conflicting colour. On card I and card II, participants should read the words or the colours as quickly as possible and the scores are the total time in seconds that the participant takes to name all the words or colours. In card III, participants should name the colours of the ink while ignoring the conflicting printed names. The score in card III, known as the interference score, was obtained by subtracting the time in seconds that the participant took to read this card from what it took to read card II. The interference score was used as an indicator of resistance to cognitive interference. Higher *z*-scores represent worse performance, but for the inclusion in the EF composite score¸ the *z*-scores were inverted, with lower *z*-scores representing worse performance. The internal consistency of this test is .93, construct validity is adequate and convergent validity for EF subdomains of working memory and conflict monitoring is -.24 and -.27, respectively (Periáñez et al., 2021; Wöstmann et al., 2013).

Tower of London test (ToL; Trujillo et al., 2015). A computerized version of the ToL test was used, which consists of three pegs of different height and three beads with different colours: yellow, blue and red. A goal configuration and a starting configuration are shown simultaneously on the screen and participants should select the minimum number of moves needed to transform the starting configuration to meet the goal, moving only one bead at a time and only when there is no other bead on top. There are five planning conditions with different degree of difficulty with moves ranging from one to five. In each condition, after nine baseline trials with feedback, there are 100 pseudo-randomized trials with no feedback and a maximum response time of 45 seconds. An average of the total number of solved correct trials of the five planning conditions was used as a total accuracy score and as an indicator of accuracy of planning. Lower *z*-scores represent worse performance. There are no psychometric properties available for this specific version, but similar computerized versions had an internal consistency of .75, adequate construct validity and convergent validity for EF subdomains of working memory and fluid reasoning of .33 and .38, respectively (Debelak et al., 2016; Unterrainer et al., 2004).

Procedure

The COGTIPS study was performed at Amsterdam University Medical Center (Amsterdam UMC) location Vrije Universiteit (VU). Participants signed an informed consent to undergo a pre-screening and were asked if they wanted to partake in a subgroup that was going to undergo neuroimaging; if so, they were screened for contraindications. Those who were included after positive pre-screening signed an informed consent for inclusion in the COGTIPS study and underwent screening of global cognitive deficits by the Montreal Cognitive Assessment, motor severity by the Unified Parkinson´s Disease Rating Scale-III, psychotic symptoms by the Schedule for Assessment of Positive Symptoms (Voss et al., 2013), depression by Beck Depression Inventory, and impulsive

control disorders by an interview. Eligible participants underwent a baseline assessment with structured interviews, questionnaires and a neuropsychological battery which included the COWAT, the Stroop and the ToL tests. After this, participants underwent MRI scans which included T1 and PSIR MRI scans. A more detailed explanation of the procedure and measures used in the COGTIPS study is described in van Balkom et al. (2019). The COGTIPS study was approved in 2016 by the Medical Ethical Committee (METc) of the VU University Medical Center with registration number: 2016.543/NL58750.029.16. The study followed the international guidelines in terms of biomedical intervention with humans (World Medical Association, 2013).

Statistical analyses

We used IBM SPSS Statistics 26.0 for performing the statistical analyses. Descriptive data was displayed using means and standard deviations. For categorical data frequencies were used. Outliers were examined using histograms and box plots.

Quality of thalamic segmentation

To compare the quality of thalamic segmentation using the two different approaches, we made qualitative and quantitative comparisons. For the qualitative comparison, we visually compared the number of voxels that differed between bilateral thalamic segmentations obtained from the two different approaches, attending to the nuclear groups where thalamic over or underestimation occurred. Less bilateral thalamic over or underestimation reflects a more accurate segmentation. For the quantitative comparison, Kolmogorov-Smirnov tests were used to check normality of total thalamic and total nuclear groups volumes. Depending on the agreement with the test assumptions, we performed paired-samples *t*-tests or Wilcoxon signed-rank tests and Pearson´s or Spearman´s correlations to examine the difference and relationship between total thalamic and total nuclear groups volumes obtained from the two different approaches. We also performed paired-samples *t*-tests or Wilcoxon signed-rank tests to examine the difference in white matter overlap of the thalamic segmentations obtained from the two different approaches. Significance level was set at $p < .05$ for all the analyses. Strong positive significant correlation coefficients reflect a high consistency between the derived volumes of the two different approaches.

Thalamic volumes and EF

To examine the relationship between thalamic and Mng volumes and EF performance, we used multiple linear regression analyses, which models are specified below. As prior studies have shown an association between age, sex and ICV and thalamic volume, age, sex and ICV were entered as nuisance regressors in all the models to control for inter-individual differences (Hughes et al., 2012; Li et al., 2014). As 76.3% of participants were in Hoehn & Yahr disease stage 2 and 2.5, and age has been associated with motor severity and disease progression, we decided not to control for disease stage or motor severity to avoid overcorrection (Levy, 2007). The assumptions of multiple linear regression analyses were checked for all the models, including multicollinearity, homoscedasticity and normal distribution of residuals.

Confirmatory analyses. To examine the relationship between total thalamic and total Mng volumes and global EF performance, two models were constructed where the independent variables were total thalamic or total Mng volumes and the dependent variable was the *z*-scores of the EF composite score. Two additional models were constructed to check the relationship between age, sex and ICV and total thalamic or total Mng volumes by entering age, sex and ICV as the independent variables and total thalamic or total Mng volumes as the dependent variables. Significance level was set at $p < .05$ for all the analyses.

Exploratory analyses. To examine the specific relationship between left or right thalamic and left or right Mng volumes and performance in each of the individual EF measures, a total of four models were constructed for each of the individual EF measures where the independent variables were left or right thalamic or left or right Mng volumes and the dependent variables were the *z*-scores of the total number of correct responses of the COWAT test, or the interference score of the Stroop test, or the total accuracy score of the ToL test. The significance level for these analyses was set at $p < .03$. This critical value was obtained using SISA (https://www.quantitativeskills.com/sisa /calculations/bonfer.htm), which adjusts α for multiple comparisons using Bonferroni´s correction while considering the mean correlation between variables ($r = .84$, 12 comparisons).

Results

Demographics and clinical characteristics

Out of the 85 participants that were initially included in this thesis project, six were excluded from analyses due to incorrect PSIR MRI scans, one was excluded for motion artifacts in the PSIR MRI scan and two were excluded for motion artifacts in the T1 MRI scans. The final sample comprised 76 participants: 29 males and 47 females. Their demographics and clinical characteristics are presented in Table 1. Nineteen participants were cognitively intact, 11 were suspected of having level II single domain cognitive impairment, 34 were suspected of having level II multiple domain cognitive impairment, and 12 were suspected of having probable PD dementia. Of those participants taking Parkinson´s disease medication, 74 were taking levodopa and one was taking duodopa. Of those taking psychiatric medication, six were taking antidepressants, 10 were taking benzodiazepines, one was taking antipsychotic, and three were taking cholinesterase inhibitors. Two participants had orthostatic hypotension.

Table 1

Demographics and clinical characteristics.

	N = 76
Age (years)	63.49 ± 7.45
Verhage education level (in %)	
3	1.32
4	5.26
5	27.63
6	36.84
7	28.95
MoCA	26.03 ± 2.14
Disease duration (years)	5.13 ± 3.39
Hoehn & Yahr disease stage (in %)	
1	9.21
1.5	7.89
2	46.05
2.5	30.26
3	6.58
UPDRS-III	19.43 ± 8.54
Psychiatric comorbidities	
BDI	7.82 ± 4.25
PAS	10.12 ± 7.22
QUIP	13.16 ± 26.30
AS	13.51 ± 4.56

Note. All variables except for Verhage education level and Hoehn & Yahr disease stage, are listed as Mean ± SD.

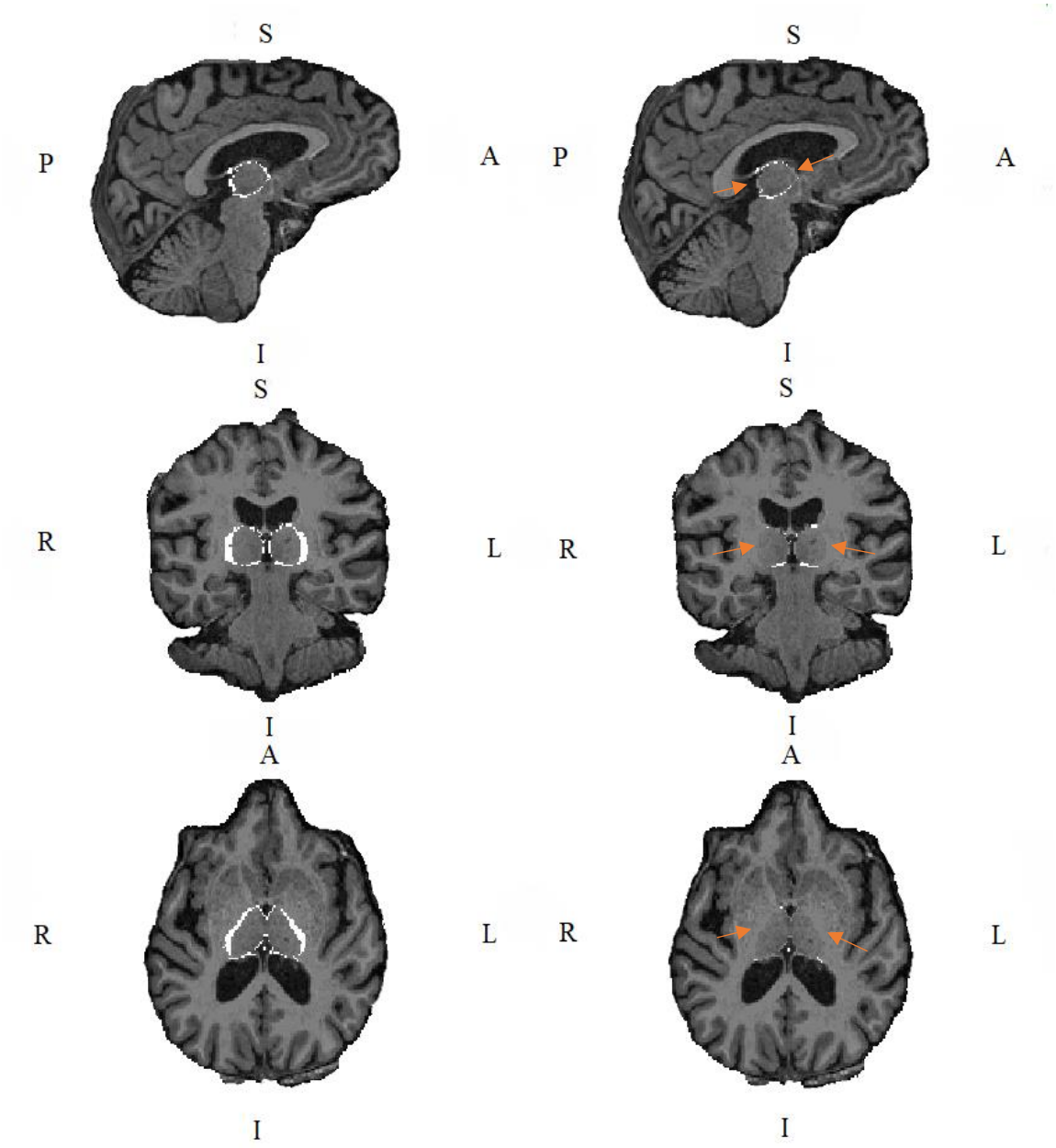
MoCA=Montreal Cognitive Assessment; UPDRS-III=Unified Parkinson´s Disease Rating Scale-Part III; BDI=Beck Depression Inventory; PAS=Parkinson´s Anxiety Scale; QUIP=Questionnaire for Impulsive-Compulsive Disorders in Parkinson´s Disease; AS=Apathy Scale.

Quality of thalamic segmentation

When compared qualitatively, we observed that thalamic segmentation using T1 MRI scans alone resulted in an overestimation of thalamic volume in all the thalamic boundaries bilaterally. The use of T1+PSIR MRI scans reduced this overestimation, especially in the lateral nuclear group bilaterally. When using T1+PSIR MRI scans, there was remaining overestimation in some thalamic boundaries, including the fornix, dorsal and ventral thalamus, and anterior and posterior thalamus bilaterally. An illustration of these differences in thalamic segmentation is presented in Figure 2.

Figure 2

Differences in thalamic segmentation using T1 (left) and T1+PSIR MRI scans (right).



Note. Additional voxels that differ between thalamic segmentations obtained from T1 and T1+PSIR MRI scans are shown in white. Arrows point to areas where thalamic segmentation was optimized.

We also compared the two different approaches quantitatively. All the assumptions of paired-samples *t*-tests and Wilcoxon signed-rank tests, and Pearson´s and Spearman´s correlations were met. Thalamic segmentation using T1+PSIR MRI scans resulted in a significantly smaller total thalamic volume compared to using T1 MRI scans alone (11,457 vs. 12,566 mm^3), [$Z = -7.57$, $p < .001$], although these volumes were significantly related ($r_s = .91$, $p < .001$). This was also the case for each of the total nuclear groups volumes (see Table 2). Thalamic segmentation of T1+PSIR MRI scans

resulted in fewer total number of outliers compared to using T1 MRI scans alone (3 vs.10). None of the approaches resulted in cerebrospinal fluid overlap. Thalamic segmentation using TI+PSIR MRI scans resulted in significantly more white matter overlap than using T1 MRI scans alone (252.01 vs. 200.68 mm^3), [$Z = -4.33$, $p < .001$]. Overall, we decided that thalamic segmentation using T1+PSIR MRI scans was the most accurate approach for thalamic segmentation in this thesis project and was used to relate the derived volumes to EF performance in the analyses reported below.

Table 2

Means, difference, and relation of total nuclear groups volumes.

	T1 Mean ± SD	T1+PSIR Mean ± SD	Difference	Relation
Ventral group (mm^3)	5,665 ± 728	4,929 ± 658	$Z = -7.57$*	$r_s = .91$*
Medial group (mm^3)	2,791 ± 260	2,665 ± 309	$t(75) = 6.77$*	$r = .85$*
Lateral group (mm^3)	302 ± 56	259 ± 48	$Z = -6.67$*	$r_s = .65$*
Pulvinar group (mm^3)	3,502 ± 400	3,357 ± 389	$t(75) = 7.52$*	$r = .91$*
Anteroventral group (mm^3)	303 ± 39	246 ± 34	$Z = -7.53$*	$r_s = .68$*

*significant at $p < .001$.

Thalamic volumes and EF

The assumptions of multiple linear regression analyses were met for all the models reported below.

Confirmatory analyses

We excluded the data of one participant from the analyses due to incomplete EF composite score. The sample $N = 75$ had a total mean thalamic volume of 11,464 mm^3 ($SD = 1,288$), total mean Mng volume of 2,670 mm^3 ($SD = 308$) and a mean *z*-score of -2.41 ($SD = 2.17$) in the EF composite score. There was a significant relationship between age ($\beta = -.17$, $t = -2.84$, $p = .006$), sex ($\beta = -.20$, $t = -2.74$, $p = .008$) and ICV ($\beta = .95$, $t = 13.62$, $p < .001$) and total thalamic volume, and between age ($\beta = -.31$, $t = -3.74$, $p < .001$), sex ($\beta = -.29$, $t = -2.94$, $p = .004$) and ICV ($\beta = .79$, $t = 8.20$, $p < .001$) and total Mng volume. Results of the multiple linear regression analyses showed no significant relationship between total thalamic or total Mng volumes and the *z*-scores of the EF composite score (see Table 3). As the relationship between *total* Mng volume and the *z*-scores of the EF composite score was not significant at $p < .05$ ($\beta = .26$, $t = 1.71$, $p = .091$), we performed two post hoc multiple regression analyses to examine whether this relationship was lateralized, including *left* or *right* Mng volumes as the independent variables, age, sex and ICV as nuisance regressors, and the *z*-scores of the EF composite score as the dependent variable. The relationship was not significant between left ($\beta = .21$, $t = 1.42$, $p = .160$) or right ($\beta = .26$, $t = 1.81$, $p = .075$) Mng volumes and the *z*-scores of the EF composite score.

Table 3

Multiple linear regression analyses for the EF composite score.

	EF composite score				
	B	*β*	*t*	*p*	R^2
Overall model					.19
Total thalamus	0.0004	.26	1.34	.218	
Age	-0.1077	-.37	-3.26	.002	
Sex	-0.4374	-.10	-0.74	.459	
ICV	< -0.0001	-.26	-1.08	.282	
Overall model					.21
Total Mng	0.0018	.26	1.71	.091	
Age	-0.0981	-.34	-2.91	.005	
Sex	-0.3426	-.08	-0.59	.559	
ICV	< -0.0001	-.21	-1.23	.225	

Exploratory analyses

We excluded the data of one participant from the analyses for the total accuracy score of the ToL test due to missing value. Means and standard deviations of left and right thalamic volumes, left and right Mng volumes and the *z*-scores of each of the individual EF measures are presented in Table 4. Results of the multiple linear regression analyses showed a positive significant relationship between left (β = .43, t = 2.26, p = .024; see Figure 3) and right (β = .51, t = 2.67, p = .009; see Figure 4) thalamic volumes and left (β = .40, t = 2.84, p = .006; see Figure 5) Mng volume and the *z*-scores of the total number of correct responses of the COWAT test, beyond the effects of age, sex and ICV (see Table 5). No significant relationship was found between left or right thalamic or left or right Mng volumes and the *z*-scores of the interference score of the Stroop test or the total accuracy score of the ToL test (see Table 6 and Table 7).

Table 4

Means of left and right thalamic and Mng volumes and individual EF measures.

	N	Mean ± SD
Left thalamus (mm^3)	76	5,708.78 ± 634.94
Right thalamus (mm^3)	76	5,748.82 ± 662.84
Left Mng (mm^3)	76	1,328.43 ± 152.40
Right Mng (mm^3)	76	1,336.94 ± 165.28
Total number correct responses COWAT (*z*-score)	76	-0.84 ± 1.00
Interference score Stroop (*z*-score)	76	5.59 ± 5.87
Total accuracy score ToL (*z*-score)	75	-0.97 ± 1.31

Figure 3

Partial plot of the relationship between left thalamic volume, controlling for age, sex and ICV, and the total number of correct responses of the COWAT test.

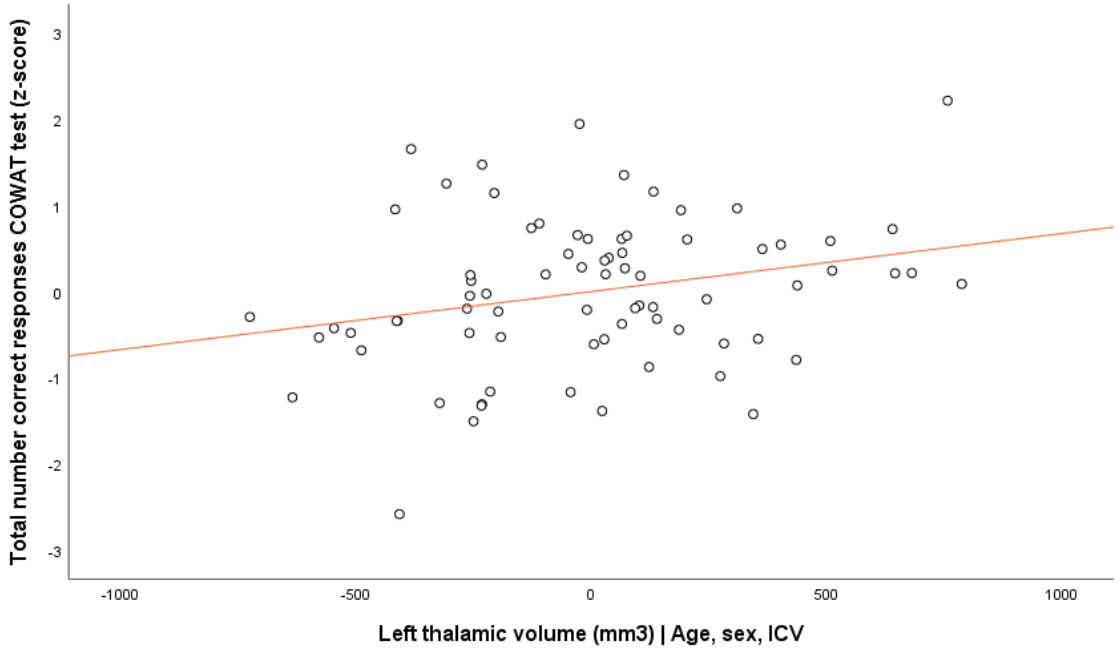


Figure 4

Partial plot of the relationship between right thalamic volume, controlling for age, sex and ICV, and the total number of correct responses of the COWAT test.

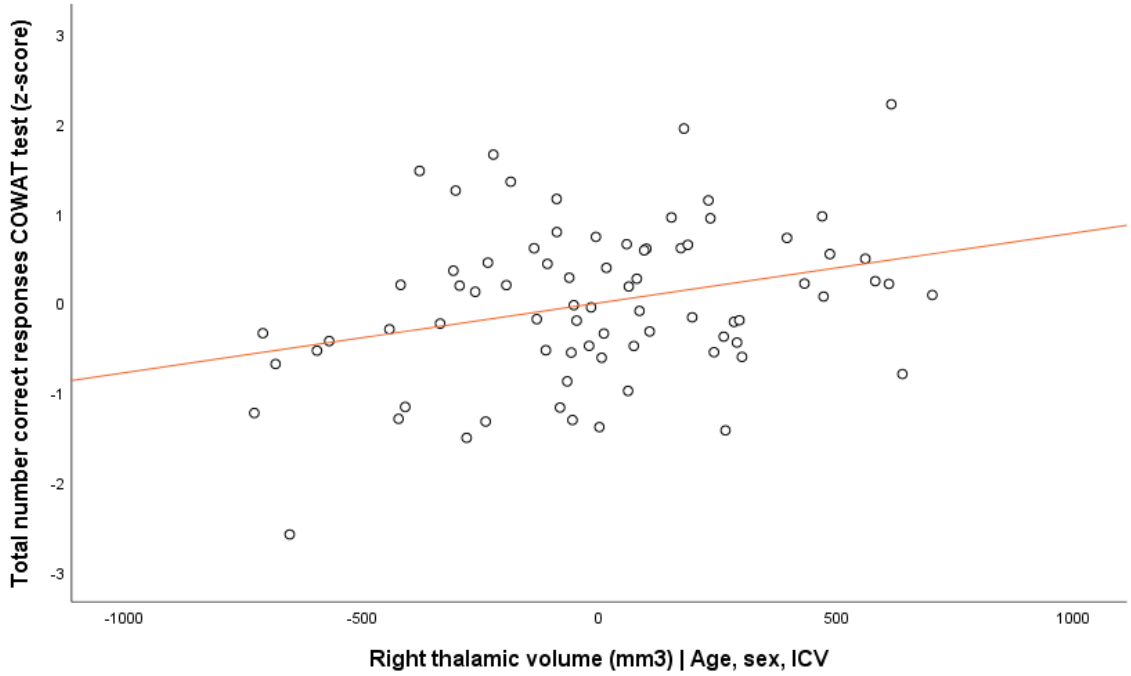


Figure 5

Partial plot of the relationship between left Mng volume, controlling for age, sex and ICV, and the total number of correct responses of the COWAT test.

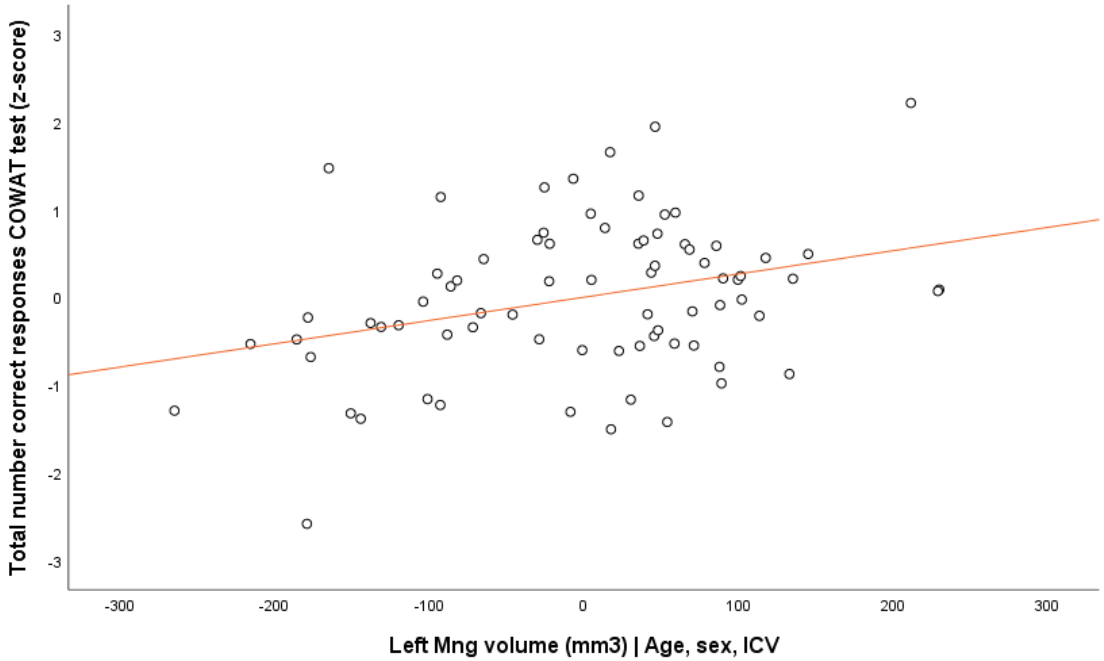


Table 5

Multiple linear regression analyses for the total number of correct responses of the COWAT test.

	Total number of correct responses of the COWAT test				
	B	*β*	*t*	*p*	R^2
Overall model					.26
Left thalamus	0.0007	.43	2.26	**.024**	
Age	-0.0018	-.01	-0.13	.899	
Sex	-0.8687	-.42	-3.30	.001	
ICV	< -0.0001	-.35	-1.66	.101	
Overall model					.28
Right thalamus	0.0008	.51	2.67	**.009**	
Age	0.0041	.03	0.29	.776	
Sex	-0.8412	-.41	-3.31	.001	
ICV	< -0.0001	-.44	-2.03	.046	
Overall model					.29
Left Mng	0.0027	.40	2.84	**.006**	
Age	0.0053	.04	0.37	.713	
Sex	-0.8188	-.40	-3.23	.002	
ICV	< -0.0001	-.27	-1.70	.096	
Overall model					.24
Right Mng	0.0014	.23	1.65	.103	
Age	0.0005	.00	0.31	.975	
Sex	-0.9016	-.44	-3.42	.001	
ICV	< -0.0001	-.13	-0.79	.433	

Table 6

Multiple linear regression analyses for the interference score of the Stroop test.

	Interference score of the Stroop test				
	B	*β*	*t*	*p*	R^2
Overall model					.15
Left thalamus	-0.0009	-.11	-0.52	.603	
Age	0.3083	.39	3.46	.001	
Sex	0.7803	.07	0.48	.632	
ICV	< 0.0001	.11	0.46	.646	
Overall model					.15
Right thalamus	-0.0013	-.15	-0.70	.488	
Age	0.2969	.38	3.22	.002	
Sex	0.6995	.06	0.43	.667	
ICV	< 0.0001	.14	0.61	.543	
Overall model					.15
Left Mng	-0.0058	-.15	-0.97	.336	
Age	0.2871	.36	3.11	.003	
Sex	0.5468	.05	0.34	.737	
ICV	< 0.0001	.13	0.71	.480	
Overall model					.17
Right Mng	-0.0075	-.21	-1.44	.155	
Age	0.2657	.34	2.85	.006	
Sex	0.2916	.02	0.18	.857	
ICV	< 0.0001	.16	0.98	.331	

Table 7

Multiple linear regression analyses for the total accuracy score of the ToL test.

	Total accuracy score of the ToL test				
	B	*β*	*t*	*p*	R^2
Overall model					.05
Left thalamus	0.0005	.23	1.05	.30	
Age	-0.0459	-.26	-2.17	.03	
Sex	-0.0409	-.02	-0.11	.92	
ICV	< -0.0001	-.18	-0.73	.47	
Overall model					.05
Right thalamus	0.0004	.21	0.95	.35	
Age	-0.0439	-.25	-1.20	.05	
Sex	-0.0525	-.02	-0.14	.89	
ICV	< -0.0001	-.17	-0.66	.51	
Overall model					.06
Left Mng	0.0017	.20	1.21	.23	
Age	-0.0420	-.24	-1.92	.06	
Sex	-0.0211	-.01	-0.06	.97	
ICV	< -0.0001	-.12	-0.66	.51	
Overall model					.07
Right Mng	0.0019	.24	1.55	.13	
Age	-0.0379	-.21	-1.72	.09	
Sex	0.0244	.01	0.06	.95	
ICV	< -0.0001	-.14	-0.82	.42	

Discussion

In this thesis project we used T1+PSIR MRI scans in FreeSurfer to optimize automatic thalamic segmentation and used the derived volumes to examine the association between thalamic and Mng volumes and EF in PD. We did not find a significant relationship between total thalamic and total Mng volumes and global EF performance. However, results of exploratory analyses showed a modest positive significant relationship between left and right thalamic volumes and left Mng volume and verbal fluency performance.

Although we did not find a significant relationship between total thalamic and total Mng volumes and global EF performance, when we performed exploratory analyses, we found that smaller left and right thalamic volumes and smaller left Mng volume were significantly related to worse verbal fluency performance, beyond the effects of age, sex and ICV. We replicated results of prior studies that showed that atrophy in the *whole* thalamus bilaterally was related to verbal fluency deficits in PD (Gerrits et al., 2016; Ibarretxe-Bilbao et al., 2010). We extended these findings by now showing a specific association between the left *Mng* volume and verbal fluency. Although to our knowledge there are no studies examining the relationship between Mng *volumes* and verbal fluency performance in PD, these results agree with Chung et al. (2017) who showed that *shape* changes in the left Mng were related to verbal fluency deficits in PD. Lateralization to the left could be related to the involvement of the left hemisphere in verbal processing rather than motor laterality, as most participants were in Hoehn & Yahr disease stage 2 and 2.5, reflecting bilateral involvement. The association between left Mng volume and verbal fluency in PD coincides with the connections of the mediodorsal nucleus with the prefrontal cortex, and intralaminar and midline nuclei with the striatum, medial prefrontal and anterior cingulate cortex, which are areas that have been related to verbal fluency (Abrahams et al., 2003; Saalmann, 2014; van der Werf et al., 2002). In addition, it also agrees with stroke and animal lesion studies that have shown a primarily role of the Mng in cognitive flexibility; a key function for verbal fluency (Saalmann, 2014; van der Werf, Jolles et al., 2003). Human post-mortem and animal model studies of PD have shown that within the Mng, the centromedian and parafascicular intralaminar subnuclei are the ones that undergo the most significant atrophy, even in early disease stages (Haliday, 2009; Henderson et al., 2000; Villalba et al., 2014). Some researchers have suggested that the significant atrophy in these intralaminar subnuclei could be related to cognitive flexibility deficits in PD (Galvan & Smith, 2011; Villalba et al., 2014). Although the centromedian and parafascicular intralaminar subnuclei were traditionally considered as "non-specific", stroke and animal lesion studies showed their involvement in specific cognitive functions, including cognitive flexibility, possibly by their role in arousal and awareness through their inputs from the reticular formation and brain stem and strong connections to the striatum (Bradfield et al., 2013; Liebermann, et al. 2013; Minamimoto et al., 2014; van der Werf et al., 2002).

Although we found a significant relationship between thalamic and Mng volumes and verbal fluency performance, we did not find a significant relationship between left or right thalamic and left

or right Mng volumes and performance in EF subdomains of resistance to cognitive interference or planning abilities. As EF, including these EF subdomains, draw on other cognitive functions, it could be the case that other thalamic nuclear groups or their combination better account for these relationships. Indeed, studies have shown that strokes in the anterior and lateral nuclear groups or their combination are related to EF deficits including resistance to cognitive interference and planning abilities (Carrera et al., 2004; Carrera et al., 2006; Ghika-Schmid & Bogousslavsky, 2000). In addition, within the Mng, the mediodorsal nucleus is the one that has been most consistently related to EF due to its strong connections with the prefrontal cortex (Parnaudeau, 2018). Apart from cognitive flexibility, animal lesion studies have shown a role of the mediodorsal nucleus in goal-directed behaviour, suppressing previous associations and working memory; functions that are required for resistance to cognitive interference and planning abilities beyond cognitive flexibility (Bradfield et al., 2013; Parnaudeau, 2018; Pickens, 2008). Indeed, in a stroke study, deficits in resistance to cognitive interference and planning abilities were related to wider lesions involving the mediodorsal, intralaminar and midline nuclei (van der Werf, Scheltens et al., 2003). In PD there is limited and contradictory evidence about the degree of neuronal degeneration and volume loss in the mediodorsal nucleus, with some in-vivo studies showing atrophy while post-mortem studies not showing a significant degeneration (Chen et al., 2020; Halliday, 2009; Henderson et al., 2000). Therefore, the degeneration of the mediodorsal nucleus in PD might not be enough to reflect in deficits in resistance to cognitive interference and planning abilities, explaining our non-significant relationship between Mng volume and performance in these EF subdomains. As we did not examine the presence of atrophy in the mediodorsal nucleus and its relationship with performance in resistance to cognitive interference and planning abilities, this interpretation is limited but opens questions for future studies.

Although the association between thalamic and Mng volumes and global EF in PD is not supported in this thesis project, thalamic and Mng volumes seem to be specifically associated with verbal fluency. In this thesis project, we cannot draw firm conclusions about the directionality of the relationship between thalamic and Mng volumes and verbal fluency in PD given the cross-sectional nature of the analyses. However, as smaller thalamic volumes have been shown to predict cognitive deficits in PD, supported by future longitudinal studies, smaller thalamic, and specifically smaller Mng volumes, could potentially facilitate early diagnosis of verbal fluency deficits (Foo et al., 2017). In PD verbal fluency deficits are common and have a psychosocial impact: they impair communication, change family dynamics, and produce a sense of loss of independence, control and confidence, which is associated with a reduced quality of life (Miller et al., 2008; Tang et al., 2020). Early diagnosis of verbal fluency deficits could facilitate early treatment implementation and effective management of verbal fluency deficits, consequently preventing their impact on psychosocial well-being and quality of life (Buono et al., 2021; Miller et al., 2008). Indeed, verbal fluency deficits occur together with other EF deficits and have been suggested as a risk factor for developing dementia (Lee et al., 2014; Torralva et al., 2015). However, as cognitive deficits in PD are heterogenous and verbal

fluency deficits do not always come first, early diagnosis of verbal fluency deficits might not be sensitive enough to predict general cognitive decline (van Balkom et al., 2016).

As a secondary objective of this thesis project, we examined whether T1+PSIR MRI scans would optimize automatic thalamic segmentation in FreeSurfer. Total thalamic and total nuclear groups volumes obtained from the two different approaches were significantly related, reflecting consistency. However, we observed that using T1+PSIR MRI scans resulted in less thalamic overestimation in all the boundaries of left and right thalamus and significantly smaller total thalamic and total nuclear groups volumes compared to using T1 MRI scans alone. Visually we found that this overestimation was significantly reduced in the lateral nuclear group, implying a better recognition of the internal capsule. In fact, there was a weaker consistency between the volumes of the lateral and anteroventral nuclear groups derived from using T1 and T1+PSIR MRI scans compared to other nuclear groups. The reduced thalamic overestimation and smaller thalamic and nuclear groups volumes obtained using T1+PSIR MRI scans are consistent with the superior grey-white matter contrast of PSIR MRI scans which in addition to T1 MRI scans can optimize thalamic segmentation (Hou et al., 2005). These findings are promising as they are in line with gold-standard manual segmentations which result in more accurate thalamic volumes than automatic segmentations (Makowski et al., 2018). Unexpectedly, thalamic segmentation using T1+PSIR MRI scans resulted in significantly more white matter overlap than using T1 MRI scans alone. Considering that this overlap was obtained using automatic white matter segmentation of T1 MRI scans in FreeSurfer, this result might not be accurate as it might include white matter that pertains to the thalamus. Overall, although using T1+PSIR MRI scans still had some shortcomings, this approach seems to be more accurate than using T1 MRI scans alone. As there are no other studies to our knowledge that examine this specific approach, results should be interpreted with caution. Replications of these results and in other populations, especially in those where atrophy is not expected, are needed to make firm conclusions.

A notable strength of this thesis project is that we are the first to our knowledge to explore the relationship between Mng volume and EF in PD and show an association between the left Mng volume and verbal fluency. We are also the first to examine and demonstrate that using T1+PSIR MRI scans can optimize automatic thalamic segmentation. This novel approach allowed us to relate EF to thalamic volumes that were likely more accurate than those of other studies that were obtained by automatic thalamic segmentation using T1 MRI scans alone in FreeSurfer (Chen et al., 2020; Gerrits et al., 2016). However, this segmentation approach also represents a possible limitation of this thesis project as it did not allow to directly compare our results to other studies examining the relationship between EF and thalamic volumes due to different volume estimations (Furlong et al., 2020; Gerrits et al., 2016). Although the Mng was relatively spared, whole thalamic volumes of this thesis project could be biased as using T1+PSIR MRI scans resulted in some remaining thalamic overestimation; future studies would ideally correct misclassifications. Although we used normalized scores of EF performance which allowed to operationalize EF in PD, we did not compare thalamic

and Mng volumes to those of a healthy control group, thus, it is unclear to what extent the association between thalamic and Mng volumes and EF is specific to PD or might generalize to other populations. Although in-vivo studies have shown thalamic atrophy in PD, in-vivo studies examining atrophy in the Mng are limited (Chen et al., 2020; Mak et al., 2014). In addition, studies have shown that some thalamic subnuclei might undergo hypertrophy as a compensatory mechanism when cortical structures are affected, resulting in whole thalamic hypertrophy or no volume change (Chen et al., 2020; McKeown, 2008). Thus, future studies would ideally compare thalamic and Mng volumes to a healthy control group and include shape and volumetric analyses for a more reliable measure of morphometric thalamic changes. Another limitation of this thesis project is that the operationalization of global EF did not encompass the entirety of EF, which could be addressed in future studies. Indeed, future work would ideally include direct measures of cognitive flexibility to help unravel the association we found between the left Mng volume and verbal fluency. Lastly, the sample consisted of more females than males, not representing the sex distribution of PD in the population, which is 1.5 times more frequent in males (Reekes et al., 2020); future studies including a more equally distributed sample would allow a better generalisation of findings.

In conclusion, in this thesis project we demonstrated that using PSIR MRI scans in addition to T1 MRI scans in FreeSurfer can optimize automatic thalamic segmentation. Although the association between thalamic and Mng volumes and global EF in PD is not supported in this thesis project, volumes seem to be specifically associated with verbal fluency. We propose that this could be due to the role of the Mng in cognitive flexibility. Supported by future longitudinal studies, these findings could have implications for early diagnosis and effective management of verbal fluency deficits in PD.

References

Aarsland, D., Andersen, K., Larsen, J. P., & Lolk, A. (2003). Prevalence and characteristics of dementia in Parkinson disease: an 8-year prospective study. *Archives of Neurology*, *60*(3), 387-392. https://doi.org/10.1001/archneur.60.3.387

Abrahams, S., Goldstein, L. H., Simmons, A., Brammer, M. J., Williams, S. C., Giampietro, V. P., Andrew, C., & Leigh, P. N. (2003). Functional magnetic resonance imaging of verbal fluency and confrontation naming using compressed image acquisition to permit overt responses. *Human Brain Mapping*, *20*(1), 29-40. https://doi.org/10.1002/hbm.10126

Aita, S. L., Beach, J. D., Taylor, S. E., Borgogna, N. C., Harrell, M. N., & Hill, B. D. (2018). Executive, language, or both? An examination of the construct validity of verbal fluency measures. *Applied Neuropsychology*, *26*(5), 441-451. https://doi.org/10.1080/23279095.2018.1439830

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, *4*(6), 561-571. https://doi.org/10.1001/archpsyc.1961.01710120031004

Benton, A. L., Sivan, A. B., des Hamsher, K., & Varney, N. R. (1994). *Contributions to neuropsychological assessment: A clinical manual*. Oxford University Press.

Bradfield, L. A., Hart, G., & Balleine, B. W. (2013). The role of the anterior, mediodorsal, and parafascicular thalamus in instrumental conditioning. *Frontiers in Systems Neuroscience*, *7*: 51. https://doi.org/10.3389/fnsys.2013.00051

Buono, V. L., Palmeri, R., De Salvo, S., Berenati, M., Greco, A., Ciurleo, R., Sobera, C., Cimino, V., Corallo, F., Bramanti, P., Marino, S., Di Lorenzo, G., & Bonanno, L. (2021). Anxiety, depression, and quality of life in Parkinson's disease: the implications of multidisciplinary treatment. *Neural Regeneration*, *16*(3), 587. https://doi.org/10.4103/1673-5374.293151

Carrera, E., & Bogousslavsky, J. (2006). The thalamus and behavior: effects of anatomically distinct strokes. *Neurology*, *66*(12), 1817-1823. https://doi.org/10.1212/01.wnl.0000219679.95223.4c

Carrera, E., Michel, P., & Bogousslavsky, J. (2004). Anteromedian, central, and posterolateral infarcts of the thalamus: three variant types. *Stroke*, *35*(12), 2826-2831. https://doi.org/10.1161/01.STR.0000147039.49252.2f

Chen, Y., Zhu, G., Liu, D., Liu, Y., Yuan, T., Zhang, X., Jiang, Y., Du, T., & Zhang, J. (2020). The morphology of thalamic subnuclei in Parkinson's disease and the effects of machine learning on disease diagnosis and clinical evaluation. *Journal of the Neurological Sciences*, *411*, 116721. https://doi.org/10.1016/j.jns.2020.116721

Chung, S. J., Shin, J. H., Cho, K. H., Lee, Y., Sohn, Y. H., Seong, J. K., & Lee, P. H. (2017). Subcortical shape analysis of progressive mild cognitive impairment in Parkinson's disease. *Movement Disorders*, *32*(10), 1447-1456. https://doi.org/10.1002/mds.27106

Chung, S. J., Yoo, H. S., Oh, J. S., Kim, J. S., Ye, B. S., Sohn, Y. H., & Lee, P. H. (2018). Effect of striatal dopamine depletion on cognition in de novo Parkinson's disease. *Parkinsonism & Related Disorders, 51,* 43-48. https://doi.org/10.1016/j.parkreldis.2018.02.048

Debelak, R., Egle, J., Köstering, L., & Kaller, C. (2016). Assessment of planning ability: Psychometric analyses on the unidimensionality and construct validity of the Tower of London Task (TOL-F). *Neuropsychology, 30(3), 346–360.* https://doi.org/10.1037/neu0000238

Ding, W., Ding, L. J., Li, F. F., Han, Y., & Mu, L. (2015). Neurodegeneration and cognition in Parkinson's disease: a review. *European Review for Medical and Pharmacological Sciences*, *19*(12), 2275-81.

Edelstyn, N. M. J., Mayes, A. R., & Ellis, S. J. (2014). Damage to the dorsomedial thalamic nucleus, central lateral intralaminar thalamic nucleus, and midline thalamic nuclei on the right-side impair executive function and attention under conditions of high demand but not low demand. *Neurocase*, *20*(2), 121-132. https://doi.org/10.1080/13554794.2012.713497

Emre, M., Aarsland, D., Brown, R., Burn, D. J., Duyckaerts, C., Mizuno, Y., Broe, G. A., Cummings, J., Dickson, D. W., Gauthier, S., Goldman, J., Goetz, C., Korczyn A., Lees, A., Levy, R., Litvan, I., McKeith, I., Olanow, W., Poewe, W., … Dubois, B. (2007). Clinical diagnostic criteria for dementia associated with Parkinson's disease. *Movement Disorders*, *22*(12), 1689-1707. https://doi.org/10.1002/mds.21507

Ewing, J. A. (1984). Detecting alcoholism: the CAGE questionnaire. *Jama*, *252*(14), 1905-1907. https:/doi.org/10.1001/jama.1984.03350140051025

Fahn, S., & Elton, R. L. (1987). UPDRS Development Committee A: Unified Parkinson's disease rating scale. In S. Fahn, C. D. Marsden, D. B. Calne & M. Goldstein (Eds.), *Recent Developments in Parkinson's Disease* (pp. 293-304). Macmillan Healthcare Information.

Fang, C., Lv, L., Mao, S., Dong, H., & Liu, B. (2020). Cognition deficits in Parkinson's disease: mechanisms and treatment. *Parkinson's Disease*, *2020*: 2076942. https://doi.org/10.1155/2020/2076942

Fischl, B. (2012). FreeSurfer. *Neuroimage*, *62*(2), 774-781. https://doi.org/10.1016/j.neuroimage.2012.01.021

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical approach for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189-198. https://doi.org/10.1016/0022-3956(75)90026-6

Foo, H., Mak, E., Yong, T. T., Wen, M. C., Chander, R. J., Au, W. L., Sitoh, Y. Y., Tan, L. C. S., & Kandiah, N. (2017). Progression of subcortical atrophy in mild Parkinson's disease and its impact on cognition. *European Journal of Neurology*, *24*(2), 341-348. https://doi.org/10.1111/ene.13205

Foster, E. R., & Hershey, T. (2011). Everyday executive function is associated with activity participation in Parkinson disease without dementia. *American Journal of Occupational Therapy*, *31*(1), 16-22. https://doi.org/10.3928/15394492-20101108-04

Furlong, L. S., Jakabek, D., Power, B. D., Owens-Walton, C., Wilkes, F. A., Walterfang, M., Velakoulis, D., Egan, G., Looi, J., & Georgiou-Karistianis, N. (2020). Morphometric in vivo evidence of thalamic atrophy correlated with cognitive and motor dysfunction in Huntington's disease: The IMAGE-HD study. *Psychiatry Research, 298*: 111048. https://doi.org/10.1016/j.pscychresns.2020.111048

Galvan, A., & Smith, Y. (2011). The primate thalamostriatal systems: anatomical organization, functional roles and possible involvement in Parkinson's disease. *Basal Ganglia*, 1, 179–189. https://doi.org/10.1016/j.baga.2011.09.001

Gerrits, N. J., van Loenhoud, A. C., van den Berg, S. F., Berendse, H. W., Foncke, E. M., Klein, M., Stoffers, D., van der Werf, Y., & 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), e0148852. https://doi.org/10.1371/journal.pone.0148852

Ghika-Schmid, F., & Bogousslavsky, J. (2000). The acute behavioral syndrome of anterior thalamic infarction: a prospective study of 12 cases. *Annals of Neurology*, *48*(2), 220-227. https://doi.org/10.1002/1531-8249(200008)48:2<220::AID-ANA12>3.0.CO;2-M

Gill, D. J., Freshman, A., Blender, J. A., & Ravina, B. (2008). The Montreal cognitive assessment as a screening tool for cognitive impairment in Parkinson's disease. *Movement disorders*, *23*(7), 1043-1046. https://doi.org/10.1002/mds.22017

Grahn, J. A., Parkinson, J. A., & Owen, A. M. (2008). The cognitive functions of the caudate nucleus. *Progress in Neurobiology*, *86*(3), 141-155. https://doi.org/10.1016/j.pneurobio.2008.09.004

Halliday, G. M. (2009). Thalamic changes in Parkinson's disease. *Parkinsonism & Related Disorders*, *15*, S152-S155. https://doi.org/10.1016/S1353-8020(09)70804-1

Hammes, J. G. W. (1971). *De Stroop Kleur-Woord Test. Handleiding.* Pearson Assessment and Information B.V.

Henderson, J. M., Carpenter, K., Cartwright, H., & Halliday, G. M. (2000). Degeneration of the centré median–parafascicular complex in Parkinson's disease. *Annals of Neurology*, *47*(3), 345-352. https://doi.org/10.1002/1531-8249(200003)47:3<345::AID-ANA10>3.0.CO;2-V

Hoehn, M. M., & Yahr, M. D. (1998). Parkinsonism: onset, progression, and mortality. *Neurology*, *50*(2), 318-318. https://doi.org/10.1212/wnl.50.2.318

Hou, P., Hasan, K. M., Sitton, C. W., Wolinsky, J. S., & Narayana, P. A. (2005). Phase-sensitive T1 inversion recovery imaging: a time-efficient interleaved technique for improved tissue contrast in neuroimaging. *American Journal of Neuroradiology*, *26*(6), 1432-1438.

Hughes, E. J., Bond, J., Svrckova, P., Makropoulos, A., Ball, G., Sharp, D. J., Edward, D., Hajnal, J., & Counsell, S. J. (2012). Regional changes in thalamic shape and volume with increasing age. *Neuroimage*, *63*(3), 1134-1142. https://doi.org/10.1016/j.neuroimage.2012.07.043

Ibarretxe-Bilbao, N., Ramirez-Ruiz, B., Junque, C., Marti, M. J., Valldeoriola, F., Bargallo, N., Juanes, S., & Tolosa, E. (2010). Differential progression of brain atrophy in Parkinson's disease with and without visual hallucinations. *Journal of Neurology, Neurosurgery & Psychiatry*, *81*(6), 650-657. http://doi.org/10.1136/jnnp.2009.179655

Iglesias, J. E., Insausti, R., Lerma-Usabiaga, G., Bocchetta, M., Van Leemput, K., Greve, D. N., van der Kouwe, A., Fischl, B., Caballero-Gaudes, C., & Alonso, P. (2018). A probabilistic atlas of the human thalamic nuclei combining ex vivo MRI and histology. *Neuroimage*, *183*, 314-326. https://doi.org/10.1016/j.neuroimage.2018.08.012

Jankovic, J. (2008). Parkinson's disease: clinical features and diagnosis. *Journal of Neurology, Neurosurgery & Psychiatry*, *79*(4), 368-376. https://doi.org/10.1136/jnnp.2007.131045

Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W., & Smith, S. M. (2012). Fsl. *Neuroimage*, *62*(2), 782-790. https://doi.org/10.1016/j.neuroimage.2011.09.015

Kaplan, E., Goodglass, H., & Weintraub, S. (2001). *Boston naming test*. Pro-ed.

Kassubek, J., Juengling, F. D., Ecker, D., & Landwehrmeyer, G. B. (2005). Thalamic atrophy in Huntington's disease co-varies with cognitive performance: a morphometric MRI analysis. *Cerebral Cortex*, *15*(6), 846-853. https://doi.org/10.1093/cercor/bhh185

Kessels, R. P., Nys, G. M., Brands, A. M., & van Zandvoort, M. J. (2004). The Location Learning Test as a measure of spatial memory: applicability of a modified administration procedure and normative data. *Tijdschrift voor Gerontologie en Geriatrie*, *35*(4), 147-152.

Kulisevsky, J., de Bobadilla, R. F., Pagonabarraga, J., Martínez-Horta, S., Campolongo, A., García-Sánchez, C., Pascual-Sedano, B., Ribosa-Nogué, R., & Villa-Bonomo, C. (2013). Measuring functional impact of cognitive impairment: validation of the Parkinson's disease cognitive functional rating scale. *Parkinsonism & Related Disorders*, *19*(9), 812-817. https://doi.org/10.1016/j.parkreldis.2013.05.007

Lanskey, J. H., McColgan, P., Schrag, A. E., Acosta-Cabronero, J., Rees, G., Morris, H. R., & Weil, R. S. (2018). Can neuroimaging predict dementia in Parkinson's disease?. *Brain*, *141*(9), 2545-2560. https://doi.org/10.1093/brain/awy211

Lee, J. E., Cho, K. H., Song, S. K., Kim, H. J., Lee, H. S., Sohn, Y. H., & Lee, P. H. (2014). Exploratory analysis of neuropsychological and neuroanatomical correlates of progressive MCI in Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry*, *85*(1), 7-16. https://doi.org/10.1136/jnnp-2013-305062

Leentjens, A. F., Dujardin, K., Pontone, G. M., Starkstein, S. E., Weintraub, D., & Martinez-Martin, P. (2014). The Parkinson Anxiety Scale (PAS): development and validation of a new anxiety scale. *Movement Disorders*, *29*(8), 1035-1043. https://doi.org/10.1002/mds.25919

Levy, G. (2007). The relationship of Parkinson disease with aging. *Archives of Neurology*, *64*(9), 1242-1246. https://doi.org/10.1001/archneur.64.9.1242

Li, W., van Tol, M. J., Li, M., Miao, W., Jiao, Y., Heinze, H. J., Bogerts, B., He, H., Walter, M., & Walter, M. (2014). Regional specificity of sex effects on subcortical volumes across the lifespan in healthy aging. *Human Brain Mapping*, *35*(1), 238-247. https://doi.org/10.1002/hbm.22168

Liebermann, D., Ploner, C. J., Kraft, A., Kopp, U. A., & Ostendorf, F. (2013). A dysexecutive syndrome of the medial thalamus. *Cortex*, *49*(1), 40-49. https://doi.org/10.1016/j.cortex.2011.11.005

Litvan, I., Goldman, J. G., Tröster, A. I., Schmand, B. A., Weintraub, D., Petersen, R. C., Mollenhauer, B., Adler, C. H., Marder, K., Williams-Gray, C., Aarsland, D., Kulisevsky, J., Rodriguez-Oroz, M., Burn, D., Barker, R., & Emre, M. (2012). Diagnostic criteria for mild cognitive impairment in Parkinson's disease: Movement Disorder Society Task Force guidelines. *Movement Disorders*, *27*(3), 349-356. https:/doi.org/10.1002/mds.24893

Mak, E., Bergsland, N., Dwyer, M. G., Zivadinov, R., & Kandiah, N. (2014). Subcortical atrophy is associated with cognitive impairment in mild Parkinson disease: a combined investigation of volumetric changes, cortical thickness, and vertex-based shape analysis. *American Journal of Neuroradiology*, *35*(12), 2257-2264. https:/doi.org/10.3174/ajnr.A4055

Makowski, C., Béland, S., Kostopoulos, P., Bhagwat, N., Devenyi, G. A., Malla, A. K., Joober, R., Lepage, M., & Chakravarty, M. M. (2018). Evaluating accuracy of striatal, pallidal, and thalamic segmentation methods: Comparing automated approaches to manual delineation. *Neuroimage*, *170*, 182-198. https://doi.org/10.1016/j.neuroimage.2017.02.069

McKeown, M. J., Uthama, A., Abugharbieh, R., Palmer, S., Lewis, M., & Huang, X. (2008). Shape (but not volume) changes in the thalami in Parkinson disease. *BMC Neurology*, *8*(1), 1-8. https:/doi.org/10.1186/1471-2377-8-8

Meyers, J. E., & Meyers, K. R. (1995). *Rey Complex Figure Test and recognition trial professional manual*. Psychological Assessment Resources.

Miller, N., Noble, E., Jones, D., Allcock, L., & Burn, D. J. (2008). How do I sound to me? Perceived changes in communication in Parkinson's disease. *Clinical Rehabilitation*, *22*(1), 14-22. https://doi.org/10.1177/0269215507079096

Minamimoto, T., Hori, Y., Yamanaka, K., & Kimura, M. (2014). Neural signal for counteracting pre-action bias in the centromedian thalamic nucleus. *Frontiers in Systems Neuroscience, 8*: 3. https://doi.org/10.3389/fnsys.2014.00003

Nasreddine, Z. S., Phillips, N. A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for MCI. *Journal of the American Geriatrics Society, 53*(4), 695-699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

Owen, A. M. (2004). Cognitive dysfunction in Parkinson's disease: the role of frontostriatal circuitry. *The Neuroscientist*, *10*(6), 525-537. https://doi.org/10.1177/1073858404266776

Owens-Walton, C., Jakabek, D., Power, B. D., Walterfang, M., Hall, S., van Westen, D., Looi, J., Shaw, M., & Hansson, O. (2021). Structural and functional neuroimaging changes associated with cognitive impairment and dementia in Parkinson's disease. *Psychiatry Research*, *312*: 111273. https://doi.org/10.1016/j.pscychresns.2021.111273

Parnaudeau, S., Bolkan, S. S., & Kellendonk, C. (2018). The mediodorsal thalamus: an essential partner of the prefrontal cortex for cognition. *Biological Psychiatry*, *83*(8), 648-656. https://doi.org/10.1016/j.biopsych.2017.11.008

Pedersen, K. F., Larsen, J. P., Tysnes, O. B., & Alves, G. (2017). Natural course of MCI in Parkinson disease: a 5-year population-based study. *Neurology*, *88*(8), 767-774. https://doi.org/10.1212/WNL.0000000000003634

Periáñez, J. A., Lubrini, G., García-Gutiérrez, A., & Ríos-Lago, M. (2021). Construct validity of the Stroop color-word test: Influence of speed of visual search, verbal fluency, working memory, cognitive flexibility, and conflict monitoring. *Archives of Clinical Neuropsychology*, *36*(1), 99-111. https://doi.org/ 10.1093/arclin/acaa034

Pickens, C. L. (2008). A limited role for mediodorsal thalamus in devaluation tasks. *Behavioral Neuroscience*, *122*(3), 659. https://doi.org/10.1037/0735-7044.122.3.659

Reekes, T. H., Higginson, C. I., Ledbetter, C. R., Sathivadivel, N., Zweig, R. M., & Disbrow, E. A. (2020). Sex specific cognitive differences in Parkinson disease. *npj Parkinson's Disease*, *6*(1), 1-6. https://doi.org/10.1038/s41531-020-0109-1

Saalmann, Y. B. (2014). Intralaminar and medial thalamic influence on cortical synchrony, information transmission and cognition. *Frontiers in Systems Neuroscience*, *8*: 83. https://doi.org/10.3389/fnsys.2014.00083

Saan, R. J., & Deelman, B. G. (1986). *De 15-woordentest A en B (een voorlopige handleiding)*. Afdeling Neuropsychologie, AZG.

Schapira, A. H., Chaudhuri, K. R., & Jenner, P. (2017). Non-motor features of Parkinson disease. *Nature Reviews Neuroscience*, *18*(7), 435. https://doi.org/10.1038/nrn.2017.62

Scharre, D. W., Chang, S. I., Murden, R. A., Lamb, J., Beversdorf, D. Q., Kataki, M., Nagaraja, H. N., & Bornstein, R. A. (2010). Self-administered Gerocognitive Examination (SAGE): a brief cognitive assessment Instrument for mild cognitive impairment (MCI) and early dementia. *Alzheimer Disease & Associated Disorders*, *24*(1), 64-71. https://doi.org/10.1097/WAD.0b013e3181b03277

Schmand, B., Groenink, SC, & Van den Dungen, M. (2008). Letterfluency: psychometric properties and Dutch standards. *Journal of Gerontology and Geriatrics*, *39*(2), 64-74. https://doi.org/ 10.1007/BF03078128

Serrano-Dueñas, M., Martínez-Martín, P., Merchán, T., Bravo, R., & Serrano, M. (2013). Properties of the Apathy Scale (AS) for use on Parkinson's patients. *Advances in Parkinson's Disease*, *2*(2), 53-57. https://doi.org/10.4236/apd.2013.22010

Smith, C. R., Cullen, B., Sheridan, M. P., Cavanagh, J., Grosset, K. A., & Grosset, D. G. (2020). Cognitive impairment in Parkinson's disease is multifactorial: A neuropsychological study. *Acta Neurologica Scandinavica*, *141*(6), 500-508. https://doi.org/10.1111/ane.13226

Sollinger, A. B., Goldstein, F. C., Lah, J. J., Levey, A. I., & Factor, S. A. (2010). Mild cognitive impairment in Parkinson's disease: subtypes and motor characteristics. *Parkinsonism & Related Disorders*, *16*(3), 177-180. https://doi.org/10.1016/j.parkreldis.2009.11.002

Song, S. K., Lee, J. E., Park, H. J., Sohn, Y. H., Lee, J. D., & Lee, P. H. (2011). The pattern of cortical atrophy in patients with Parkinson's disease according to cognitive status. *Movement Disorders*, *26*(2), 289-296. https://doi.org/10.1002/mds.23477

Starkstein, S. E., Mayberg, H. S., Preziosi, T., Andrezejewski, P., Leiguarda, R., & Robinson, R. G. (1992). Reliability, validity, and clinical correlates of apathy in Parkinson's disease. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *4*(2), 134-139. https://doi.org/10.1038/nrn.2017.62

Tang, Y., Liang, X., Han, L., Peng, F., Shen, B., Yu, H., Shen, Y., Shen, C., Yu, J., Wang, J., & Wang, J. (2020). Cognitive function and quality of life in Parkinson's disease: a cross-sectional study. *Journal of Parkinson's Disease*, *10*(3), 1209-1216. https://doi.org/10.3233/JPD-202097

Torralva, T., Laffaye, T., Báez, S., Gleichgerrcht, E., Bruno, D., Chade, A., Ibañez, A., Manes, F., Gershanik, O., & Roca, M. (2015). Verbal fluency as a rapid screening test for cognitive impairment in early Parkinson's disease. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *27*(3), 244-247. https://doi.org/10.1176/appi.neuropsych.14060139

Trujillo, J. P., Gerrits, N. J., Vriend, C., Berendse, H. W., van den Heuvel, O. A., & van der Werf, Y. D. (2015). Impaired planning in Parkinson's disease is reflected by reduced brain activation and connectivity. *Human Brain Mapping*, *36*(9), 3703-3715. https://doi.org/10.1002/hbm.22873.

Unterrainer, J. M., Rahm, B., Kaller, C. P., Leonhart, R., Quiske, K., Hoppe-Seyler, K., Meier, C., Müller, C., & Halsband, U. (2004). Planning abilities and the Tower of London: is this task measuring a discrete cognitive function?. *Journal of Clinical and Experimental Neuropsychology*, *26*(6), 846-856. https://doi.org/10.1080/13803390490509574

van Balkom, T. D., Berendse, H. W., van der Werf, Y. D., Twisk, J. W., Zijlstra, I., Hagen, R. H., Berk, T., Vriend, C., & van den Heuvel, O. A. (2019). COGTIPS: a double-blind randomized active controlled trial protocol to study the effect of home-based, online cognitive training on cognition and brain networks in Parkinson's disease. *BMC Neurology*, *19*(1), 1-13. https://doi.org/10.1186/s12883-019-1403-6

van Balkom, T. D., Vriend, C., Berendse, H. W., Foncke, E. M., van der Werf, Y. D., van den Heuvel, O. A., & Klein, M. (2016). Profiling cognitive and neuropsychiatric heterogeneity in Parkinson's disease. *Parkinsonism & Related Disorders*, *28*, 130-136. https://doi.org/10.1016/j.parkreldis.2016.05.014

van der Werf, Y. D., Jolles, J., Witter, M. P., & Uylings, H. B. (2003). Contributions of thalamic nuclei to declarative memory functioning. *Cortex*, *39*(4-5), 1047-1062. https://doi.org/10.1016/s0010-9452(08)70877-3

van der Werf, Y. D., Scheltens, P., Lindeboom, J., Witter, M. P., Uylings, H. B., & Jolles, J. (2003). Deficits of memory, executive functioning and attention following infarction in the thalamus; a study of 22 cases with localised lesions. *Neuropsychologia*, *41*(10), 1330-1344. https://doi.org/10.1016/s0028-3932(03)00059-9

van der Werf, Y. D., Witter, M. P., & Groenewegen, H. J. (2002). The intralaminar and midline nuclei of the thalamus. Anatomical and functional evidence for participation in processes of arousal and awareness. *Brain Research Reviews*, *39*(2-3), 107-140. https://doi.org/10.1016/s0165-0173(02)00181-9

Varela, C. (2014). Thalamic neuromodulation and its implications for executive networks. *Frontiers in Neural Circuits*, *8*: 69. https://doi.org/10.3389/fncir.2014.00069

Vasconcellos, L. F., Pereira, J. S., Adachi, M., Greca, D., Cruz, M., Malak, A. L., & Charchat-Fichman, H. (2018). Volumetric brain analysis as a predictor of a worse cognitive outcome in Parkinson's disease. *Journal of Psychiatric Research*, *102*, 254-260. https://doi.org/10.1016/j.jpsychires.2018.04.016

Verhage, F. (1964). *Intelligentie en leeftijd: Onderzoek bij Nederlanders van twaalf tot zevenenzeventig jaar*. Van Gorcum Assen.

Villalba, R. M., Wichmann, T., & Smith, Y. (2014). Neuronal loss in the caudal intralaminar thalamic nuclei in a primate model of Parkinson's disease. *Brain Structure and Function*, *219*(1), 381-394. https://doi.org/10.1007/s00429-013-0507-9

Visser, M., Leentjens, A. F., Marinus, J., Stiggelbout, A. M., & van Hilten, J. J. (2006). Reliability and validity of the Beck depression inventory in patients with Parkinson's disease. *Movement disorders*, *21*(5), 668-672. https://doi.org/10.1002/mds.20792

Voss, T., Bahr, D., Cummings, J., Mills, R., Ravina, B., & Williams, H. (2013). Performance of a shortened Scale for Assessment of Positive Symptoms for Parkinson's disease psychosis. *Parkinsonism & Related Disorders*, *19*(3), 295-299. https://doi.org/10.1016/j.parkreldis.2012.10.022

Wechsler, D. (2000). Wechsler adult intelligence scale - third edition. Nederlandstalige bewerking. Pearson Assessment and Information B.V.

Weeland, C. J., Vriend, C., van der Werf, Y., Huyser, C., Hillegers, M., Tiemeier, H., White, T., & van den Heuvel, O. A. (2021). Thalamic Subregions and Obsessive-Compulsive Symptoms in

2,500 Children From the General Population. *Journal of the American Academy of Child & Adolescent Psychiatry*. Advance online publication. https://doi.org/10.1016/j.jaac.2021.05.024

Weintraub, D., Mamikonyan, E., Papay, K., Shea, J. A., Xie, S. X., & Siderowf, A. (2012). Questionnaire for impulsive-compulsive disorders in Parkinson's Disease–Rating Scale. *Movement Disorders, 27*(2), 242-247. https://doi.org/10.1002/mds.24023

World Medical Association (2013). Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA, 310*, 2191-2194. https://doi:10.1001/jama.2013.281053

Wöstmann, N. M., Aichert, D. S., Costa, A., Rubia, K., Möller, H. J., & Ettinger, U. (2013). Reliability and plasticity of response inhibition and interference control. *Brain and cognition, 81*(1), 82-94. https://doi.org/10.1016/j.bandc.2012.09.010

Zhang, Y., Brady, M., & Smith, S. (2011). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging, 20*(1), 45-57. https://doi.org/10.1109/42.906424

Appendix A

Inclusion and exclusion criteria of the COGTIPS study

Inclusion criteria	**Measurement**	**Defined by**
Significant subjective cognitive complaints	Parkinson´s disease Cognitive Functional Rating Scale (Kulisevsky et al., 2013)	Score>3
Mild to moderate disease stage	Hoehn & Yahr disease stage (Hoehn & Yahr, 1998)	Score<4
Access to computer or tablet with access to internet.	Phone interview	-
Capability to use keyboard and computer mouse		
Signed informed consent	-	-
General exclusion criteria	**Measurement**	**Defined by**
Indication for dementia syndrome	Self-administered Gerocognitive Examination (Scharre et a., 2010)	Score<14
	Montreal Cognitive Assessment (Nasreddine et al., 2005)	Score <22
Current drug or alcohol abuse	CAGE-AID interview (Ewing, 1984)	Score<1
Inability to undergo extensive neuropsychological assessments or eight weeks of home-based cognitive intervention	-	-
Moderate to severe depressive symptoms	Beck depression inventory (Beck et al., 1961)	Score>18
Presence of one or more impulse control disorders	ICD criteria interview	Positive screening
Psychotic symptoms. Benign hallucinations with insight are not exclusion criterion	Schedule for Assessment of Positive Symptoms – Parkinson´s disease (Voss et al., 2013)	Positive screening
Traumatic brain injury	Phone interview	Cerebral contusion with (1) loss of consciousness for >15 minutes and (2) posttraumatic amnesia >1 hour
Exclusion criteria for magnetic resonance imaging	**Measurement**	**Defined by**
A space occupying lesion	Assessment by radiologist	-
Significant vascular abnormalities	Assessment by radiologist	Fazekas>1
Severe claustrophobia	MRI safe screening questionnaire	Positive screening
Presence of metal in the body	-	-
Pregnancy	-	-
Difficulty with or shortness of breath during 60 minutes of lying still	-	-

Appendix B

Cognitive measures employed in the COGTIPS study

Test name	Author (year)	Construct	Outcome	Worse performance
Boston Naming test	Kaplan et al. (2001)	Confrontational naming	Total number of correct responses	Low score
Category and Letter Fluency tests	Schmand et al. (2008)	Semantic and phonemic memory and strategic search	Total number of correct responses	Low score
Computerized adaptation of Tower of London	Trujillo et al. (2015)	Planning and problem-solving	Total number of solved correct trials (accuracy score) and average reaction time	Low score in total number of correct trials and high score in reaction time
Location Learning test	Kessels et al. (2014)	Visuospatial memory (STM and LTM)	Total number of errors in immediate and delayed recall trials	High score
Pentagon Copy from Mini-Mental State Examination	Folstein et al. (1975)	Visuospatial abilities	Correct drawing	Low score
Rey Auditory Verbal Learning test	Saan and Deelman (1986)	Verbal memory (STM & LTM)	Total number of correct responses in immediate and delayed recall trials	Low score
Rey Complex Figure test	Meyers and Meyers (1995)	Visuospatial memory (STM & LTM)	Total score in copy, immediate and delayed recall	Low score
Stroop Colour-Word test	Hammes (1971)	Sensitivity to interference	Total time in card III - total time in card II (interference score)	High score
Visual Form Discrimination test	Benton et al. (1994)	Visuospatial abilities	Total score	Low score
Wechsler Adult Intelligence Scale III – Digit Span	Weschler (2000)	Working memory	Total score of forward, backward and sorting trials	Low score

Note. STM=short term memory; LTM=long term memory.